Essays in Firm Productivity:

Trade, Geography and Recession

NICK JACOB

Submitted for the degree of Doctor of Philosophy in Economics

University of Sussex

June 2020

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

I also hereby declare that Chapter 3 and Chapter 4 are co-authored with Professor Giordano Mion. My contribution to Chapter 3 was significant. I conducted the majority of the data analysis and contributed to the empirical strategy while writing much of the paper. I estimate I completed 40% of the work. In Chapter 4, I cleaned, managed and performed substantially all of the data analysis, contributed to the empirical approach and took the lead in writing the paper. I estimate I completed 70% of the work. I sole-authored Chapter 2.

Signature:

Nick Jacob

UNIVERSITY OF SUSSEX

Nick Jacob

Doctor of Philosophy in Economics

Essays in firm productivity: Trade, geography and recession

Summary

This thesis answers three research questions in the field of firm-level productivity, drawing on recent empirical frameworks that allow me to use detailed production datasets to simultaneously estimate firm heterogeneities in productivity, consumer demand and price-cost markups.

Trade Reform Redux: Prices, Markups and Product Quality revisits the widely-studied Indian tariff liberalization that began in 1991 to show that rising product price-cost markups, documented in the recent literature are linked to quality upgrading responses to lower tariffs on imported intermediates.

In *On The Productivity Advantage of Cities*, a co-authored paper, we shed new light on the nature of agglomeration externalities for which most theory focuses on differences in technical ability of firms across space, but for which most evidence relies on productivity measures that may be biased by differences in prices across space. We provide new evidence from France that shows little heterogeneity in the technical efficiency -– the ability to turn a basket of inputs into physical outputs, or what we call quantity TFP -– displayed by firms in dense and less-dense areas, but some heterogeneity in the prices firms are able to charge, as well as a positive correlation between efficient allocation of resources within regions and the density of regions.

And in *The UK's Great Demand Recession*, a co-authored paper, we estimate changes in revenue total factor productivity, i.e., TFP estimated using deflated sales data; quantity TFP; consumer demand and other measures for UK firms before, during and after the Great Recession. Our results show weakness in quantity TFP and demand pushed down productivity in manufacturing and services.

This thesis contributes to the existing literature in its use of these datasets and methods to provide new understanding of firms' responses to a range of economic forces, with implications for both theory and policy in all three settings.

Acknowledgements

I am indebted to my advisor Giordano Mion who has – always coolly and calmly – guided me through the process. I also thank other faculty of the Economics Department at Sussex for helpful comments and the opportunities to discuss my work, including Peter Holmes, Rob Eastwood, Andy Newell and Barry Reilly.

The ever-changing cast of the economics PhD community at Sussex has been a wonderful support, and a group from which I have learnt so much. In particular, I would like to thank Eugenia Go, Antonia Schwarz and Marta Schoch for their, usually successful, efforts in cheering me up during the more difficult moments.

I have received helpful comments from seminar participants including Andy McKay, Richard Disney, Ingo Borchert, Richard Tol, Cecila Poggi, Eva-Maria Egger, Marta Schoch, Antonia Schwarz, Edgar Salgado Chavez and Hector Gutierrez Rufrancos at the University of Sussex, and Stephen Millard and other seminar participants at the Bank of England.

I also thank the Centre for Economic Performance at the London School of Economics for providing access to the CASD data in a convivial but secure environment, and the UK Secure Data Service for their patience in dealing with enquiries.

This thesis was supported by an Economic and Social Sciences Research Council (ESRC) studentship in the University of Sussex's ESRC Doctoral Training Centre. In addition, the Economics department at Sussex was generous in funding the acquisition of the Prowess data.

Chapter 3 is supported by a public grant overseen by the French National Research Agency (ANR) as part of the 'Investissments d'Avenir' program (reference: ANR-10-EQPX-17-Centre d'accès sécurisé aux données – CASD).

Dedication

This thesis is for Lucy, whose love and support has made it, and so much more, possible.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis consists of three substantive chapters, each concerned with what might be thought of as distinct, but long-standing, questions in economic research: the reasons why denser regions appear to be more productive than less dense regions, what happens to firms when tariffs on international trade are reduced, and what explains the productivity growth slowdown following the financial crisis of 2008/9. What links these questions, and what is novel about my research, are the tools and data I use to answer them. I use three different firm-product level datasets that provide both price and quantity information on firms' sales and production and apply recent empirical frameworks that allow me to simultaneously estimate firm-product level productivity, markups and product demand/quality. This approach contributes to the existing literature in its use of these datasets and methods by providing new understanding of firms' responses to a range of economic forces, with implications for both theory and policy.

My work draws heavily from two key papers: De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) (henceforth DGKP) and Forlani, Martin, Mion and Muûls (2016) (henceforth FMMM). These papers are related to the literature on heterogeneous markups and productivity, notably Hall (1986) and Olley and Pakes (1996) and developed in Ackerberg et al. (2015). DGKP and FMMM are similar in their motivation to unravel heterogeneity in revenue TFP (total factor productivity estimated using deflated sales, or what we call TFP-R). TFP-R is a mixture of supply-side differences between firms, notably physical productivity (what we call quantity total factor productivity, or TFP-Q), and demand-side differences in prices which could be due to differences in input and/or output quality, demand and markups.

DGKP builds on the proxy variable approach to estimate a production function allowing for heterogeneity in TFP-Q as well as heterogeneity in demand and markups by using output prices in a control function for unobserved input prices. FMMM provides an alternative estimation framework that also allows firm heterogeneity in TFP-Q, demand and markups by exploiting a revenue equation that is explicit about the type of demand faced by the firm, as well as a quantity equation. And because it is explicit about the nature of demand, FMMM allows construction of demand measures which are consistent with several underlying consumers' preferences (representative consumer and discrete choice models) and market structures (monopolistic competition, monopoly and standard forms of oligopoly). Crucially, these demand measures do not depend on the choice of estimation technique and so can also be recovered using the DGKP estimation approach.

Demand heterogeneity in FMMM has a direct interpretation from a representative consumer's utility function as its relative willingness-to-pay for one variety over another, i.e., a measure of 'perceived quality'. Therefore, demand heterogeneity in FMMM captures both actual quality differences, i.e., differences in the physical attributes of products that are reflected in consumers' willingness-to-pay, and 'appeal' differences, i.e., the capacity of certain firms, through marketing, distribution channels and branding, to sell more than others despite their products being similar to those of their competitors. It thus encompasses notions of both product quality (my focus in Chapters 2 and 3) and product demand (my focus in Chapter 4).

I borrow these approaches and apply them in the three different settings. *Trade Reform Redux: Prices, Markups and Product Quality* (Chapter 2) revisits the early 1990s Indian tariff liberalization episode studied in DGKP to show that rising product price-cost markups are linked to manufacturing firms' quality upgrading responses to lower tariffs on imported intermediates. I find that conditional on product quality, there was no rise in average firm markups over the 1989-1997 period, i.e., quality upgrading fully explains the rise in markups without any need for market power explanations. Further, I find that there was substantial heterogeneity in firm responses such that all of the unconditional rise in markups came from 'laggard' varieties, i.e., those that started the period with the lowest sales, total factor productivity and markups.

In *On The Productivity Advantage of Cities* (Chapter 3), we estimate TFP-R for French manufacturing firms using both the DGKP and FMMM techniques and use the decomposition provided by FMMM to disentangle TFP-Q, quality/demand, markups and production scale. Our analysis suggests that the TFP-R advantage of denser areas is mainly driven by higher prices charged rather than differences in TFP-Q. At the same time, firms in denser areas sell higher quantities, and generate higher revenues, despite higher prices. These and other results we document suggest that firms in denser areas are able to charge higher prices because they sell higher demand/quality products. Finally, while the correlation between firm revenue TFP and firm size is positive in each location, it is also systematically related to density: firms with higher (lower) TFP-R account for a larger (smaller) share of total revenue in denser areas. These patterns thus amplify in aggregate regional-level figures any firm-level differences in productivity across space.

And in *The UK's Great Demand Recession* (Chapter 4) we turn the focus to UK firms' poor productivity performance since the 2008/9 financial crisis, again using the methods of DGKP and FMMM to both measure demand and its changes over time and distinguish between TFP-Q and TFP-R. This in turn allows us to measure how changes in TFP-Q, demand and markups ultimately affected revenue TFP, as well as labour productivity, over the Great Recession. Our findings suggest that UK firms' poor productivity performance post-recession is due to both a weakening of demand and TFP-Q pushing down sales, markups, revenue TFP and labour productivity.

The results suggest important implications for policy. In both Chapters 3 and 4, for instance, we aggregate firm-level measures to regional and national levels and show that policy based on revenue productivity alone, without an understanding of its underlying price and quantity productivity components, can be ill-directed. In particular, for the UK we highlight that a large part, though by no means all, of the UK productivity 'puzzle' can be explained by poor post-shock demand. For France we show that regional disparities in productivity are caused to an extent by product composition effects and product quality differences between firms from high-density areas compared to low-density areas, as well as an inferior allocation of resources between firms within low-density areas compared to high-density areas. The conventional approach that diagnosed low revenue TFP would recommend policy be focused on measures to boost the technical efficiency of firms (whether nationally, or in lagging regions). Our results, however, stress the importance of marketing, connectedness and other attributes which affect the prices that firms can charge (in the French setting), and the overall level of demand in the economy (in the UK setting). Likewise, the finding in Chapter 2 that markups on average rise following a reduction in tariffs, might imply policy makers should be concerned with ameliorating firms' market power. But if the rising markups are associated with an increase in product quality, such that when taking quality/product appeal into account cost changes are passed through to prices, as I show, then consumers did indeed benefit from the tariff liberalisation.

While all of the chapters in this thesis share a common approach to estimating a quantity-based production function, there are operational differences in the implementation, and so, at the risk of repetition, I set out in detail – in either the main text or the accompanying appendices – the full framework used in each. To give an overview here, in Chapter 2 my goal is to extend the analysis of DGKP and so in order to first replicate the key results from that paper I make use of the code made available online for that purpose. This approach considers a Translog production function which has the benefit of ensuring that variation in markups across and between firms is not only due to variation in the share of intermediate inputs among firms. Our benchmark results in Chapters 3 and 4 on the other hand consider Cobb-Douglas production functions, simplifying estimation and reducing the number of parameters to estimate (although we do provide results for the Translog as robustness checks). For both these settings, we also use the alternative estimation procedure set out in FMMM. Additionally, for the UK we estimate revenue productivity using a restricted version of the FMMM framework on firms in the services industries. In the France setting, as in that of India, because of the institutional frameworks, we consider it appropriate to treat labour as a semi-flexible input, while in the UK we treat it as a fully-flexible input.

DGKP and FMMM both provide a treatment of multi-product firms and I make use of both. In Chapter 2 I follow DGKP which, after estimating production function parameters from single-product firms, assigns firm-level inputs to product-level outputs by numerically solving a set of equations with an assumption that productivity is constant across products within a firm. The FMMM approach to this problem is to allow productivity to vary across products, but assumes that demand/quality is constant within firms.

Finally, there are some operational differences driven by data constraints. In Chapter 2, I reassemble the raw production data used by DGKP from a commercially-available database of large and medium-sized manufacturing firms, and show that my dataset closely replicates the key results of the original paper. In this, I am able to follow DGKP and estimate production functions separately for each 2-digit National Industrial Classification industry. Our French data is the most comprehensive as we have access to full balance sheet information for the population of firms in metropolitan France, and this data matches well to the production data, allowing us to perform production function estimations at the 2-digit SIC industry level. The UK, however, only carries out a limited survey of firms – which also forces us to estimate firm-level capital stocks from investment data – and once we match this data with our production data we are left with too few firms to run separate production function estimations, so instead group the data in a single estimation.

Chapter 2

Trade Reform Redux: Prices, Markups and Product Quality

Abstract

This paper revisits the widely-studied Indian tariff liberalization that began in 1991 to show that the rising markups documented in recent work by De Loecker et al. (2016) (DGKP) are linked to quality upgrading responses to lower tariffs on imported intermediates. To do this, I reassemble the detailed product-level dataset used by DGKP and closely follow their approach to estimating marginal costs and markups, while drawing on Forlani et al. (2016) to obtain measures of product quality/appeal consistent with several underlying types of consumer preferences and market structures. In doing so, I find that the markups rise documented in DGKP is entirely driven by an increase in product quality rather than an increase in marker power, i.e., conditional on product quality, there was no rise in average firm markups over the 1989-1997 period. Furthermore, there was substantial heterogeneity in firm responses such that all of the unconditional rise in markups came from 'laggard' firms, i.e., those that started the period with the lowest sales, total factor productivity and markups. These results indicate that the tariff liberalization benefited Indian consumers more than what a simple reading of the rising markups result would suggest.

2.1 Introduction

In 1991 India began a programme of substantially liberalising its economy in response to a severe balance of payments crisis. It slashed external tariffs and relaxed its regimes for foreign direct investment and industrial licencing requirements while also undertaking macroeconomic reforms such as reducing the fiscal deficit, devaluing the rupee and liberalising capital markets (Sivadasan, 2009). The reforms have been associated with a sharp rise in imports, particularly of intermediate goods (Goldberg et al., 2009), higher aggregate and within-firm productivity (Harrison et al., 2013; Topalova and Khandelwal, 2011), the introduction of new product varieties (Goldberg et al., 2010a) and the exit of the least productive informal firms (Nataraj, 2011), among other findings. India's pre-reform tariff wall was exceptionally high, with rates averaging 101%, and it varied in height substantially such that the across-the-board cuts that brought average rates down to 32.7% by 1997 generated substantial variation in the impact of the reform across industries. Researchers have typically treated this variation as exogenous to firm and industry characteristics and used it to estimate causal impacts of lower tariffs on various aspects of firm performance. Central to this paper are the results presented by De Loecker et al. (2016) (henceforth DGKP) showing that the trade reform spurred increases in average product markups.

While much of the theory (Bernard et al., 2003; Melitz and Ottaviano, 2008) predicts that by lowering barriers to trade, firms will be forced down their cost curves and will lower prices and markups, DGKP shows empirically that these pro-competitive effects of lower tariffs in output markets can be swamped by the marginal cost-reducing effects of lower tariffs on intermediate inputs which are not passed through to prices. A rise in markups is usually attributed to a less competitive market or institutional rigidities and is therefore welfare-reducing. However, if the rise in markups comes about through a rise in product quality/appeal and concomitant consumer willingness-to-pay then it would be of less concern to policymakers.

In this paper I answer the question, did Indian manufacturing firms' markups rise following the trade reform because firms raised product quality and/or because of an increase in market power? Theoretically and empirically, a rise in product quality, has been shown to be a consequence of more open trade in developing countries (Goldberg et al., 2009;

Verhoogen, 2008; Kugler and Verhoogen, 2011; Antoniades, 2015), at least for some firms, but the link to the rising markups documented by DGKP is less well understood. To answer my research question, I reassemble the raw production data on Indian manufacturing firms used by DGKP and run the same procedures to estimate marginal costs and markups. I then use the insights of Forlani et al. (2016), henceforth FMMM, to construct measures of product quality/appeal consistent with several underlying types of consumer preferences and market structures.

The first contribution of my paper is to a literature that theoretically and empirically documents the effects of increasing trade openness on product quality. I show that quality upgrading was an important factor in the rise of markups during the Indian trade reform documented by DGKP. Indeed, when controlling for product quality/appeal, the trade reform did not have any impact on average firm-product markups. Key to explaining this result is that the decline in marginal costs via lower input tariffs observed in DGKP is actually associated with a rise in output quality: higher-quality foreign inputs enabled firms to produce higher-quality outputs at lower marginal costs while maintaining high prices. This finding complements existing empirical studies. Goldberg et al. (2009) provides descriptive evidence that product quality rose in India over 1987-2000, documenting fast growing imports of intermediates which were both rising in unit values and variety, indicating that technological restraints were being relaxed and higher quality inputs being used. For Mexico in the early 1990s, Verhoogen (2008) shows that exporters upgraded quality in response to a devaluation.

The importance of lower tariffs on imported inputs to firm productivity and product quality upgrading has been stressed in many studies. Notably, Amiti and Konings (2007) estimates that the productivity gain for importing firms in Indonesia from a reduction of input tariffs is more than the double the gain from the equivalent reduction in tariffs on output.[1] Topalova (2010) provide similar evidence for the Indian trade reform and Goldberg et al. (2010b) provides evidence that lower input tariffs increased the rate of new product entry, as well as TFP and research and development investment.

Standard heterogeneous firms trade models à la Melitz predict a negative correlation between prices and firm size, but empirically this is only true in some industries, a fact

[1] Input tariffs are usually defined as the weighted average tariff on imports of intermediate inputs, with the weights derived from detailed industry level input-output tables, see Section 2.4.1.

explained by differences in the scope for quality differentiation (Khandelwal, 2010; Verhoogen, 2008; Kugler and Verhoogen, 2011). Models that allow firm heterogeneity in both productivity and product quality, such as Antoniades (2015), feature the usual pro-competitive effects of lower tariffs as more firms enter the market, the least productive firms exit and surviving firms lower prices and markups. But by also modelling endogenous product quality, they also predict that trade openness raises returns to innovation, leading to higher quality, prices and markups.

The second contribution of this paper is to document the extent of this heterogeneity in the impacts of the trade reform across firms. More specifically, I show that the impacts for firm-products that begin the period as low-performing versus high-performing are mediated through different channels. The initially most productive firm-products do not adjust markups in response to greater competition from imports, but do make efficiency improvements that are passed through to prices. Firm-products that are initially poor performers, however, respond to lower input tariffs by increasing output quality, allowing them to increase quantities sold and reduce marginal costs, without a concomitant lowering of prices, so leading to sharply higher markups.

These findings are in line with some of the literature. For example, evidence from a developing country that it is the least productive firms that respond most strongly to tariff reduction comes from Fan et al. (2015, 2018). Studying the effects of China's entry into the World Trade Organization, Fan et al. (2015, 2018) find that lagging exporting firms upgraded output and inputs quality and prices faster than pre-reform leaders. Key to these results, as highlighted in their theoretical model, is the presence of scale effects and non-Hicks neutral productivity. However, many other researchers reach the opposite conclusion. For instance, in the Antoniades (2015) model it is the most productive firms that will respond to competition by raising quality, and the least productive by reducing it. At the same time, Verhoogen (2008) highlights the importance of scale effects, such that larger, more productive exporters find it more profitable to cover the fixed costs of quality upgrading in response to better access to foreign markets. On the empirical side, Amiti and Khandelwal (2013) shows that greater competition in markets due to lower tariffs have a quality-raising effect only for products close to the world technological frontier, with the opposite effect for products far away from it. However, for countries that rank lowly for institutional quality in the World Bank Doing Business surveys, there is no distance-from-frontier effect.

To tackle these questions, I build on the analysis of DGKP. First, I reassemble the finely-granular production data, an unbalanced panel of medium and large-sized Indian manufacturing firms that contains detailed information on product-level sales and production quantity. I then follow, as precisely as possible, the estimation procedures used by DGKP to obtain marginal costs and markups at the firm-product-year level for the period 1989-1997, and show that my dataset closely replicates the key results of the original paper. The DGKP procedures, which I describe in detail in Section 2.2 allow for estimating a quantity-based production function and recovering key product-level measures for multi-product firms.

The reason why DGKP showed only that markups rose but not why, is because testing demand-side features, such as increases in quality and/or product appeal, requires explicitly modelling demand which involves laying down a number of strong assumptions about preferences and market structure, such as in Khandelwal (2010), Amiti and Khandelwal (2013) and Feenstra (2018), or else relies on choosing a value for a (constant) elasticity of substitution between products, sometimes using the values for the United States from Broda and Weinstein (2006) as in Stiebale and Vencappa (2018).

I however, build upon FMMM to construct measures of the demand faced by firms for their products which are consistent with several underlying types of consumer preferences (representative consumer and discrete choice models) and market structures (monopolistic competition, monopoly and standard forms of oligopoly) and analyze what happens to these demand measures when Indian firms face lower tariffs in output and inputs markets.[2] Demand heterogeneity in FMMM has a direct interpretation from a representative consumer's utility function as its relative willingness-to-pay for one variety over another, i.e., a measure of 'perceived quality'. Therefore, demand heterogeneity in FMMM captures both actual quality differences, i.e., differences in the physical attributes of products that are reflected in consumers' willingness-to-pay, and 'appeal' differences, i.e., the capacity

[2] A paper with a similar approach and setting, but a different research question to mine (mergers and acquisitions), is Stiebale and Vencappa (2018). Stiebale and Vencappa (2018) uses the multi-product DGKP quantity production function estimation procedure to obtain firm-product level markups and the FMMM approach to directly obtain estimates of product quality/appeal from the estimated markups and data, as in this paper.

of certain firms, through marketing, distribution channels and branding, to sell more than others despite their products being similar to those of their competitors.

The identification strategy employed in DGKP and this paper relies on tariff changes over the period being exogenous to firms' pricing and production decisions. Topalova and Khandelwal (2011) provided evidence that this is the case, for the period up to the Indian election in 1997, but not thereafter, so I restrict analysis to this period. The tariff liberalisation was accompanied by a number of other reforms to dismantle what was known as the license Raj and it could be argued that the empirical strategy needs to explicitly recognise the correlations between reforms. DGKP show convincingly that controlling for industry-level de-licensing reforms does not qualitatively change their results and so I maintain a focus on tariff changes.

The remainder of this paper is organised as follows. Section 2.2 sets out the DGKP estimation procedure for marginal costs and markups and Section 2.3 the FMMM approach for calculating measures of product quality. Section 2.4 describes the data I use and provides a comparison to DGKP. Section 2.5 reports results and Section 2.6 concludes.

2.2 DGKP model and implementation

2.2.1 Production function estimation

DGKP explain the impact of trade reform on markups, using price and quantity data from multi-product firms. To do this, they develop a framework that builds on the simple rule – first highlighted in Hall (1986) and implemented in De Loecker and Warzynski (2012) among others – that markups can be pinned down by taking the ratio of the output elasticity of materials to the share of materials in revenue. Developing this framework is the main contribution of DGKP, allowing a way to get at the output elasticity from a quantity-based production function, and at the materials share in revenues from a multi-product firm setting, in order to study the effects of trade policy on markups.

The approach extends the proxy variable (or control function) approach to estimating the parameters of a production function due to Olley and Pakes (1996) and Levinsohn and Petrin (2003). These earlier approaches attempted to deal with the endogeneity problem inherent in measuring TFP as the residual from a regression of the log of output on the log of an index of inputs that arises because firms observe their own TFP when making input choices (the simultaneity bias). Olley and Pakes (1996) uses investment expenditure and Levinsohn and Petrin (2003) uses materials expenditure to control for unobserved productivity by inverting the firm's demand function (a monotonic function of the unobserved productivity) for the input. However, researchers typically use firm revenue, rather than quantity, data for estimation, biasing production function coefficients while also conflating productivity and marginal costs estimates with price and demand variation across firms. To deal with this issue, DGKP start by writing a (log) quantity production function for firm f producing product j at time t:

$$q_{fjt} = f_j(\boldsymbol{x}_{fjt}; \boldsymbol{\beta}) + \omega_{ft} + \epsilon_{fjt}, \tag{2.1}$$

where $\boldsymbol{x}_{fjt}$ is a vector of log physical inputs, consisting of labour, materials and capital (l_{fjt}, m_{fjt}, k_{fjt}), $\boldsymbol{\beta}$ is the vector of production function coefficients, ω_{ft} is TFP[3] and ϵ_{fjt} is an error term.

Writing the production function in terms of physical output and inputs creates two challenges not faced in the usual Olley and Pakes (1996) and Levinsohn and Petrin (2003) approaches to estimation: input quantities are rarely observed, only input expenditures (as in my data); and these expenditures are at the firm level and not attributed to products within multi-product firms. Each firm-product-specific input (labour, materials, capital) is defined as:

$$x_{fjt} = \rho_{fjt} + \tilde{x}_{ft} - w^x_{fjt}, \tag{2.2}$$

where $\rho_{fjt} = \ln \tilde{\rho_{fjt}}$ is the share of firm input expenditure allocated to product j and w^x_{fjt} is the firm-product-specific input price. Writing $\boldsymbol{w_{fjt}}$ for the vector of firm-product-specific input prices, and substituting in Equation (2.1), DGKP obtain:

$$q_{fjt} = f_j(\tilde{\boldsymbol{x}}_{ft}; \boldsymbol{\beta}) + A(\rho_{fjt}, \tilde{\boldsymbol{x}}_{ft}, \boldsymbol{\beta}) + B(\boldsymbol{w}_{fjt}, \rho_{fjt}, \tilde{\boldsymbol{x}}_{fjt}, \boldsymbol{\beta}) + \omega_{ft} + \epsilon_{fjt}, \tag{2.3}$$

Equation (2.3) highlights the two challenges. The term $A(.)$ represents an input allocation bias, and the term $B(.)$ an input price bias since both are functions of the input expenditures.

[3]Note that TFP in the DGKP model is firm-time specific, while in the multi-product extension of FMMM it firm-product-time specific. I return to this in Section 2.3 below.

2.2.2 Input price bias and single-product firm estimation

Making the assumption that the production technology is product-specific, it is clear that when considering only single-product firms, the term $A(.) = 0$ since $\tilde{\rho}_{fjt} = 1$ and so the production function can be estimated using only single-product firms and then in a later stage used to assign inputs within multi-product firms. Equation (2.3) considering only single-product firms simplifies to:

$$q_{ft} = f(\tilde{\boldsymbol{x}}_{ft}; \boldsymbol{\beta}) + B(\boldsymbol{w}_{ft}, \tilde{\boldsymbol{x}}_{ft}, \boldsymbol{\beta}) + \omega_{ft} + \epsilon_{ft}, \tag{2.4}$$

Key to DGKP's approach to dealing with the $B(.)$ term is to use a control function for unobserved input prices. The insight here is to note that higher input prices imply higher quality inputs creating higher quality outputs, as in Verhoogen (2008), allowing DGKP to write firm-specific input prices as a function of output prices as well as market shares, product dummies, exporter status and tariffs.

$$w_{it}^{x} = w_t \left(p_{it}, ms_{it}, \mathbf{D}_f, EXP_{it} \right). \tag{2.5}$$

Substituting into Equation (2.4):

$$B(\boldsymbol{w}_{it}, \tilde{\boldsymbol{x}}_{it}, \boldsymbol{\beta}) = B((p_{it}, ms_{it}, EXP_{it}) \times (1, \tilde{\boldsymbol{x}}_{it}), \boldsymbol{\beta}, \boldsymbol{\delta}), \tag{2.6}$$

where $\boldsymbol{\delta}$ is an additional coefficient vector representing the parameters of the $w(.)$ function. These coefficients will be retrieved, allowing an estimate of firm-specific input prices.

The usual Levinsohn and Petrin (2003) control function is used for unobserved productivity by inverting the materials demand equation given by:

$$\tilde{m}_{ft} = m_t \left(\omega_{ft}, \tilde{k}_{ft}, \tilde{l}_{ft}, p_{ft}, \mathbf{D}_f, \mathbf{ms}_{ft}, EXP_{ft}, \tau_{ft}^{input}, \tau_{ft}^{output} \right) \tag{2.7}$$

so that,

$$\omega_{ft} = h_t(\tilde{\mathbf{x}}_{ft}, \boldsymbol{Z}_{ft}), \tag{2.8}$$

where $\boldsymbol{Z}_{ft}$ contains all the variables that determining input demand except input expenditures and productivity, i.e. $\boldsymbol{Z}_{ft} = \{p_{ft}, \mathbf{D}_f, \mathbf{ms}_{ft}, EXP_{ft}, \tau_{ft}^{input}, \tau_{ft}^{output}\}$

A potential source of bias in using only single-product firms for the estimation comes from productivity or inputs use affecting firms' choices to add additional products. DGKP control for this in much the same way as Olley and Pakes (1996) control for selection into single-plant vs. multi-plant establishments, that is non-parametrically estimating the probability of a firm being single-product in $t+1$ using its information set in t, i.e. a polynomial of inputs and interactions of inputs, tariffs, market share, a polynomial of prices and year dummies. Using Equations (2.6), and (2.8), DGKP write (2.4) as,

$$q_{ft} = \phi_t(\tilde{\mathbf{x}}_{ft}, \mathbf{Z}_{ft}) + \epsilon_{ft}, \tag{2.9}$$

in order to estimate $\hat{\phi}_{ft}$, i.e. output quantity purged of measurement error and unanticipated shocks. Equations (2.4), (2.6) and (2.9) then give,

$$\omega_{ft}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right) = \hat{\phi}_{ft} - f\left(\tilde{\boldsymbol{x}}_{ft}; \boldsymbol{\beta}\right) - B\left((p_{ft}, \boldsymbol{ms}_{ft}, \mathbf{D}_f, EXP_{ft)}) \times (1, \tilde{\mathbf{x}}_{ft}); \boldsymbol{\delta}\right). \tag{2.10}$$

Estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ can then follow by assuming that productivity follows a law of motion given by

$$\omega_{ft} = g\left(\omega_{ft-1}, \tau^{input}_{ft-1}, \tau^{output}_{ft-1}, EXP_{ft-1}, SP_{ft}\right) + \xi_{ft}, \tag{2.11}$$

where SP_{ft} is the estimated probability that a firm remains single-product in the next period. The law of motion for productivity given the parameters, $\xi_{ft}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right)$ allows construction of moment conditions,

$$E\left(\xi_{ft}(\boldsymbol{\beta}, \boldsymbol{\delta})\mathbf{Y}_{ft}\right) = 0, \tag{2.12}$$

where $\mathbf{Y}_{ft}$ contains labour and capital along with lagged materials and interaction terms, as well as lagged market shares, tariffs and interactions with the inputs, and this is estimated using the GMM method due to Wooldridge (2009).

2.2.3 Input allocations and multi-product firms

After estimating the $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ coefficients using the single-product firms, DGKP write the production function as,

$$\hat{q}_{fjt} = f(\tilde{\mathbf{x}}_{ft}, \hat{\boldsymbol{\beta}}, \hat{w}_{fjt}, \rho_{fjt}) + \omega_{ft}, \tag{2.13}$$

and note that it can be split into terms that contain the input allocations ρ_{fjt} and those that don't

$$\hat{q}_{fjt} - f_1(\tilde{\mathbf{x}}_{ft}, \hat{\boldsymbol{\beta}}, \hat{w}_{fjt}) = f_2(\tilde{\mathbf{x}}_{ft}, \hat{\boldsymbol{\beta}}, \hat{w}_{fjt}, \rho_{fjt}) + \omega_{ft}. \tag{2.14}$$

Collecting all terms formed of estimates and data on the left hand side, there are now J_{ft} equations for each firm. Together with the restriction that the input allocation shares have to add up to 1 across products within firms,

$$\sum_j exp(\rho_{fjt}) = 1, \tag{2.15}$$

this implies that the J_{ft} input allocations for each firm and ω_{ft} can be recovered from a system of $J_{ft} + 1$ equations which are solved numerically. At this stage it is worth noting that recovering these input allocations in this way requires the assumption that TFP is specific to firms across products, ω_{ft} while in parts of the analysis I use a firm-product specific measure, $a_{fjt} = \hat{q}_{fjt} - \bar{q}_{ft}$.

2.2.4 Markups

Having estimated the parameters of the production function and input shares across products within firms, it remains only to use data and estimates to recover markups and marginal costs. This is straightforward following the markup rule due to Hall (1986) and used in De Loecker and Warzynski (2012) that markups can be pinned down by taking the ratio of the output elasticity of materials to the share of materials in revenue. This rule to pin-down markups is consistent with many hypotheses on product market structure (monopolistic competition, monopoly and standard forms of oligopoly). The proof is as follows.

Firms minimise cost of a variable input free of adjustment costs and exhibit no market power in input markets (the cost of materials W_{Mit} is allowed to be firm-time specific but it is given to the firm). We can write marginal cost as:

$$\frac{\partial C_{fjt}}{\partial Q_{fjt}} = \frac{\partial C_{fjt}}{\partial M_{fjt}}\frac{\partial M_{fjt}}{\partial Q_{fjt}} = W_{Mfjt}\frac{\partial M_{fjt}}{\partial Q_{fjt}}. \tag{2.16}$$

Markups over marginal cost are defined as:

$$\mu_{fjt} \equiv \frac{P_{fjt}}{\frac{\partial C_{fjt}}{\partial Q_{fjt}}}. \tag{2.17}$$

Combining, we have:

$$\frac{P_{fjt}}{\mu_{fjt}} = W_{Mfjt}\frac{\partial M_{fjt}}{\partial Q_{fjt}}.$$

Multiplying by Q_{fjt} and dividing by M_{fjt} on both sides we get:

$$\frac{P_{fjt}Q_{fjt}}{M_{fjt}\mu_{fjt}} = \frac{R_{fjt}}{M_{fjt}\mu_{fjt}} = W_{Mfjt}\frac{\partial M_{fjt}}{\partial Q_{fjt}}\frac{Q_{fjt}}{M_{fjt}} = W_{Mfjt}\frac{\partial m_{fjt}}{\partial q_{fjt}}.$$

Re-arranging we finally have:

$$\mu_{fjt} = \frac{\frac{\partial q_{fjt}}{\partial m_{fjt}}}{\frac{W_{Mfjt}M_{fjt}}{R_{fjt}}} = \frac{\frac{\partial q_{fjt}}{\partial m_{fjt}}}{s_{Mfjt}}. \tag{2.18}$$

Having estimated the Translog production function:

$$\begin{aligned} q_{ft} \quad = \quad & \beta_l l_{ft} + \beta_{ll} l_{ft}^2 + \beta_m m_{ft} + \beta_{mm} m_{ft}^2 + \beta_k k_{ft} + \beta_{kk} k_{ft}^2 + \\ & \beta_{lm} l_{ft} m_{ft} + \beta_{lk} l_{ft} k_{ft} + \beta_{mk} m_{ft} k_{ft} + \beta_{lmk} l_{ft} m_{ft} k_{ft} + a_{ft}, \end{aligned} \tag{2.19}$$

using the method outlined above, the materials share of revenues in product j (s_{Mfjt})is recovered from Equations (2.14) and (2.15) and then the output elasticity of materials ($\frac{\partial q_{fjt}}{\partial m_{fjt}}$)can be calculated for all firm-products as $\beta_m + 2\beta_{mm}m_{fjt} + \beta_{ml}l_{fjt} + \beta_{mk}k_{fjt} + \beta_{lmk}l_{fjt}k_{fjt}$ because of the assumption that single-product and multi-product firms use the same technology in production.

2.3 FMMM demand structure and assumptions

The DGKP framework described above allows estimation of heterogeneous and variable markups but does not shed light on how firms are able to raise markups during the Indian trade reform since it argues that imposing an Industrial Organisation-style demand system, such as Khandelwal (2010) or Verhoogen (2008) is inappropriate. FMMM, however, provides an approach that is consistent with several underlying types of consumer preferences (representative consumer and discrete choice models) and market structures (monopolistic competition, monopoly and standard forms of oligopoly). This Section provides an outline of the FMMM approach.

2.3.1 Demand heterogeneity

Returning to standard profit maximization (marginal revenue equal to marginal costs), FMMM write the elasticity of revenue R_{ft} with respect to quantity Q_{ft} as one over the profit maximizing markup:

$$\frac{\partial r_{ft}}{\partial q_{ft}} = \frac{\partial R_{ft}}{\partial Q_{ft}} \frac{Q_{ft}}{R_{ft}} = \frac{\partial C_{ft}}{\partial Q_{ft}} \frac{Q_{ft}}{P_{ft} Q_{ft}} = \frac{\frac{\partial C_{ft}}{\partial Q_{ft}}}{P_{ft}} = \frac{1}{\mu_{ft}}, \tag{2.20}$$

Again, this result holds under different assumptions about demand (representative consumer and discrete choice models) and product market structure (monopolistic competition, monopoly and standard forms of oligopoly). The log revenue function, i.e., the function relating log revenue to log quantity, is both unknown and potentially different across firms, but equation (2.20) provides the slope of the firm-specific log revenue function for firm f while data on the actual log revenue r_{ft} and log quantity q_{ft} provide us with a point where such firm-specific log revenue function cuts through the (q, r) space. If we now linearize the log revenue function around the observed data point (q_{ft}, r_{ft}) with a slope given by $\frac{1}{\mu_{ft}}$ we can uniquely pin down an intercept for this linearized log revenue function on the r axis. We use such intercept $\tilde{\lambda}_{ft}$ as our measure of demand heterogeneity:[4]

$$\tilde{\lambda}_{ft} \equiv r_{ft} - \frac{\partial r_{ft}}{\partial q_{ft}} q_{ft} = r_{ft} - \frac{q_{ft}}{\mu_{ft}}. \tag{2.21}$$

Given our definition of $\tilde{\lambda}_{ft}$ observed firm log revenue is simply

$$r_{ft} = \tilde{\lambda}_{ft} + \frac{1}{\mu_{ft}} q_{ft}, \tag{2.22}$$

[4]To simplify notation we ignore components that are constant across firms in a given time period or within a product category. Those constants will be captured in our empirical analysis by a suitable set of dummies.

and so $\tilde{\lambda}_{ft}$ is a firm-specific log revenue shifter[5] corresponding to the log price firm f would face if selling one unit of its product.[6]

While being general and intuitive, this measure of demand heterogeneity also maps to more formal and explicit differences in the underlying structure of preferences. In particular, FMMM show that $\tilde{\lambda}_{ft} = \frac{\lambda_{ft}}{\mu_{ft}}$ where λ_{ft} is a parameter characterizing differences in utility derived from the consumption of products sold by different firms. More specifically, consider a representative consumer who maximises at each point in time t a differentiable utility function $U(.)$ subject to budget B_t:

$$\max_{Q} \left\{ U\left(\tilde{Q}\right) \right\} \text{ s.t. } \int_f P_{ft} Q_{ft} \mathrm{d}f - B_t = 0$$

where $\tilde{Q}$ is a vector of elements $\Lambda_{ft} Q_{ft}$ and $\lambda_{ft} = \log(\Lambda_{ft})$. Therefore, while the representative consumer chooses quantities Q, these quantities enter into the utility function as $\tilde{Q}$ and Λ_{ft} can be interpreted as a measure of the perceived quality/appeal of a particular variety. FMMM show that the log revenue function corresponding to the above preferences $r(q_{ft}, \lambda_{ft})$ can be approximated, around the observed profit-maximizing solution, by the linear function:

$$r_{ft} \simeq \frac{1}{\mu_{ft}} (q_{ft} + \lambda_{ft}), \tag{2.23}$$

and so λ_{ft} is:

$$\lambda_{ft} \simeq \mu_{ft} r_{ft} - q_{ft}. \tag{2.24}$$

Two things are worth nothing at this stage. First, (2.24) is valid as a first-order linear approximation and is the counterpart of (2.21) meaning that the log revenue shifter $\tilde{\lambda}_{ft}$, what FMMM label demand heterogeneity, maps via markups into differences in product appeal across firms' varieties $\lambda_{ft} = \tilde{\lambda}_{ft} \mu_{ft}$. Second, while the shape of the function relating revenue to quantity and product appeal will depend upon the specific underlying preferences, FMMM show that (2.23) applies to any preferences structure that can be used to model monopolistic competition and for which a well-behaved differentiable utility

[5] Demand heterogeneity is the variation in revenue that is not explained by variation in quantities, i.e., two firms selling the same quantity but generating a different revenue (because of a different price). Therefore, demand heterogeneity is a firm-specific log revenue shifter given quantity (or equivalently a firm-specific log price shifter given quantity).

[6] At the intercept point $q_{ft} = 0$ and so we have $Q_{ft} = 1$ from which $R_{ft} = P_{ft}$ and $r_{ft} = p_{ft} = \tilde{\lambda}_{ft}$. Note this has no implications whatsoever about the presence/absence of a choke price.

function exists.[7] This includes standard CES preferences as well as generalized CES preferences (Spence, 1976)[8], CARA preferences (Behrens et al., 2014), HARA preferences (Haltiwanger et al., 2018), Translog preferences (Feenstra, 2003) as well as the class of Variable Elasticity of Substitution (VES) preferences discussed in Zhelobodko et al. (2012) and Dhingra and Morrow (2019).

In order to focus on demand, I rewrite equation (2.22) as $p_{ft} + q_{ft} = \tilde{\lambda}_{ft} + \frac{1}{\mu_{ft}} q_{ft}$ to be clear about the log demand curve rather than the log revenue function:

$$p_{ft} = \tilde{\lambda}_{ft} + \frac{1}{\mu_{ft}} q_{ft} - q_{ft} = \tilde{\lambda}_{ft} + \frac{1-\mu_{ft}}{\mu_{ft}} q_{ft} \tag{2.25}$$

For my purposes in this paper the key thing I take from the FMMM framework is that it allows imposing minimal assumptions on preferences and market structure to deliver consumer demand for a variety governed by the two heterogeneities: the price elasticity of demand which delivers the slope of the log demand curve at the profit-maximising point, $(1-\mu_{ft})/\mu_{ft}$, and the demand shifter, λ_{ft}/μ_{ft}.

The FMMM framework is also useful in linking revenue-based TFP and quantity-based TFP. The production function with Hicks-neutral TFP, such as the Translog used here in Equation (2.19) can be written as $q_{ft} = \bar{q}_{ft} + a_{ft}$ where $\bar{q}_{ft}$ is the index of inputs use that we label log scale.[9] Finally, by defining revenue TFP as $TFP^R_{ft} \equiv r_{ft} - \bar{q}_{ft}$ and using equation (2.22) while substituting we get:

$$TFP^R_{ft} = \frac{a_{ft}}{\mu_{ft}} + \tilde{\lambda}_{ft} + \frac{1-\mu_{ft}}{\mu_{ft}} \bar{q}_{ft} \tag{2.26}$$

meaning that TFP^R_{ft} is a (non-linear) function of quantity-based TFP a_{ft}, the log revenue shifter $\tilde{\lambda}_{ft}$, the profit-maximizing markup μ_{ft} and log production scale $\bar{q}_{ft}$.

[7] FMMM also show λ_{ft} is a measure characterizing differences in utility in the oligopoly model developed in Atkeson and Burstein (2008) and further refined in Hottman et al. (2016)

[8] In the case of CES and generalized CES preferences (2.24) holds as an equality because the log revenue function is linear in both q_{ft} and λ_{ft}.

[9] For example, with the Translog production technology:

$$\begin{gathered}\bar{q}_{ft} = \beta_l l_{ft} + \beta_{ll} l^2_{ft} + \beta_m m_{ft} + \beta_{mm} m^2_{ft} + \beta_k k_{ft} + \beta_{kk} k^2_{ft} + \\ \beta_{lm} l_{ft} m_{ft} + \beta_{lk} l_{ft} k_{ft} + \beta_{mk} m_{ft} k_{ft} + \beta_{lmk} l_{ft} m_{ft} k_{ft},\end{gathered}$$

.

2.4 Data

DGKP use the Prowess dataset[10] produced by the Centre for Monitoring Indian Economy (CMIE), for the period 1989-2003. This privately-managed subscription based service collects data on 45,000 listed and unlisted Indian companies from 1988-2018. For the period 1989-2003 and narrowing the scope to manufacturers, it covers 5,000 firms, representing around 60%-70% of industrial activity (De Loecker et al., 2016), in an unbalanced panel. It is not a representative sample of firms but makes use of the detailed reporting required by firms under the 1956 Companies Act that means CMIE can collate panel data for the largest firms in the Indian economy. The reporting requirements are broad, and include, for each manufactured product, the value, quantities and units of sales, data that is not available in a representative cross-sectional sample such as the Annual Survey of Industry (ASI) for the trade reform period. In total Prowess contains approximately 14,000 variables collected from annual and interim financial reports, stock exchanges and regulatory sources.

DGKP provide replication files – Stata and Matlab syntax and an intermediate data file containing firm-product level data required to generate the published figures and tables. This subset of Prowess variables is not sufficient to perform my analysis since I require measures of all the firm-product inputs in order to re-estimate the production function to get at the FMMM measures of firm heterogeneity.

I therefore independently access the raw data from the Prowess database. While CMIE continue to provide Prowess to researchers, the service has undergone changes since DGKP downloaded their data in 2005 which make it both easier to work with and more difficult to perfectly replicate the dataset. CMIE uses a proprietary product classification that can be concorded to the Indian National Industry Classification (NIC) but when DGKP obtained the data in 2005, it did not link the product names reported by firms to the classification.[11] Recent vintages (snapshots produced semi-annually) of the data are linked to standardised names and then to product codes but DGKP had to manually map names[12]. DGKP

[10]This dataset was also used by Stiebale and Vencappa (2018); Goldberg et al. (2010b); Chakraborty and Raveh (2018) and others. See Goldberg et al. (2010b) for more details of the construction.

[11]It also used a 12 digit classification, while it now uses a 20 digit classification. In practice the manufacturing data uses a maximum of 16 of the available 20 digits, with 97.3% of the observations in my data identified using 12 digits or fewer.

[12]In the 2018 vintage I use, there are 16,550 unique product names linked to 2987 standardised names. For instance, firms use 49 unique names to report production of 'Steel castings'.

Table 2.1: Variable names and descriptions

Variable	Prowess 2005 description	Prowess 2018 description
Labour	Total wage bill which includes bonuses and contributions to employees' provident funds	Compensation to employees which includes bonuses, pensions, social security and other items
Intermediates	Consumption of commodities by an enterprise in the process of manufacturing or transformation into product	Raw materials, stores & spares
Capital	Gross fixed assets, which includes movable and immovable assets	Gross fixed assets

Source: DGKP for Prowess 2005 variable descriptions. Centre for Monitoring Indian Economy (CMIE) for Prowess 2018 names.

kindly supplied their original mapping of product codes to NIC 1998 but changes to the Prowess and NIC classifications[13] mean that it is not possible to replicate the exact product mapping. I instead use the current (updated 2018) CMIE classification mapped to NIC 2004 using a file provided to me by CMIE and then manually concord the handful of changes between NIC 1998 and NIC 2004.

The DGKP intermediate data file contains firms from a limited subset of manufacturing industries. It includes all NIC 2-digit industries except for division 16 (Tobacco products), 18 (Wearing apparel), 19 (Leather), 20 (Wood and wood products), 22 (Publishing and printing), 23 (Coke and refined petroleum), 30 (Office, accounting and computing machinery), 32 (Communication equipment and 33 (Medical, precision and optical instruments). The omitted industry divisions contain few observations (less than 1,000 per industry over the 1989-2003 period) in my data, explaining the omission by DGKP. I use the same industries as DGKP.

Prowess contains several measures of materials, labour costs and capital, and I use the variable closest in description to those reported in DGKP. Table 2.1 reports changes in names. I follow DGKP in obtaining two-digit NIC wholesale price indices.[14] to deflate sales and inputs[15]

[13]NIC is based on the Standard Industrial Classification and underwent a substantial revision in 2008.

[14]Downloaded from the Department for Promotion of Industry and Internal Trade at http://eaindustry.nic.in/ .

[15]Multi-product firms in the data often produce in multiple two-digit industries. I deflate these firm-level inputs by assigning the firm to the industry from which comes its largest proportion of output. For product

The Prowess data is delivered in separate modules, of which I make use the companies' (standardised) standalone annual financial statements, which contain information on sales, employment costs, raw materials use and capital stock; and the products module which contains information on values, quantities and units of product sales.

I do some initial cleaning of the raw data, dropping duplicate observations (firms sometimes report more than once in a 365-day period as they change accounting year-end). Most Indian firms do not use calendar years for financial reporting purposes while the tariff is calendar year based. Following Kalemli-Ozcan et al. (2015) which uses similar firm-level data, I assign observations to calendar year $t-1$ if the firm's accounting year-end falls earlier than June 1.

From the products file, I drop lines that are reported to be clubbed together with another (unreported) product, before summing sales quantities and values using the most detailed (20-digit) CMIE product definition. I then combine the product and financial data. This gives me two measures of the value of production – from total sales of produced goods and from the sum of product sales. I drop all firm-year observations where these separately reported figures are different by +\-10%. Finally, I drop firms with reported raw materials use greater than the sum of product sales and those that report negative raw materials use.

Table 2.2: Summary statistics for full sample by industry sector

Sector	Share output	All firms	Single product	Num products
15 Food products & beverages	0.11	317	126	198
17 Textiles & apparel	0.10	318	173	100
21 Paper & paper products	0.02	87	64	38
24 Chemicals	0.28	435	170	696
25 Rubber & plastic	0.07	129	78	99
26 Non-metallic mineral products	0.07	106	73	89
27 Basic metals	0.14	226	110	139
28 Fabricated metal products	0.01	63	41	63
29 Machinery & equipment	0.05	96	46	224
31 Electrical machinery & communications	0.03	63	39	123
34 Motor vehicles & trailers	0.11	77	44	102
Total	1.00	1,917	964	1,871

Notes: Indian National Industrial Classification 1998. All columns show annual averages. 'All firms' is unique firms in full estimation sample, 'Single product' refers to firms producing one unique product and 'Num products' is the number of unique products within each sector.

sales values I use the product's appropriate industry deflator.

The DGKP replication file[16] contains 53,807 firm-product-year observations, substantially fewer than in my replicated dataset (61,448), and particularly so in the crucial years for the analysis of trade reform, 1989-1997 which accounts for 44% of the original sample (24,031 observations) but 55% of the new sample (33,587 observations). It is not clear why the datasets differ by so much. Enquiries made to CMIE indicate that no new firms are added to the older data, so the discrepancies are more likely due to either the product matching procedures or some differences in the preliminary data-cleaning steps undertaken in DGKP and my sample.

The distribution of firms across industrial sectors, and between single- and multi-product firms, shown in Table 2.2 is however qualitatively similar in DGKP. Chemicals and Basic metals account for the largest shares of average annual output in the new sample, as they do in DGKP (28% and 14% vs 26% and 16% respectively). The average number of unique firms per year is similar (1,971 in the new sample vs. 1,970 in DGKP). There are somewhat fewer single-product firms in the new sample (964 average per year vs. 1,047) as the product data is more granular in the new sample, containing 1,871 unique products compared to 1,400 in DGKP.

2.4.1 Tariff data

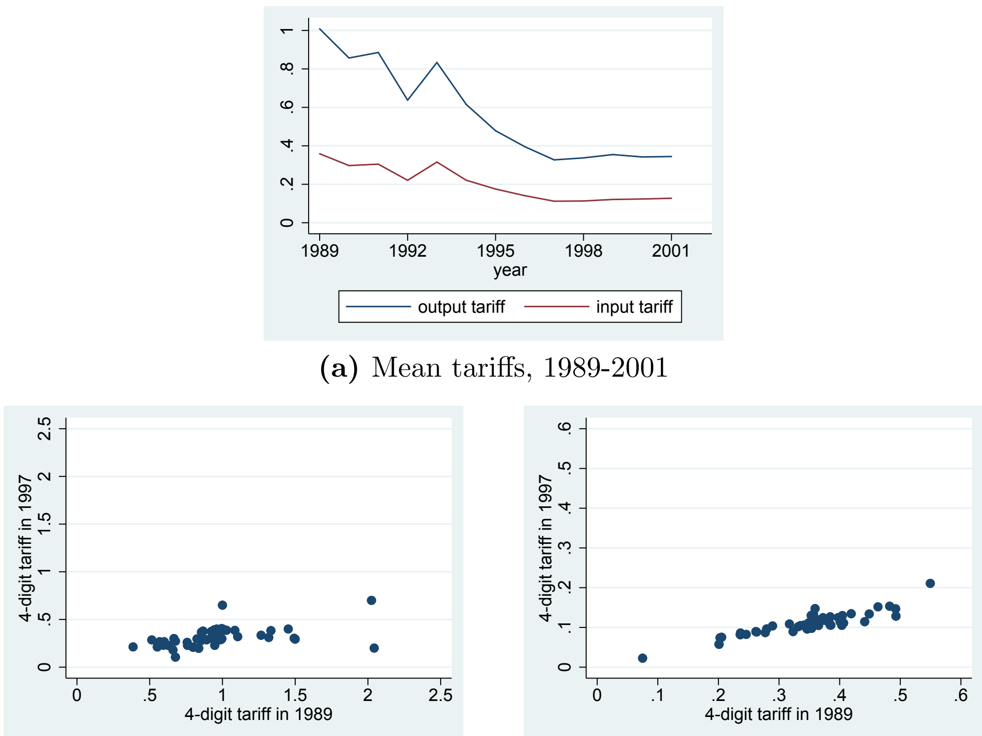
I take the industry level tariffs for 1989-2001 directly from the DGKP intermediate data file. These tariffs come from Topalova (2010). DGKP construct input tariffs using the Indian input-output tables for 1993-4 using the standard definition of $\tau_{it}^{input} = \sum_k s_{ki}\tau_{kt}^{output}$, where s_{ki} is the share of industry k in industry i and τ_{kt}^{output} is the tariff on products in industry k at time t. These tariffs were collected at the six-digit HS level and concorded using Debroy and Santhanam (1993) to the Indian National Industrial Classification. The tariffs are almost all at the four-digit NIC 1998 industry level with a handful at three-digit and two-digit levels.

The identification strategy employed in DGKP and this paper relies on tariff changes over the period being exogenous to firms' pricing and production decisions. Topalova and Khandelwal (2011) provide evidence that this was case up to the Indian election in 1997. The period up to 1997 is therefore the focus of DGKP and my paper. Panel (a) of Figure

[16] Available as 'estimationfile.dta' in the Supplemental Material at https://doi.org/10.3982/ECTA11042.

2.1 shows how the mean output and input tariffs faced by producers in my sample changed over the 1989-2001 period. Mean output tariffs began the period at 101% and declined to 32.7% by 1997, while mean input tariffs in the sample fell from 39.9% to 11.2%. Tariffs fell in every industry, with those producers facing the highest tariffs in 1989 seeing the largest reductions by 1997, as shown in panels (b) and (c). This pattern of liberalisation is helpful for the empirical strategy, first because there is a large amount of variation in tariff changes over time between industries, and second because these across-the-board cuts indicate that there was little or no favouritism shown to particular industries based on pre-reform characteristics, as shown more formally in Topalova and Khandelwal (2011).

Figure 2.1: Industry tariffs during the trade reform period



(a) Mean tariffs, 1989-2001

(b) Industry-level output tariffs, 1989 vs 1997

(c) Industry-level input tariffs, 1989 vs 1997

Notes: Panel (a) plots mean (unweighted) tariffs per year in the estimation sample. See text for calculation of input tariffs. Panels (b) and (c) plot the tariffs facing 80 4-digit industries in 1989 against the tariffs faced in 1997. Panel (b) shows output tariffs and panel (c) shows input tariffs. Source of tariffs is DGKP.

2.5 Estimation and results

2.5.1 Production function estimation

I follow DGKP in estimating a Translog production function using 1989-2003 data on single-product firms to recover the vector of input coefficients, $\boldsymbol{\beta}$ and the parameter vector $\boldsymbol{\delta}$, and then using the multi-product input assignment procedure to recover the input shares, ρ_{fjt}. Appendix A-1 describes in detail the close quantitative and qualitative replication of the DGKP procedure with my dataset. I then calculate markups using equation (2.18), and then estimates of TFP (a), product appeal/demand (λ), the demand shifter (λ/μ) and the demand slope ($1/\mu$) from equations (2.19), (2.21) and (2.24). Finally in preparation of the estimation sample summarised in Table 2.3, I trim outliers of markups, demand (λ) and the demand shifter (λ/μ)[17] to provide a sample size of 27,504 firm-product-year observations, substantially more than the 21,246 in the DGKP estimation sample.

Panel A of Table 2.3 shows summary statistics at the firm-year level. On average, firms produce 2.26 products in the average year, while more than 90% produce four or fewer products. Panel B, showing the firm-product-year level, highlights the large amount of variation in TFP between producers and the considerably larger variation in product appeal, the demand shifter and the demand slope.

Of the 27,504 firm-product-year observations in the sample, 9,448 or 34% have markups of less than 1, appearing to violate standard profit maximisation assumptions. However, as DGKP point out, multi-product firms maximise profits across products. As in DGKP, in my data, average firm markups (weighted by annual within-firm sales shares) are below 1 for only 10% of firms while the analysis also considers firm-product fixed effects so using variation within firm-products over time rather then depending on levels.

2.5.2 Replication and discussion of main DGKP results

DGKP exploit variation within firm-product pairs to show (1) the pass-through of marginal cost to prices is incomplete; (2) that the large fall in output tariffs over the 1989-1997

[17] DGKP trim the top and bottom three percentiles of the markup distribution when presenting their main results. I trim on within-firm changes as this is the variation used in the analysis, although results are not substantially different.

Table 2.3: Firm and firm-product summary statistics

	mean	sd	p10	p50	p90
Panel A: Firm-level variables					
firm sales	322.35	1110.84	14.25	100.49	651.31
firm value added	158.86	577.21	4.30	38.45	319.93
firm products	2.26	2.01	1.00	2.00	4.00
Observations: 14,504					
Panel B: Firm-product-level variables					
sales	137.91	517.16	2.37	39.58	303.18
value added	64.73	451.87	-20.04	12.85	170.43
materials	73.19	240.84	2.32	24.90	155.73
labour	11.56	41.32	0.27	2.93	26.07
capital	109.07	499.91	2.50	26.20	198.62
$\ln price$	-1.68	3.03	-5.10	-2.51	2.56
$\ln quantity$	5.15	3.40	0.54	5.47	9.26
$\ln mc$	-1.79	3.14	-5.38	-2.43	2.44
TFP-R	0.89	1.46	-0.96	0.96	2.51
TFP a	2.57	3.14	-1.88	3.32	6.10
product appeal λ	2.94	14.24	-5.20	-0.02	12.84
markup μ	2.01	2.96	0.20	1.34	4.25
production scale $\bar{q}$	2.58	1.82	0.26	2.73	4.71
demand shifter λ/μ	-5.65	26.90	-16.09	-0.02	3.99
demand slope $1/\mu$	2.43	6.17	0.24	0.75	5.04
Observations: 27,504					

Notes: Summary statistics refer to firm-year observations in Panel A and firm-product-year observations in Panel B. Sales, value-added, materials, labour and capital measured in 1981-1982 Indian rupees.

period is associated with only a small fall in prices; (3) that the fall in input tariffs reduced marginal costs – and hence via (1), generated higher markups; and (4) that when controlling for marginal cost in a regression of markups on output tariffs so as to estimate the pure pro-competitive effect of trade reform, that these effects helped to moderate the rise in markups.

I replicate these results with the new Prowess data, leaving comparisons and full analysis to Appendix A-1 but here showing the main specification (2.27) as it provides the starting point for my analysis.

Table 2.4: OLS regressions of prices, marginal costs and markups on output and input tariffs

	(1) $\ln price$	(2) $\ln marg.cost$	(3) $\ln markup$
τ^{output}	0.1788	0.0664	0.1125
	(0.0590)***	(0.0591)	(0.0647)*
τ^{input}	-0.0385	0.7862	-0.8247
	(0.5066)	(0.4408)*	(0.2952)***
R^2	0.03	0.02	0.01
N	27,504	27,504	27,504
Firm-product FEs	Yes	Yes	Yes
Sector-Year FEs	Yes	Yes	Yes
Overall impact	-11.2	-23.9***	12.7**
Standard error	10.6	9.0	6.2

Notes: Corresponds to DGKP, Table IX, p.491. Standard errors clustered by industry. All regressions include a constant, firm-product fixed effects and sector-year fixed effects. The last two rows use the average decline in output tariffs of 68 percentage points and the average decline of 24 percentage points in input tariffs to compute the mean and standard error of the overall impact on prices. R^2 is the within R-squared. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

$$\ln price_{fjt} = \alpha_{fj} + \alpha_{st} + \beta_1 \tau_{it}^{output} + \beta_2 \tau_{it}^{input} + \epsilon_{fjt} \tag{2.27}$$

Prices are regressed on output and input tariffs, along with firm-product fixed effects and a full set of year-times-sector (two-digit) dummies. The tariffs are at the (four-digit) industry level and so the β coefficients in equation (2.27) are interpreted as the difference in the percentage change in prices for a firm in an industry that experiences a 1 percentage point greater change in tariffs than another. Results are shown in column (1) of Table 2.4, while columns (2) and (3) replace prices with marginal costs and markups as the dependent variable. Since log prices are the sum of log marginal costs and log markups, the coefficients of interest also sum across the three regressions, i.e. the effect of output tariffs on prices is equal to the sum of the effects of output tariffs on marginal costs and on markups. Standard errors are clustered at the four-digit industry level, the "treatment" level at which tariffs vary across firms.

The coefficient on output tariffs in column (1) is positive and significant, showing that a fall in output tariffs of 10 percentage points acts to reduce output prices by 1.788%, conditional on input tariffs. DGKP argue that this effect can be thought of as an increase in competition, acting either to reduce markups or improve X-efficiencies.

A fall in input tariffs, however, is not passed through to prices at any reasonable significance level. The effects of changes in input tariffs, conditional on output tariffs, would be associated with access to cheaper or more varieties of intermediates that act to reduce marginal costs – indeed, DGKP argue that this is the only channel for input tariffs – but, in this specification there is no effect on prices.

Column (2) shows results for marginal costs, where a change in output tariffs has no significant affect, implying that competition does not lead firms to reduce costs. However, a 10 percentage point fall in input tariffs is associated with a 7.86% fall in marginal costs, implying that firms do benefit from lower prices of intermediates. Column (3) shows changes in markups. The reduction in output tariffs, leading to lower prices but largely unchanged marginal costs weakly acts to reduce markups. However, it is the reduction in input tariffs that sharply reduces marginal costs without affecting prices and acts strongly to raise markups.

Table 2.4 also shows the overall impact of the fall in tariffs on prices over the 1989-1997 period, which can be computed by multiplying the average fall in tariffs – of 68 percentage points and 24 percentage points for output and input tariffs – by β_1 and β_2 respectively, and is shown together with the standard error in the final two rows. Average firm prices, net of sector-year fluctuations, fell by only 11.2% due to the trade reform and were not statistically different from zero. However, the reduction in tariffs reduced firms' marginal costs by 23.9% – far more than 11.2% reduction in factory-gate prices, leaving the combined impact of the tariff changes to raise markups by a statistically significant 12.7%.

DGKP's main claim is that these results show that firms are the chief beneficiaries of trade reform via higher markups, at least in the short run before any dynamic effects are taken into account. They provide weaker evidence that firm-products that begin the period with top decile markups reduce markups more than other firms. However, the result is also consistent with firms upgrading the quality of existing products. DGKP argue quality upgrading alone would not be a sufficient mechanism to produce these results. If there was complete passthrough of marginal cost changes to prices, then quality upgrading would not result in higher markups. And, since quality is costly, upgrading, conditional on productivity and input prices, would imply higher marginal costs whereas DGKP find lower marginal costs in response to tariff reform.

By using the measures of demand following FMMM I show that quality upgrading is an important part of the story, and that these demand changes are much stronger for varieties that have weaker measures of performance when they are first observed in the data. For better established varieties, efficiency improvements made in response to tariff reform are more important.

2.5.3 Key features of FMMM measures

Prices, demand parameters and TFP

Tables 2.5, 2.6 and 2.7 provide OLS regressions that show correlations between key features of the data and MULAMA model measures. In these regressions, I use product-unit dummies and year dummies while reporting robust standard errors clustered at the firm level in order to show the relationships across firms. Table 2.5 column (1) shows that the association across firms between TFP and the demand shifter is positive, an important feature of the Prowess data that stands in contrast to that found in other settings that consistently find a negative correlation (Forlani et al. (2016) and see other chapters in this thesis). This feature of the data in this paper is robust when also considering firm-unit-product fixed effects, to the untrimmed data, to further trimming of outliers and to considering changes over time in the variables. While the MULAMA model allows this correlation to be unrestricted, the result is somewhat surprising considering the usual interpretation of a negative correlation as a quality-quantity trade-off due to higher-quality/demand goods necessitating higher-cost inputs. It is however consistent with models such as Antoniades (2015) in which higher quality products are associated with high productivity (low marginal cost) producers. Briefly putting this result to one side, both TFP and the demand shifter in relation to quantities and prices behave in ways consistent with economic intuition. Higher TFP is associated with larger quantities (column 2) and lower prices (column 3), while higher demand is associated with larger quantities (column 4) and higher prices (column 5). Finally, column (6) shows that conditional on TFP, the demand shifter is associated with higher prices.

Table 2.5: OLS regressions involving prices, demand parameters and TFP

	(1) ln demand shifter λ/μ	(2) ln *quantity*	(3) ln *price*	(4) ln *quantity*	(5) ln *price*	(6) ln *price*
TFP a	6.778	0.840	-0.177			-0.251
	(0.427)***	(0.021)***	(0.018)***			(0.021)***
ln demand shifter λ/μ				0.007	0.007	0.011
				(0.001)***	(0.001)***	(0.001)***
Product FEs	yes	yes	yes	yes	yes	yes
N	27,504	27,504	27,504	27,504	27,504	27,504
R^2	0.47	0.87	0.92	0.82	0.92	0.93

Notes: Standard errors clustered by firm. All regressions include product-unit and year dummies. R^2 is the within R-squared.
$p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Demand parameters and costs

In order to better understand the relationship between TFP and demand shown in column (1) of Table 2.5, Table 2.6 provides OLS regressions of various model measures on log marginal costs and log input prices. Column (1) shows a positive and significant marginal cost to prices pass-through elasticity, very similar in size to the 0.413 (with 0.029 standard error) within-firm results reported in Appendix Table A.2. Higher marginal costs are also, as expected, associated with lower TFP in column (2). Consistent with the positive relationship between TFP and the demand shifter, lower marginal costs are associated with higher demand, whether measured by the relative product appeal, λ (column (3)) or the demand shifter, λ/μ (column(4)). This is despite one of the identifying assumptions required when using output prices in the control function for input prices/quality in equation (2.5), being that higher quality/demand outputs are made from more expensive inputs. Clearly, this result needs further explanation. The key is that it is firm-product input prices and not marginal costs that appear in the control function and columns (5)-(8) show the relationship of these input prices with marginal costs, output prices and demand.

DGKP make use of an assumption, following typical O-ring theory (Kremer, 1993), that high quality (and cost) labour, materials and capital are complementary and therefore their prices can be captured using a single firm-product specific price index. Operationally, these prices, w_{fjt}, are retrieved from the single product production function estimation which yields the $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ vectors in the input price bias control function in equation (2.4), $B(\boldsymbol{w}_{ft}, \tilde{\boldsymbol{x}}_{ft}, \boldsymbol{\beta})$, and then using equation (2.5) to calculate estimates of the input prices, $w_{it}^x = w_t\left(p_{it}, ms_{it}, \mathbf{D}_f, EXP_{it}\right)$.

Table 2.6: OLS regressions of cost measures on prices, demand parameters and TFP

	(1) ln *price*	(2) ln TFP a	(3) ln product appeal λ	(4) ln demand shifter λ/μ	(5) ln $price^{inputs}$	(6) ln *price*	(7) ln product appeal λ	(8) ln demand shifter λ/μ
ln marg. cost	0.452	-0.633	-2.870	-4.130	0.345			
	(0.016)***	(0.017)***	(0.217)***	(0.344)***	(0.015)***			
ln $price^{inputs}$						0.953	0.807	1.558
						(0.022)***	(0.189)***	(0.290)***
N	27,504	27,504	27,504	27,504	27,504	27,504	27,504	27,504
R^2	0.95	0.96	0.38	0.44	0.37	0.97	0.32	0.41
Product FEs	yes	yes	yes	yes	yes	yes	yes	yes

Notes: Standard errors clustered by firm. All regressions include product-unit and year dummies. R^2 is the within R-squared.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Column (5) confirms that firm-products with higher cost inputs have a higher marginal cost of production. Higher cost inputs are also associated with higher prices (column (6)) with an elasticity very close to 1, a result that provides comfort that the input price control function is correctly specified. Finally, columns (7) and (8) show that firms with higher input prices do indeed produce higher relative-demand products and have a demand curve with a higher intercept term. Together these results suggest that in the Indian setting higher quality/demand products are produced with more expensive inputs, but at lower marginal costs.

One explanation could be that the more technically efficient firms are those producing for higher-value markets, even within detailed product categories, as in for example, Antoniades (2015), or that because our demand measures capture perceived quality, a product sold widely in a national market will have a higher demand shift and product appeal measure than a product sold only locally at the same price.

Prices, marginal costs and markups on MULAMA measures

Table 2.7 provides OLS regressions of prices, marginal costs and markups on TFP, demand heterogeneity and production scale, showing the conditional correlations for these MULAMA measures across firms within products (columns (1), (3) and (5)) and within firm-products (columns (2),(4) and (6)). Here, I follow DGKP to make results consistent with the analysis of trade reform by including sector-by-year fixed effects to control for macroeconomic fluctuations at the 2-digit sector level, and once again, standard errors are

Table 2.7: OLS regressions of prices, marginal costs and markups on MULAMA model measures

	(1) ln *price*	(2) ln *price*	(3) ln marg. cost	(4) ln marg. cost	(5) ln markup	(6) ln markup
TFP a	-0.283	-0.346	-1.013	-0.854	0.730	0.508
	(0.017)***	(0.038)***	(0.016)***	(0.031)***	(0.014)***	(0.035)***
ln demand shifter λ/μ	0.009	0.012	-0.002	-0.005	0.010	0.017
	(0.001)***	(0.001)***	(0.000)***	(0.001)***	(0.001)***	(0.001)***
scale $\bar{q}$	-0.319	-0.339	-0.309	-0.362	-0.010	0.023
	(0.018)***	(0.031)***	(0.022)***	(0.033)***	(0.009)	(0.022)
Sector-year FEs	yes	yes	yes	yes	yes	yes
Product FEs	yes		yes		yes	
Firm-Product FEs		yes		yes		yes
N	27,504	27,504	27,504	27,504	27,504	27,504
R^2	0.94	0.99	0.95	0.99	0.85	0.96

Notes: Standard errors clustered by firm. All regressions include sector$x$year dummies. Regressions in columns (1), (3) and (5) include product dummies and columns (2), (4) and (6) firm-product dummies. R^2 is the within R-squared.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

clustered by firm.

Whether considering product or firm-product fixed effects, higher TFP (conditional on demand and firm size) is associated with lower prices, sharply lower marginal costs, and higher markups. Likewise, firms with higher demand, conditional on TFP and production scale (note that since $q_{fjt} = a + \overline{q_{fjt}}$ this effectively controls for production quantity) is associated with higher prices, and as discussed above, somewhat lower marginal costs, with a combined effect of higher markups. Finally, bigger production scale is associated with lower prices and lower marginal costs, but no significant effect on markups, conditional on TFP and demand.

2.5.4 MULAMA demand measures and trade reform

Having established these key relationships, I turn to the relationship between demand and tariffs, re-running equation (2.27) but with the demand shifter and demand slope as the dependent variable. Table 2.8 report the results.

The positive and significant coefficient on output tariffs in column (1) shows how lower tariffs reduce residual demand amid higher import competition, and this shift in the demand curve results in a lower markup and consequently more elastic demand – a higher (i.e. flatter) slope ($1/\mu$), which is seen in column (4). But lower input tariffs act to oppose

Table 2.8: OLS and IV regressions of MULAMA demand measures on output and input tariffs

	ln demand shifter λ/μ			ln demand slope $1/\mu$		
	(1)	(2)	(3)	(4)	(5)	(6)
τ^{output}	4.567	4.786	4.599	-1.234	-1.343	-1.266
	(1.883)**	(1.923)**	(2.118)**	(0.602)**	(0.647)**	(0.640)**
τ^{input}	-32.907	-30.315	-22.768	11.364	10.071	7.361
	(12.557)**	(13.211)**	(11.278)**	(4.708)**	(5.132)*	(3.647)**
R^2	0.02	0.07	0.09	0.02	0.15	0.14
N	27,504	27,504	14,922	27,504	27,504	14,922
Control for MC	no	yes	IV	no	yes	IV
Firm-product FEs	yes	yes	yes	yes	yes	yes
Sector-year FEs	yes	yes	yes	yes	yes	yes

Notes: Standard errors clustered by 4-digit industry. All regressions include firm-product fixed effects and sector×year dummies. Columns (2) and (4) control for log marginal cost (coefficients not shown), columns (3) and (6) instrument for log marginal cost with lagged marginal cost and TFP a. R^2 is within-R-squared. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

these changes, shifting demand upwards and, via markups, making it less elastic. This higher demand could be the result of 1) an increase in product quality, or 2) an indirect effect of tariffs via lower marginal costs which allows a product to be sold more widely. To show the role of quality, in columns 2-3 and 5-6 I control for marginal costs first using OLS and then, to account for measurement error, using an instrumental variables approach. DGKP in some specifications instrument for marginal costs using lagged marginal cost and τ^{input} to deal with measurement error in the estimated variable. They note that it is only necessary for the instruments to be uncorrelated with the measurement error, i.e. there is no serial correlation in the measurement error of marginal cost. In these regressions I instrument marginal costs with lagged marginal costs and TFP, a_{fjt}, instead of lagged marginal costs and τ^{input} since the latter is already included in its own right.[18] The coefficients and standard errors on output tariffs barely change and, although the coefficients on input tariffs are somewhat smaller they remain significant, indicating that the effect of input tariffs on demand does not come indirectly via marginal cost changes. This is evidence that trade reform prompts firms to upgrade product quality.

Taken together, these results show that lower tariffs affect demand through both the slope and position of the demand curve and indicate that that firms are using the availab-

[18]The IV requires a lagged variable, reducing the sample size and I check that the columns (1) and (2) results are consistent when using the IV sample.

ility of cheaper and/or higher quality inputs to increase output quality and are charging correspondingly higher markups. This is a departure from the DGKP story that firms' market power lets them retain the benefits of lower input prices from the tariff liberalisation. If product quality increases as a result of liberalisation then there are real gains for consumers, and the apparent limited pass through of marginal costs to prices is not the full story. The natural question to ask is, what happens to prices, marginal costs and markups conditional on product demand and quality?

Revisting prices, markups and trade reform

To explore this question further I return to the analysis of the main DGKP results shown in Table 2.4, and now control for the demand curve parameters, λ/μ and $1/\mu$ in equation (2.27). With this reduced form, it is important to note that the results have less of a causal flavour than those above, since the demand parameters are themselves not exogenous. Table 2.9 reports the results.

Columns (1), (3) and (5) reproduce results from Table 2.4 above, for comparison and columns (2), (4) and (6) add the demand shifter and slope as controls. In column (2), the addition of controls results in a lower coefficient on output tariffs, though it remains significant, implying that if demand had not changed due to the pro-competitive effects of liberalization in output markets, prices would still have fallen. Another way to think of this is that since we are holding the log demand curve constant, the price changes must be due to firms making efficiency improvements that shift the supply curve. Consistent with this, the coefficient on output tariffs in the column (4) regression of marginal cost becomes much larger and significant, indicating that lower output tariffs led to reductions in marginal costs that were more than fully passed through to prices. This finding is a departure from DGKP, which takes a different approach to estimating the pro-competitive effects of the reform, and finds that most of the output tariff effect in column (5) is due increased competition lowering price-marginal cost margins, see Appendix Table A.5. However, that approach depends on the assumption that input tariffs only affect markups via reductions in marginal costs, and not, as we have seen directly via higher quality inputs.

The coefficients on input tariffs change even more when controlling for demand. In column (2), the larger and positive (though still not statistically significant) coefficient points towards firms raising quality and being able to sustain higher prices. In other words, had output quality remained constant, lower input tariffs would have put downwards pressure on prices.

The result in column (4) is perhaps surprising, showing that when input tariffs fall, marginal costs do not change conditional on demand. It appears that marginal costs only fell when input tariffs were reduced because of quality changes, implying that lower marginal cost inputs were of a type that raised output quality. Again, this result stems from the finding in Tables 2.5 and 2.6 that quality and efficiency are positively correlated, while marginal costs and quality/demand are negatively correlated.

The overall impact of the price and marginal cost changes net of the effects of demand and quality, are a greater reduction of prices, 15.5% (now significant at the 10% level) instead of 11.2%, and a smaller, statistically insignificant effect on marginal costs of a 8.1% reduction instead of a strongly significant 23.9% reduction.

Column (6) shows that when controlling for product quality, the trade reform did not have any impact on average firm markups. Firms reduce prices and costs through lower input costs and increasing efficiencies when tariffs fall, at the same time producing higher quality products in response to the availability of cheaper and higher quality inputs. The higher markups documented in DGKP appear to be entirely driven by this process.

Table 2.9: OLS regressions of prices, marginal costs and markups on tariffs with controls for demand

	ln *price*		ln marg.cost		ln markup	
	(1)	(2)	(3)	(4)	(5)	(6)
τ^{output}	0.1788	0.1368	0.0664	0.1688	0.1125	-0.0320
	(0.0590)***	(0.0419)***	(0.0591)	(0.0787)**	(0.0647)*	(0.0548)
τ^{input}	-0.0385	0.2504	0.7862	-0.1397	-0.8247	0.3901
	(0.5066)	(0.3920)	(0.4408)*	(0.6274)	(0.2952)***	(0.3330)
R^2	0.03	0.07	0.02	0.15	0.01	0.31
N	27,504	27,504	27,504	27,504	27,504	27,504
Controls for demand		Yes		Yes		Yes
Firm-product FEs	Yes	Yes	Yes	Yes	Yes	Yes
Sector-Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Overall impact	-11.2	-15.5*	-23.9***	-8.1	12.7**	-7.4
Standard error	10.6	9.1	9.0	12.1	6.2	6.2

Notes: Standard errors clustered by industry. All regressions include a constant, firm-product fixed effects and sector×year fixed effects. Controls for demand in columns (2), (4) and (6) are the demand shifter λ/μ and the demand slope $1/\mu$. R^2 is the Within R-squared. The final two rows use the average decline in output tariffs of 68pp and the averge decline of 24pp in input tariffs to compute the mean and standard error of the overall impact, net of the controls, on the dependent variable. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

2.5.5 Firm-product heterogeneity and trade reform

There is abundant empirical evidence that the effects of trade liberalization on product quality are heterogeneous (for example, see Fan et al. (2015, 2018); Khandelwal (2010); Kugler and Verhoogen (2011); Verhoogen (2008); Manova and Zhang (2012)) and typically the size and productivity of firms are the characteristics that determine the heterogeneous effects, as supported by theoretical models such Kugler and Verhoogen (2011), Antoniades (2015) and Fan et al. (2018). The results above that show that demand-side effects of trade reform are key to understanding firm pricing and markup decisions, is also suggestive of heterogeneous effects since these are factors that vary across firms and products. I investigate the heterogeneous effects of the Indian trade reform by considering the differential effects on varieties depending on their pre-treatment performance. Guided by the existing empirical and theoretical evidence, I define a size dummy using the market share of each variety within a four-digit industry in the year in which the variety first appears in the data, assigning the dummy a value of 1 if it has an above-median market share and 0 otherwise.[19]

[19] DGKP use a similar method to categorise high markup firms (top quintile) and find some evidence that the pro-competitive effects of trade reform are felt more strongly by firms that begin the period with relatively higher markups.

Table 2.10: Summary statistics for low- and high-share varieties

Variable	Low share	High share	Δ
$\ln quantity$	-0.520	0.695	1.215***
$\ln price$	-0.035	0.047	0.082***
$\ln marg.cost$	0.175	-0.234	-0.409***
$\ln markup$	-0.220	0.537	0.758***
TFP a	-0.204	0.273	0.477***
demand shifter λ/μ	-1.843	2.461	4.304***
demand slope $1/\mu$	3.533	0.950	-2.583***

Notes: All variables except $1/\mu$ de-meaned by product-unit fixed effects. Mean values shown. Low share/High share refers to firm-products sales value in each half of the 4-digit industry distribution in the first year in which the firm-product appears in the data. The difference in means, Δ, is shown in the final column along with significance based on a t-test of the equality of means with unequal variance. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 2.10 shows sample means across the high and low market share varieties, and confirms some key findings in the literature. High market share varieties outperform peers in all dimensions. They sell higher quantities at higher prices and with much higher markups, while productivity (TFP-Q) is higher.[20] Marginal costs lower are lower for these varieties, while they face higher, less elastic demand.[21]

I then re-run the reduced form equation (2.27) using as the dependent variable in turn, log prices, marginal costs, markups, quantities, TFP, the demand shifter and demand slope, and interact the time-invariant size dummy variable with both output and input tariffs, running the following specification,

$$\begin{aligned}\ln Z_{fjt} =& \alpha_{fj} + \alpha_{st} + \beta_1 \tau_{it}^{output} + \beta_2 \tau_{it}^{input} + \\ & \gamma_1(\tau_{it}^{output} * top_{fj}) + \gamma_2(\tau_{it}^{input} * top_{fj}) + \epsilon_{fjt},\end{aligned} \tag{2.28}$$

where $\ln Z_{fjt}$ is log quantity, log price, log marginal cost, log markup, log TFP-Q, the demand shifter and demand slope, and top_{fj} is the size dummy that takes the value 1 if a variety has an above-median market share within a four-digit industry when it first appears in the data and 0 otherwise.

Table 2.11 reports and decomposes the results into the total effects over the 1989-1997 period, providing evidence that the changes in prices, marginal costs and markups were

[20] Positive correlations between either output or export prices and productivity, and output prices and market shares are features noted by, among others, Verhoogen (2008), Kugler and Verhoogen (2011) and Manova and Zhang (2012).

[21] The authors cited above argue that these correlations require heterogeneous product quality.

mediated through very different channels for varieties with low initial market shares compared to those with high shares. Column (1) shows the overall effect on each variable due to the trade reform in the same way as the final two rows of Table 2.4 after estimating equation (2.28) without the interaction terms and then multiplying the β_1 and β_2 coefficients by the 68 percentage point drop in output tariffs and the 24 point drop in input tariffs over the period respectively. Columns (2) and (3) separate out the changes due to output and input tariffs respectively. The effects on low share varieties are calculated from the β coefficients estimated in equation (2.28) and the high share varieties from the β and γ coefficients. For the log-levels regressions (i.e., all except the demand shifter and slope) these numbers are percentage changes in each variable from 1989-1997, while for the levels-levels they are absolute changes.

What is striking about these results is how similar tariff-induced price declines of 10.6% and 11.9% for low and high share varieties respectively are derived through different channels. High share varieties respond to lower output tariffs with increases in technical efficiency which are almost fully passed through to marginal costs and prices (and hence unchanged markups), while demand parameters are unchanged.[22] Lower input tariffs for these varieties raise and steepen the demand curve, allowing firms to increase quantities while leaving marginal costs, prices and markups largely unchanged. But for the low-share varieties, markups change a lot. Indeed, almost all of the 12.69% increase in markups under the DGKP methodology is due to a 18.5% increase in markups for these varieties. Lower input tariffs raise quality much more strongly for the low-share firm-products than the high-share, allowing them to increase quantities sold without adjusting prices. Efficiency rises very strongly through the inputs channel as well.

The finding that it is the low-share varieties that benefit most from lower tariffs is similar with the evidence from Fan et al. (2018), that laggard firms in China when it entered the WTO upgraded output and inputs quality (and prices) faster than pre-entry leaders. However, it that case, the evidence concerns exporting firms rather than my paper with its focus on all products whether exported or not. Nataraj (2011) documents large increases in productivity among the smallest, informal firms in India during the 1990s, but this result also concerns firms not products and, further, the main mechanism

[22]There still may be underlying changes in demand heterogeneity, but any demand increases due to quality improvements are offset by a contraction in demand due to tougher competition.

for this is via reallocation from the exit of the least efficient small firms due to foreign competition. This reallocation channel is one that I have not considered, Prowess being unsuitable for analysis of firm entry and exit because of its focus on large firms that rarely exit the market. However, while selection effects – the exit of the least productive firms and varieties – could be a concern for my analysis, existing evidence at the firm and product level suggests that it was not a major driver of Indian economic performance over the period. Goldberg et al. (2010b), using the Prowess data finds that product dropping by firms was rare over the same period, with far less churn than comparable exercises in advanced economies, while Harrison et al. (2013) and Bollard et al. (2013) using broader datasets than Prowess also point to within-firm changes as being far more important than between-firm reallocations in driving up Indian productivity.

Table 2.11: Decomposition of impact of trade reform on low- and high-share producers

Variable	(1) Change due to trade reform	(2) ...of which due to output tariffs	(3) ... of which due to input tariffs
$\ln price^{\dagger}$	-11.23	-12.18***	0.95
low share	-10.56	-14.26**	3.70
high share	-11.89	-9.92***	-1.97
$\ln mc^{\dagger}$	-23.92***	-4.52	-19.40*
low share	-29.07***	-0.05	-29.02***
high share	-17.80**	-9.24*	-8.56
$\ln \mu^{\dagger}$	12.69**	-7.66*	20.35***
low share	18.51**	-14.21*	32.72***
high share	5.91	-0.68	6.58
$\ln quantity^{\dagger}$	22.03***	-1.04	23.07***
low share	32.35***	0.35	32.00***
high share	8.77	-2.98	11.75*
$a^{\dagger}$	21.90**	6.39	15.51
low share	27.03***	2.28	24.75**
high share	15.78**	10.73**	5.05
$\lambda/\mu^{\ddagger}$	5.01**	-3.11**	8.12**
low share	5.54**	-5.32*	10.85**
high share	4.55**	-0.70	5.25***
$1/\mu^{\ddagger}$	-1.96**	0.84**	-2.80**
low share	-2.19**	1.57	-3.76*
high share	-1.75**	0.05	-1.79***

Notes: Compiled using results from Table A.6. The average decline in output tariffs of 68pp and the average decline of 24pp in input tariffs are used to compute the mean of the impact on each dependent variable shown in the first column, for (a) the whole sample, (b) low market share firms and (c) high market share firms. Column 1 shows the overall impact of trade reform, summing the impacts of input and output tariffs. Column 2 shows the output tariff component and column 3 the input tariff component. †: percentage changes in variable. ‡: absolute change in variable. Significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

2.6 Conclusions

In this paper I reassemble the detailed product-level dataset used by DGKP and closely follow its approach to estimating marginal costs and markups, while drawing on Forlani et al. (2016) to obtain measures of consumer demand for products conditional on price without the need for onerous demand-side assumptions.

I first show that the impacts of the tariff liberalization on prices (a reduction of 11.2%), marginal costs (a reduction of 23.9%) and markups (an increase of 12.7%) are driven by changes to the demand faced by firms for their products, most notably by the reductions in input tariffs that allow firms to raise product quality. Conditional on this demand, prices fell more (by 15.5%), marginal costs fell less (by 8.1%) and markups if anything also fell, (by 7.4%). While both DGKP and my paper show factory gate prices falling by less than marginal costs, implying that firms received most of the benefits of the trade liberalization, the finding that product quality rose along with markups suggests that there were also big benefits for consumers.

Second, I show significant heterogeneity in firm product performance, measured by market shares, was such that almost all of the unconditional (on demand/quality) effects of lower tariffs on markups were due to improvements in initially laggard varieties. While average markups increased by 12.7% due to lower tariffs, the effects on below-median market share varieties resulted in an 18.5% increase, with only a 5.9% increase for above-median market share varieties. The initially poor performers saw the largest declines in marginal costs and largest increases in product quality associated with lower tariffs on imported intermediates.

The literature on the growth in revenue TFP in India after the 1991 reforms strongly supports the role of within-firm TFP growth rather than reallocation towards more productive firms (e.g. Harrison et al. (2013), Bollard et al. (2013)). A possible mechanism would have been a within-firm reallocation towards higher productivity products (Harrison et al., 2013; Bernard et al., 2010), the opposite of the effect I find and also inconsistent with Goldberg et al. (2010b) that shows little evidence of Indian firms dropping products during the period. The results presented in my paper offer an alternative mechanism consistent with both the growth of within-firm productivity and limited product churn:

that the input tariff reductions allowed firms to raise the quality, productivity and scale of laggard products. Further research into the dynamics of changes in product quality within multi-product firms in the Indian context would help build robustness to the results I have presented.

Chapter 3

On The Productivity Advantage of Cities

Abstract

Ever since Marshall (1890) agglomeration externalities have been viewed as the key factor explaining the existence of cities and their size. However, while the various micro foundations of agglomeration externalities stress the importance of Total Factor Productivity (TFP), the empirical evidence on agglomeration externalities rests on measures obtained using firm revenue or value-added as a measure of firm output: revenue-based TFP (TFP-R). This paper uses data on French manufacturing firms' revenue, quantity and prices to estimate TFP and TFP-R and decompose the latter into various elements. Our analysis suggests that the revenue productivity advantage of denser areas is mainly driven by higher prices charged rather than differences in TFP. At the same time, firms in denser areas are able to sell higher quantities, and generate higher revenues, despite higher prices. These and other results we document suggest that firms in denser areas are able to charge higher prices because they sell higher demand/quality products. Finally, while the correlation between firm revenue TFP and firm size is positive in each location, it is also systematically related to density: firms with higher (lower) TFP-R account for a larger (smaller) share of total revenue in denser areas. These patterns thus amplify in aggregate regional-level figures any firm-level differences in productivity across space.

3.1 Introduction

A stylized fact of economic geography is that the productivity of firms increases with city size and urban density (Combes and Gobillon, 2015), and a large literature going back to Marshall (1890) explores the question of why cities have this productivity advantage. Micro-foundations put forward for these agglomeration externalities are now typically grouped under the headings sharing, matching, learning and sorting (Duranton and Puga, 2004; Combes et al., 2008) and include different forms of knowledge spillovers between firms, costly trade, pro-competitive effects of city size, and sorting of workers (Syverson, 2011). The empirical literature suggests a rather consistent, across countries and years, range for the elasticity of productivity with respect to city size of 0.04-0.07 (Rosenthal and Strange, 2004). However, while theoretical micro-foundations for agglomeration externalities rest on differences across space in total factor productivity (TFP), i.e., the capacity to turn inputs into more physical output, empirical work has so far considered what we call revenue TFP measures (TFP-R), i.e., productivity calculated using revenue as a measure of output and so the capacity to turn inputs into more revenue.

To be more specific, researchers typically try to measure TFP as the residual obtained by estimating a production function through a regression of some measure of firm output on inputs. One key problem with this in practice is that usually the only output measures available are gross revenues or value-added, and not quantities. Revenues are of course made up of price and quantities. Even though industry-level price deflators are usually available they are of little use if the goal of the analysis is to pick up differences in productivity across space because they do not take into account differences in prices across locations. More broadly, revenue-based measures of productivity will pick up any heterogeneity in firm-level prices, confounding efforts to measure 'true' physical TFP. This heterogeneity in prices across firms could be due to many factors including firm-level demand shifters, markups and production scale. At a regional level, for instance, if firms in larger cities systematically sell higher-priced, higher-quality goods, the econometrician working with a measure of revenue TFP will overstate the impact of city size on TFP. At the same time, establishing that part of the observed revenue productivity advantage of cities is due to factors other than technical efficiency would require a substantial reconsideration of agglomeration economies and in particular of the related mechanisms and policy implications.

In order to address these issues we make use of high-quality and detailed quantity, prices and revenue data on products produced by French manufacturing firms. This type of data is becoming more widely available, allowing researchers to measure firm-specific TFP while considering the presence of other forms of heterogeneities across firms, and that is what we do in this paper. More specifically, we build upon the framework developed in Forlani et al. (2016) – henceforth FMMM – that allows us to measure heterogeneity in TFP, demand and markups across firms while further providing an exact decomposition of revenue TFP. We employ the FMMM framework to measure these heterogeneities at the firm level and subsequently aggregate them at the location level to analyze differences in TFP, demand and markups across space.

We first highlight two strong patterns in the data relating revenue TFP and density. First, a substantial portion of the revenue productivity advantage of denser areas stems from product composition effects: denser areas are specialised in products generating a higher revenue TFP. Second, the way one aggregates firm-level data into regional-level data matters considerably for the measurement of the elasticity of revenue TFP with respect to density. More specifically, magnitudes are considerably larger when considering a weighted (by firm revenue or employment cost) as opposed to un-weighted data aggregation, while weighted results are also in line with the range suggested by regional-level studies (Rosenthal and Strange, 2004). These patterns are driven by the relationship between firm revenue TFP and firm size (as measured by either revenue or employment cost) being positive in each location but systematically related to density: firms with higher (lower) TFP-R account for a larger (smaller) share of total revenue in denser areas. One way of interpreting this is that the market better allocates market shares across firms with heterogeneous productivities in denser areas so amplifying in aggregate regional-level figures any firm-level differences in productivity across space. These findings have important implications for regional policy. For example, they suggest that achieving regional convergence is not only about improving the TFP or the revenue TFP of firms in lagging regions but also increasing (decreasing) the relative size of the most (least) productive firms in those regions which might be hindered more than in denser regions by factors like inputs misallocation.

Concerning the factors driving the revenue productivity advantage of firms in denser areas that remains after accounting for the product composition and aggregation effects described above, we start by highlighting how a properly defined and measured revenue TFP should equal TFP plus the log price. Using information from the raw data we first document that prices are higher in denser areas. At the same time, quantities sold at these higher prices are also higher and so are revenues. This suggests that products sold by firms located in denser areas face a higher demand. Using measures obtained with the FMMM framework we subsequently establish that marginal costs are higher while markups are lower in denser areas. Furthermore, there is no overall significant relationship between TFP and density and so the revenue TFP advantage of denser areas is mainly driven by higher prices.[1] By using complementary information from exports data, we also provide evidence that prices charged and quantities sold by firms located in denser areas are also higher when conditioning to a given destination market so suggesting that products sold by firms located in denser areas are of higher (actual and/or perceived) quality. The above results have further implications for regional policy. In particular, the current policy approach is based on the presumption that firms in lagging regions are characterized by a lower TFP and so interventions are directed towards increasing their technical efficiency. However, our evidence suggests that interventions should rather promote firms' product quality and marketing capabilities in order to increase revenue TFP in lagging regions.

In terms of data we make use of Eurostat's Products of the European Community (Prodcom) dataset. Prodcom consists of surveys, standardized across the European Union, of firm-level production that cover over 90% of output in manufacturing industries at a detailed (8-digit) level. We use the French Prodcom provided by the Institut National de la Statistique et des Études Économiques (INSEE) for the 2008-15 period. Firm balance sheet and location information comes from the Fichier Approché des Résultats d'Esane (FARE) database and cover the same period 2008-2015. We use Zones d'Emploi (ZE) as our spatial unit, a measure of local labour markets similar in construction to UK Travel-To-Work-Areas.

In order to provide reassurance about the robustness of our results we employ two es-

[1] Our results refer to the aggregate of manufacturing products. Therefore, it might well be the case that, for some specific products, there is a positive and significant relationship between TFP and density. Indeed, we provide one such example in our analysis: 'Ready mixed concrete'.

timation techniques: the estimation procedure developed in FMMM as well the procedure described in De Loecker et al. (2016) – henceforth DGKP. Indeed, under the assumptions laid down in FMMM, the same revenue productivity decomposition holds and both estimation procedures are valid. We find results of both procedures to be qualitatively identical and also quantitatively very similar. We further provide a number of additional results showing that our key findings are little affected by whether we focus on the sample of single-product firms or the larger sample of single-product and multi-product firms, by whether we employ the number of full-time equivalent employees or the total wage bill to measure the labour input, by whether we consider firm revenue or firm wage bill to weight observations, by whether we eliminate the Paris area or not, as well as by whether we use a Cobb-Douglas or a Translog production function.

Our paper is closely related to the literature on the measurement of agglomeration economies. Rosenthal and Strange (2004) and Combes and Gobillon (2015) provide summaries of this literature and agree on a range for the key elasticity of productivity with respect to density of 0.04-0.07.[2] These findings are robust to the endogeneity of current economic density and in particular to the use of long lags of historical density as instruments for current density (Ciccone and Hall, 1996; Ciccone, 2002). However all these findings, including Combes et al. (2012),[3] relate to measures of revenue TFP. By contrast, we use data on quantity, prices and revenue to measure TFP and, via the decomposition provided in FMMM, we unravel the revenue TFP advantage of denser areas into its components.

Our paper is also related to the literature on firm TFP measurement on which Olley and Pakes (1996) has had a deep impact. The key endogeneity issue addressed in Olley and Pakes (1996) is omitted variables: the firm observes and takes decisions based on productivity shocks that are unobservable to the econometrician. Yet, the econometrician observes firm decisions (investments) that do not impact productivity today and that can (under certain conditions) be used as a proxy for productivity shocks. This proxy

[2]See also Combes et al. (2008), Mion and Naticchioni (2009) and De La Roca and Puga (2017) for estimates of the elasticity of worker-level wages with respect to density.

[3]Combes et al. (2012) use revenue TFP measures to establish whether and how the productivity advantage of large cities is due to agglomeration, measured as a right-shifting and dilation of the productivity distribution, as opposed to firm selection, measured as a left-truncation of the productivity distribution. In our analysis we do not explore the issues of right-shifting, dilation and left-truncation. However, our framework could be used to establish if and how much the distribution of each component of revenue TFP is subject to right-shifting and dilation and/or left-truncation.

variable approach to tackle the issue of unobservable productivity shocks has been further developed in Levinsohn and Petrin (2003), Wooldridge (2009) and Ackerberg et al. (2015) while De Loecker et al. (2016) and Forlani et al. (2016) provide frameworks consistent with the presence of heterogeneity across firms in TFP, demand and markups.

The outline of the remainder of this paper is as follows. Section 3.2 provides details on the data we use. Section 3.3 presents the model and revenue TFP decomposition of FMMM while further providing highlights of the estimation procedures. Section 3.4 presents our main results and findings while Section 3.5 contains a number of additional results and robustness checks. Section 3.6 concludes while the Appendix provides additional Tables and details of the estimation procedures.

3.2 Data

This Section describes the data used to study productivity and agglomeration in France. Our analysis focuses on the period 2008-2015. The core data required to estimate firm-level revenue productivity using standard methodologies comprises revenue (and/or value-added), labour, intermediates and capital. For these variables we turn to FARE, an annual census of French firms carried out by INSEE.

From the FARE dataset we take firm labour, intermediates and capital variables. The capital stock variable represents the reported book value of capital while intermediates is the value of intermediate inputs and services. For labour, we use the number of full-time equivalent employees. Some productivity studies use the firm wage bill instead on the grounds that this controls in some way for the ability of workers. We prefer to use the number of full-time equivalent employees as benchmark, while providing additional results obtained using the wage bill, for the following reasons. Our aim is not to establish what share of the productivity advantage of denser areas is related to workers' skills and abilities (possibly due to sorting of better workers across space), but rather to establish how much of the observed revenue-based productivity advantage of firms located in denser areas is due to actual TFP differences as opposed to demand and markups differences. In this light, we prefer to use a measure of the labour input allowing our firm-level revenue TFP and quantity TFP to incorporate differences in workers' skills and abilities across locations. Furthermore, as discussed in Section 3.4, using the number of full-time equivalent employees allows us to more clearly establish whether products sold by firms located in denser locations actually require more inputs to be produced as opposed to more expensive inputs.

FARE can be matched, via the unique firm identifier (SIREN code), to another dataset, the 'Répertoire des entreprises et des établissements', providing us with the location of the establishments of each firm. We use information on the municipality (commune) which we subsequently match to the corresponding 'Zone d'Emploi' (ZE), a measure of local labour markets similar in construction to the UK Travel-To-Work-Areas, of which there are 297 in mainland France (excluding overseas territories and Corsica). In order to give a more causal flavor to our results, in some of our regressions we instrument for current density building on an approach that is standard in the literature: using long-lagged historical densities as instruments for current densities (Combes and Gobillon, 2015). In particular, we use population density in 1831, 1861 and 1891 as our instruments. In doing so, we had to take into account two additional issues. First, historical censuses did not record municipalities which had a population of less than 5,000 in their respective years. At the ZE-level, this still leaves 24 ZEs as having zero population in 1831 so they exert no weight in subsequent regressions. Second, historical municipalities do not exactly match those of today. Several no longer exist having been subsumed over the course of 150 years of administrative changes. We deal with these by manually matching to the modern ZE.

In our investigations, we consider firms as the unit of analysis and restrict our attention to firms whose establishments (if more than one) are all located in the same ZE so that we can uniquely associate a firm to a ZE at a given point in time. In this respect, we believe that the most natural unit of analysis for productivity, demand and markups heterogeneity is the firm and not the establishment. Furthermore, inputs and outputs data are available at the level of the firm and not the establishment and so measuring productivity, demand and markups heterogeneity across establishments would necessarily involve debatable assignment procedures. In doing so, while applying some cleaning to the data,[4] we end up with 628,940 firm-year observations in NACE two-digit industries

[4]We eliminate firm-year observations with non-positive values of revenue and/or intermediates and/or the wage bill and/or capital and/or value added. We then apply a small trimming (top and bottom 0.5%) based on the distribution of the following four ratios: intermediates over sales, wage bill over sales, capital over sales and sales in t over sales in $t-1$. We further apply a final trimming based on the ratio between intermediates plus the wage bill over sales and also drop 2 digit sections with less than 500 observations.

10-32 (Manufacturing), which we label the FARE sample.

Quantity TFP estimation requires data on production quantities and sales values, and that information is available in the Products of the European Community (Prodcom) dataset at a detailed product level. Prodcom is a firm-level survey of manufacturing and production carried out by EU national statistical agencies using an 8-digit nomenclature established by Eurostat. The first four digits correspond to the 'Nomenclature Statistique des Activités Economiques dans la Communauté Européenne' (NACE) revision 2, and the first six digits to the 'Classification of Products by Activity' (CPA) with the last 2 digits adding further detail. There are approximately 3,800 Prodcom codes per year.[5] The Prodcom survey captures at least 90% of production in all the four digit industries covered by the survey.

Table 3.1: CPA 13.10.61: Cotton yarn (other than sewing thread)

PRODCOM	Description
13.10.61.32	Yarn of uncombed cotton, not per retail sale, for woven fabrics (excluding for carpets and floor coverings)
13.10.61.33	Yarn of uncombed cotton, not per retail sale, for knitted fabrics and hosiery
13.10.61.35	Yarn of uncombed cotton, not per retail sale, for other uses (including carpets and floor coverings)
13.10.61.52	Yarn of combed cotton, not per retail sale, for woven fabrics (excluding for carpets and floor coverings)
13.10.61.53	Yarn of combed cotton, not per retail sale, for knitted fabrics and hosiery
13.10.61.55	Yarn of combed cotton, not per retail sale, for other uses (including carpets and floor coverings)
13.10.61.60	Cotton yarn, per retail sale (excluding sewing thread)

Illustrating the advantages of highly disaggregated data, Table 3.1 shows an extract from the 2014 Prodcom list for the six-digit code 13.10.61: 'Cotton yarn (other than sewing thread)'. As can be appreciated from Table 3.1, the eight-digit product breakdown is quite detailed and working at this level of disaggregation allows us to take into account rich differences in technology, demand and degree of competition across finely defined products.

[5]In order to deal with Prodcom codes changing over time we use the correspondence Tables provided by Eurostat RAMON and apply the methodology described in Van Beveren et al. (2012) to construct a time-consistent products breakdown. In practice, from 2008 to 2015 there have been only minor changes in Prodcom codes.

Table 3.2: Summary statistics, €000s

	(1) Fare sample mean (sd)	(2) Prodcom sample mean (sd)	(3) SP+MP sample mean (sd)	(4) SP sample mean (sd)
turnover	3,521.4	7,268.0	4,148.7	6,058.1
	(132,954.4)	(221,462.6)	(43,741.7)	(21,636.6)
value-added	994.1	2,015.9	1,141.7	1,689.1
	(19,578.0)	(26,607.0)	(6,778.2)	(4,493.5)
employees	14.6	29.2	19.9	28.5
	(245.4)	(400.3)	(203.3)	(60.7)
wage bill	718.2	1470.7	1002.5	1380.6
	(15,430.3)	(22,841.9)	(11,517.6)	(3,265.9)
materials	2,527.3	5,252.1	3,007.0	4,369.0
	(115,529.0)	(195,743.2)	(37,627.6)	(18,015.9)
capital	1,992.4	4,128.1	2,071.9	3,337.9
	(93,695.4)	(156,855.7)	(13,595.9)	(15,217.1)
value-added/worker	63.8	68.0	65.0	61.9
	(273.1)	(93.8)	(76.6)	(61.6)
Observation	firm-year	firm-year	firm-prod-year	firm-prod-year
Number of Observations	628,940	201,261	189,017	55,432

The Fare sample includes firms with complete balance sheet data in NACE 2 industries 10-32 that remain after an initial cleaning of the data. The Prodcom sample includes the subset of such firms that are in the Prodcom dataset. In both samples, an observation is a firm-year combination. SP and MP refer to single-product and multi-product firms in the Prodcom sample that have been subject to further data cleaning. We consider two samples: 1) the sample of SP and MP; 2) the sample of SP. In both samples an observation is a firm-product-year combination. For SP a firm-product-year combination corresponds to a unique firm-year combination. Monetary values are in current thousands of euros.

The Fare sample can be matched to Prodcom by means of the unique firm identifier (SIREN code). We label the matched sample, comprising 201,261 firm-year observations, as Prodcom. We subsequently applied the following cleaning procedures:

- Drop products whose unit of measure is not consistent over time
- Drop observations with missing quantity and/or value information
- Drop observations with extreme prices (top and bottom 2.5% within a given 8-digit product)
- Keep firms only if recorded Prodcom sales are consistent with FARE revenues
- We identify single-product firms as those firms that produce only one product or produce a product representing 90% or more of their total sales.
- Drop 2 digit sections with less than 500 single-product firms observations
- Drop observations corresponding to extreme markups values (top and bottom 1%)

This leaves us with a sample of 189,017 (121,004) firm-product-year (firm-year) observations for single-product and multi-product firms combined (SP+MP sample) and 55,432 firm-product-year observations for single-product firms only (SP sample). Clearly, for single-product firms a firm-product-year combination corresponds to a unique firm-year combination. Table 3.2 reports summary statistics for various samples.

Table 3.3: Industry groupings

NACE Industries	Description
13+14+15	Textiles; Wearing apparel; Leather and related products
16+17	Wood and wood products; Paper and paper products
18	Printing and reproduction of recorded media
20+22	Chemicals and chemical products; Rubber and rubber products
23+24	Other non-metallic mineral products; Basic metals
25	Fabricated metal products, except machinery and equipment
26+27+28	Computer, electronic and optical products; Electrical equipment; Machinery and equipment n.e.c.
29+30	Motor vehicles, trailers and semi-trailers; Other transport equipment
31	Furniture
32	Other manufacturing

There are several NACE sections missing from the SP+MP and SP samples. Section 19 (Manufacture of coke and refined petroleum products) is not part of Prodcom. Sections 10-12 (Manufacture of food products, beverages and tobacco products) are covered, and in many countries typically provide both a large number of observations, and contribution to economy-wide production. However, in France the Prodcom data for these sections is collected and stored separately to the main survey and we do not have access to it. We exclude section 21 (Manufacture of pharmaceuticals, medicinal chemical and botanical products) and section 30 (Manufacture of other transport equipment) when dropping sections with less than 500 single-product firms observations. Finally, we apply some aggregation across sections in order to increase the number of observations for industry-specific production function estimations ending up with the industry grouping reported in Table 3.3.

3.3 The MULAMA model: TFP-R decomposed

This Section follows FMMM and in particular we provide here the single-product firm version of the model. See FMMM and Appendix C for the multi-product firm extension of the model. The model is labelled MULAMA because of the names of the 3 heterogeneities it allows for: markups **MU**, demand **LAM**bda and quantity productivity **A**. Crucially, the MULAMA model allows to derive an exact decomposition of revenue-based TFP in terms of the underlying heterogeneities so bridging the gap between quantity TFP and revenue TFP.

In our empirical investigations, we perform estimations and provide results based on both the single-product firms sample and the larger sample of single and multi-product firms. There are pros and cons for each of the two samples. On the one hand, the sample of single and multi-product firms is larger and more representative of the population of French manufacturing firms. On the other hand, using single-product firms requires fewer assumptions in order to measure markups, demand and productivity heterogeneity. In particular, as discussed in more detail in DGKP and FMMM, the key operational issue with multi-product firms is the assignment of inputs to outputs. Produced quantities and generated revenues are observable for the different products of each firm in databases like ours. However, information on inputs used for a specific product is typically not available. Therefore, in order to handle multi-product firms, one needs to lay down additional assumptions in order to solve the problem of assigning inputs to outputs. We provide in Appendix C a full description of the procedure used to assign inputs to outputs. As in DGKP, our procedure first requires estimating the parameters of the production function using single-product firms only.

3.3.1 Demand heterogeneity

In what follows we index firms by i and time by t and denote with lower case the log of a variable (for example r_{it} denotes the natural logarithm of revenue R_{it}). Standard profit maximization (marginal revenue equal to marginal costs) implies that the elasticity of revenue R_{it} with respect to quantity Q_{it} is one over the profit maximizing markup:

$$\frac{\partial r_{it}}{\partial q_{it}} = \underbrace{\frac{\partial R_{it}}{\partial Q_{it}}}_{\text{marginal revenue}} \frac{Q_{it}}{R_{it}} = \underbrace{\frac{\partial C_{it}}{\partial Q_{it}}}_{\text{marginal cost}} \frac{Q_{it}}{P_{it}Q_{it}} = \frac{\frac{\partial C_{it}}{\partial Q_{it}}}{P_{it}} = \frac{1}{\mu_{it}}, \tag{3.1}$$

where $\mu_{it} = \frac{P_{it}}{\frac{\partial C_{it}}{\partial Q_{it}}}$ is the profit maximizing markup. This result comes from static profit maximization and holds under different assumptions about demand (representative consumer and discrete choice models) and product market structure (monopolistic competition, monopoly and standard forms of oligopoly).

Despite the log revenue function, i.e., the function relating log revenue to log quantity, being both unknown and potentially different across firms, equation (3.1) provides us with the slope of the firm-specific log revenue function for firm i while data on the actual log revenue r_{it} and log quantity q_{it} referring to firm i provide us with a point where such a firm-specific log revenue function cuts through the $(q,\, r)$ space. If we now linearize the log revenue function around the observed data point $(q_{it},\, r_{it})$ with a slope given by $\frac{1}{\mu_{it}}$ we can uniquely pin down an intercept for this linearized log revenue function on the r axis. We use such intercept $\tilde{\lambda}_{it}$ as our measure of demand heterogeneity:[6]

$$\tilde{\lambda}_{it} \equiv r_{it} - \frac{\partial r_{it}}{\partial q_{it}} q_{it} = r_{it} - \frac{q_{it}}{\mu_{it}}. \tag{3.2}$$

Given our definition of $\tilde{\lambda}_{it}$ observed firm log revenue is simply

$$r_{it} = \tilde{\lambda}_{it} + \frac{1}{\mu_{it}} q_{it}, \tag{3.3}$$

and so $\tilde{\lambda}_{it}$ is a firm-specific log revenue shifter[7] corresponding to the log price firm i would face if selling one unit of its product.[8]

While being general and intuitive, this measure of demand heterogeneity also maps to more formal and explicit differences in the underlying structure of preferences. In particular, FMMM show that $\tilde{\lambda}_{it} = \frac{\lambda_{it}}{\mu_{it}}$ where λ_{it} is a parameter characterizing differences in

[6] To simplify notation we ignore components that are constant across firms in a given time period or within a product category. Those constants will be captured in our empirical analysis by a suitable set of dummies.

[7] Demand heterogeneity is the variation in revenue that is not explained by variation in quantities, i.e., two firms selling the same quantity but generating a different revenue (because of a different price). Therefore, demand heterogeneity is a firm-specific log revenue shifter given quantity (or equivalently a firm-specific log price shifter given quantity).

[8] At the intercept point $q_{it} = 0$ and so we have $Q_{it} = 1$ from which $R_{it} = P_{it}$ and $r_{it} = p_{it} = \tilde{\lambda}_{it}$. Note this has no implications whatsoever about the presence/absence of a choke price.

utility derived from the consumption of products sold by different firms. More specifically, consider a representative consumer who maximises at each point in time t a differentiable utility function $U(.)$ subject to budget B_t:

$$\max_{Q}\left\{U\left(\tilde{Q}\right)\right\} \text{ s.t. } \int_i P_{it}Q_{it}\mathrm{d}i - B_t = 0$$

where $\tilde{Q}$ is a vector of elements $\Lambda_{it}Q_{it}$ and $\lambda_{it} = \log(\Lambda_{it})$. Therefore, while the representative consumer chooses quantities Q, these quantities enter into the utility function as $\tilde{Q}$ and Λ_{it} can be interpreted as a measure of the perceived quality/appeal of a particular variety. FMMM show that the log revenue function corresponding to the above preferences $r(q_{it}, \lambda_{it})$ can be approximated, around the observed profit-maximizing solution, by the linear function:

$$r_{it} \simeq \frac{1}{\mu_{it}}(q_{it} + \lambda_{it}), \tag{3.4}$$

and so λ_{it} is:

$$\lambda_{it} \simeq \mu_{it}r_{it} - q_{it}. \tag{3.5}$$

Two things are worth nothing at this stage. First, (3.5) is valid as a first-order linear approximation and is the counterpart of (3.2) meaning that the log revenue shifter $\tilde{\lambda}_{it}$, what FMMM label demand heterogeneity, maps via markups into differences in product appeal across firms' varieties $\lambda_{it} = \tilde{\lambda}_{it}\mu_{it}$. Second, while the shape of the function relating revenue to quantity and product appeal will depend upon the specific underlying preferences, FMMM show that (3.4) applies to any preferences structure that can be used to model monopolistic competition and for which a well-behaved differentiable utility function exists.[9] This includes standard CES preferences as well as generalized CES preferences (Spence, 1976)[10], CARA preferences (Behrens et al., 2014), HARA preferences (Haltiwanger et al., 2018), Translog preferences (Feenstra, 2003) as well as the class of Variable Elasticity of Substitution (VES) preferences discussed in Zhelobodko et al. (2012) and Dhingra and Morrow (2019). For example, in the case of CARA preferences, which are non-homothetic, the underlying utility behind heterogeneity in product appeal

[9] FMMM also show λ_{it} is a measure characterizing differences in utility in the oligopoly model developed in Atkeson and Burstein (2008) and further refined in Hottman et al. (2016)

[10] In the case of CES and generalized CES preferences (3.5) holds as an equality because the log revenue function is linear in both q_{it} and λ_{it}.

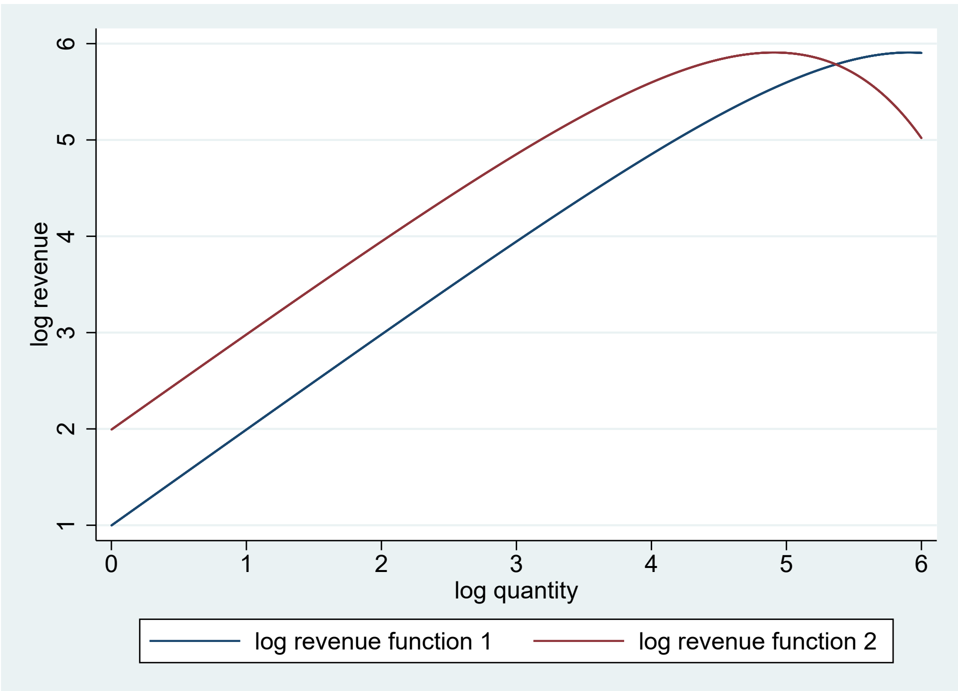
across firms would be:

$$U\left(\tilde{Q}\right) = \int_{\Omega_t} [1 - \mathrm{e}^{-\alpha Q_{it} \Lambda_{it}}] \mathrm{d}i,$$

where Ω_t is the set of varieties available at time t.

Finally, FMMM provide examples suggesting that a log-linear approximation of the revenue function, which is behind both the construction of $\tilde{\lambda}_{it}$ and its interpretation as a markup-adjusted measure of product appeal, works well in many specifications. For example, Figure 3.1 plots two CARA log revenue functions obtained using two different values for product appeal: λ_{it}=1 for log revenue function 1 and λ_{it}=2 for log revenue function 2.[11] As can be appreciated from Figure 3.1, a linear approximation looks both reasonable and accurate for most of the relevant part of the two log revenue functions, i.e., within the range where log revenue (and revenue) is increasing because the marginal revenue is positive and the demand is elastic.

Figure 3.1: CARA log revenue function examples



3.3.2 Markups and marginal costs

As far as markups are concerned FMMM build upon a result, first highlighted in Hall (1986) and implemented in De Loecker and Warzynski (2012) and DGKP among others, based on cost-minimization of a variable input free of adjustment costs (materials in

[11]The other parameters are α=0.001 and the lagrange multiplier κ_t=0.001.

our empirical implementation) and price-taking behaviour on the inputs side (the cost of materials W_{Mit} is allowed to be firm-time specific but it is given to the firm). The proof goes as follows. Starting from the definition of marginal cost:

$$\frac{\partial C_{it}}{\partial Q_{it}} = \frac{\partial C_{it}}{\partial M_{it}}\frac{\partial M_{it}}{\partial Q_{it}} = W_{Mit}\frac{\partial M_{it}}{\partial Q_{it}}.$$

Now define the markup as:

$$\mu_{it} \equiv \frac{P_{it}}{\frac{\partial C_{it}}{\partial Q_{it}}}.$$

We thus have:

$$\frac{P_{it}}{\mu_{it}} = W_{Mit}\frac{\partial M_{it}}{\partial Q_{it}}.$$

Multiplying by Q_{it} and dividing by M_{it} on both sides we get:

$$\frac{P_{it}Q_{it}}{M_{it}\mu_{it}} = \frac{R_{it}}{M_{it}\mu_{it}} = W_{Mit}\frac{\partial M_{it}}{\partial Q_{it}}\frac{Q_{it}}{M_{it}} = W_{Mit}\frac{\partial m_{it}}{\partial q_{it}}.$$

Re-arranging we finally have:

$$\mu_{it} = \frac{\frac{\partial q_{it}}{\partial m_{it}}}{\frac{W_{Mit}M_{it}}{R_{it}}} = \frac{\frac{\partial q_{it}}{\partial m_{it}}}{s_{Mit}}. \tag{3.6}$$

The simple rule to pin-down markups is consistent with many hypotheses on product market structure (monopolistic competition, monopoly and standard forms of oligopoly) and consists in taking the ratio of the output elasticity of materials ($\frac{\partial q_{it}}{\partial m_{it}}$) to the share of materials in revenue ($s_{Mit} \equiv \frac{W_{Mit}M_{it}}{R_{it}}$). Measuring the output elasticity of materials requires estimation of the coefficients of the production function while the share of materials in revenue is directly observable in most datasets (including ours). For example, in the case of a Cobb-Douglas production function with 3 inputs (labour L, materials M and capital K) and with (log) quantity TFP being labeled as a_{it}, log quantity is:

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + a_{it}, \tag{3.7}$$

and so the output elasticity of materials is constant and equal to α_M meaning that $\mu_{it} = \frac{\alpha_M}{s_{Mit}}$. When instead considering a Translog production function log quantity is:

$$q_{it} = \sum_{x\in\{m,l,k\}} \left[\alpha_X x_{it} + \frac{1}{2}\alpha_{XX}(x_{it})^2\right] + \alpha_{MK}m_{it}k_{it} + \alpha_{ML}m_{it}l_{it} + \alpha_{LK}l_{it}k_{it} + a_{it}, \tag{3.8}$$

and so:

$$\mu_{it} = \frac{\alpha_M + \alpha_{MM} m_{it} + \alpha_{ML} l_{it} + \alpha_{MK} k_{it}}{s_{Mit}}.$$

Therefore, with estimates of the production function coefficients at hand, (3.6) can be used to recover firm-specific markups. At the same time, using information on prices and markups, one can recover the marginal cost:

$$MC_{it} = \frac{P_{it}}{\mu_{it}}. \tag{3.9}$$

Finally, with markups as well as log quantity and log revenue, (3.2) can be used to get a measure of demand heterogeneity $\tilde{\lambda}_{it}$.

3.3.3 Quantity TFP

The last step to close the model involves estimating the parameters of the production function and so recover quantity TFP a_{it} and subsequently markups, marginal costs and demand heterogeneity as explained above. There are many different hypotheses, and related estimation procedures, one can use in order to achieve this and in what follows we provide two examples.

One readily available approach to estimate the production function, that is consistent with the underlying presence of heterogeneity in markups and demand, is provided in DGKP. This methodology relies on the popular proxy variable approach pioneered by Olley and Pakes (1996) and in particular, starting from the conditional input demand for materials, adds to such a function a number of observables (prices and market shares in particular) to proxy for unobservables (markups and demand heterogeneity in our framework) while further imposing invertibility of the conditional input demand for materials. More specifically, DGKP build on the GMM approach outlined in Wooldridge (2009) and in particular consider the leading case of an AR(1) process for productivity:

$$a_{it} = \phi_a a_{it-1} + G_{ar} + \nu_{ait}, \tag{3.10}$$

where G_{ar} represents geographical factors affecting productivity (like the density of economic activities),[12] and ν_{ait} stands for productivity shocks that are iid and represent

[12] The index r denotes the region where firm i is located at time t. In our empirical analysis, we use for G_{ar} both the log of the 2009 population and the log of the land area of region r. Given our relatively short time frame (2008-15), it would not make much sense to consider a time-varying population.

innovations with respect to the information set of the firm in $t-1$. Therefore, *productivity shocks* ν_{ait} are uncorrelated with past values of all firm-level variables (capital, revenue, quantity, etc.) including productivity. However, the *productivity level* a_{it} is allowed to be correlated with past and present firm-level variables and in particular is a variable considered by the firm when making choices in t.

Under the (usual) additional assumption that capital is predetermined in t, i.e., capital is chosen beforehand and cannot adjust *immediately* to shocks ν_{ait} occurring in t,[13] the firm will thus consider capital as given in t and will choose the optimal amount of materials in order to minimize costs based on the given values of capital k_{it} and TFP a_{it} as well as the price of materials W_{Mit}. Such optimal amount will in general be a deterministic function $h(.)$ of k_{it}, a_{it} and W_{Mit}. Furthermore, with underlying differences in markups and demand, $h(.)$ will also depend on markups μ_{it} and product appeal λ_{it}. Finally, if labour has also been chosen prior to t (because it is like capital difficult to adjust in the wake of short-term shocks ν_{ait}), then $h(.)$ will also contain l_{it}: $m_{it} = h(k_{it}, l_{it}, a_{it}, W_{Mit}, \mu_{it}, \lambda_{it})$. If $h(.)$ is globally invertible with respect to a_{it}, the inverse function $a_{it} = g(k_{it}, l_{it}, m_{it}, W_{Mit}, \mu_{it}, \lambda_{it})$ exists and is well behaved and so one can use a semi-parametric polynomial approximation of $g(.)$ in order to proxy for the unobservable (to the econometrician) quantity TFP a_{it}. Furthermore, given also W_{Mit}, λ_{it} and μ_{it} are unobservable (to the econometrician), DGKP suggest using regional variables G_r as well as the observable output price and market share of firm i as proxies for W_{Mit}, λ_{it} and μ_{it} in the semi-parametric approximation of $g(.)$,[14] that so becomes a function of observables only. Operationally, $g(.)$ is thus approximated by a polynomial function in the 3 inputs, G_r, the output price and the market share. We provide more details on the DGKP approach and estimation procedure in Appendix A.

Two shortcomings of the DGKP approach are related to its implicit assumptions and the amount of identifying variation. More specifically, existence and invertibility of a suitable conditional input demand for materials implies making implicit assumptions about demand and market structure that are nor readily verifiable. Furthermore, in the main estimation procedure described in DGKP firm market share (de facto firm revenue) and price in $t-1$ are, among other things, added as covariates in a regression where quantity

[13] Capital can nonetheless adjust to shocks ν_{ait} at time t+1.

[14] DGKP cite Kugler and Verhoogen (2011) who document how producers of more expensive products also use more expensive inputs so suggesting that observable output prices could be reasonably used to proxy for unobservable input prices.

at time t in on the left-hand side. Therefore, there might be little variation left to precisely identify technology parameters.[15]

In an attempt to address these two issues FMMM develop an alternative estimation methodology that does not rely on the proxy variable approach. More specifically, FMMM use both the first-order approximation of the log revenue function (3.4) and the production function equation to recover technology parameters. Indeed, FMMM are sufficiently explicit about demand to be able to explicitly write the log revenue function in terms of observables and heterogeneities and use both this and the production function equation to estimate technology parameters. The key disadvantage of this methodology is that one has to be explicit about the process governing the evolution of product appeal λ_{it} and in particular FMMM assume it follows an AR(1) process.[16] In our analysis, we further allow for product appeal to be related to geographical factors G_r^λ which is a straightforward extension of FMMM. More specifically, in our implementation of the FMMM procedure we use:

$$\lambda_{it} = \phi_\lambda \lambda_{it-1} + G_{\lambda r} + \nu_{\lambda it}, \tag{3.11}$$

where $G_{\lambda r}$ represents geographical factors affecting demand (like the density of economic activities),[17] and $\nu_{\lambda it}$ stands for product appeal shocks that are iid and represent innovations with respect to the information set of the firm in $t-1$. However, we do not impose (in line with FMMM) any constraints on the correlation between product appeal shocks $\nu_{\lambda it}$ and quantity TFP shocks ν_{ait} and so ultimately we do not impose a priori any constraints on the correlation between product appeal λ_{it} and quantity TFP a_{it}. Indeed, our results confirm previous findings in FMMM of a negative correlation between product appeal (as well as demand heterogeneity) and quantity TFP irrespective of whether we use the FMMM or the DGKP procedure. This is suggestive of a trade-off between the appeal/perceived quality of a firm's products and their production cost which is in line with findings in the demand system literature (Ackerberg et al., 2007). We provide more

[15] DGKP use the market share in their preferred Translog production function specification. When using a Cobb-Douglas production function, DGKP argue that there is no need to use the market share.

[16] λ_{it} captures consumers' perception of a firm's products quality and appeal; something that arguably does not change much from one year to another. It takes years of effort and costly investments to firms to establish their brand and build their customers' base very much like it takes years of effort and costly investments to firms to put in place and develop an efficient production process for their products. FMMM thus argue that there are profound similarities between the processes of productivity (typically modelled as an autoregressive process) and product appeal.

[17] In our empirical analysis, we use for $G_{\lambda r}$ both the log of the 2009 population and the log of the land area of region r.

details on the FMMM approach and estimation procedure, which builds upon both (3.10) and (3.11), in Appendix A.

3.3.4 TFP-R decomposed

To appreciate how the MULAMA model is useful in linking revenue-based TFP and quantity-based TFP note that, with standard Hicks-neutral TFP, one can write the log of the production function as $q_{it} = \bar{q}_{it} + a_{it}$ where $\bar{q}_{it}$ is an index of inputs use that we label log scale.[18] And by defining revenue TFP as $TFP_{it}^{R} \equiv r_{it} - \bar{q}_{it}$ and using equation (3.3) while substituting we get:

$$TFP_{it}^{R} = \frac{a_{it}}{\mu_{it}} + \tilde{\lambda}_{it} + \frac{1-\mu_{it}}{\mu_{it}}\bar{q}_{it}, \tag{3.12}$$

meaning that TFP_{it}^{R} is a (non-linear) function of quantity-based TFP a_{it}, the log revenue shifter $\tilde{\lambda}_{it}$, the profit-maximizing markup μ_{it} and log production scale $\bar{q}_{it}$. (3.12) can also be made linear by considering markups-adjusted quantity TFP and log scale ($\tilde{a}_{it} = \frac{a_{it}}{\mu_{it}}$ and $\tilde{\bar{q}}_{it} = \frac{(1-\mu_{it})\bar{q}_{it}}{\mu_{it}}$):

$$TFP_{it}^{R} = \tilde{a}_{it} + \tilde{\lambda}_{it} + \tilde{\bar{q}}_{it}, \tag{3.13}$$

so that TFP_{it}^{R} differences across firms located in different regions can be decomposed as the sum of differences in $\tilde{a}_{it}$, $\tilde{\lambda}_{it}$ and $\tilde{\bar{q}}_{it}$ across such firms. In this respect, we note again that while the Urban Economics literature has focused on models featuring differences in quantity TFP across space, the empirical evidence we have gathered so far is at best about revenue TFP and in this respect our framework can shed new light on the determinants of differences in TFP_{it}^{R} across space.

3.3.5 A few last remarks

In our empirical investigations, we perform estimations and provide results based on both the DGKP and FMMM estimation procedures while considering the former as the baseline procedure. In both cases, we consider the Cobb-Douglas production function (3.7) as the leading case while providing some robustness results based on the Translog production function (3.8). In all instances we assume, in light of the features of the heavily regulated French labour market, that labour is predetermined, i.e., it cannot immediately adjust to short-term productivity or demand shocks. Furthermore, we measure the labour input

[18]For example, with a Cobb-Douglas production technology $\bar{q}_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it}$.

with the number of full-time equivalent employees, as in Combes et al. (2012),[19] while providing some robustness results where we use the total wage bill to measure the labour input. Crucially, we will see later on that our key findings are little affected by whether we use the DGKP or the FMMM estimation procedure, by whether we employ the number of full-time equivalent employees or the total wage bill to measure the labour input as well as whether we use a Cobb-Douglas or a Translog production function. Last but not least, we also provide results based on both the single-product firms sample and the larger sample of single and multi-product firms while considering the latter as our preferred sample. Again, our key findings are little affected by which sample we use.

Three last operational issues are worth noting. First, as customary in productivity analyses, we correct (in all estimations) for the presence of measurement error in output (quantity and revenue) and/or unanticipated (to the firm) shocks using the methodology described in DGKP and on which we provide key highlights in Appendix B. Second, we perform TFP estimations separately for each two-digit industry (NACE Sections) and consider a full battery of 8-digit product dummies, as well as year dummies. Indeed, quantity in the data is measured in units (kilograms, litres, number of items, etc.) that are specific to each 8-digit product and so quantity TFP a_{it} can be reasonably compared across firms and space only within an 8-digit product category. For similar reasons, λ_{it} can also be reasonably compared across firms and space only within an 8-digit product category. Therefore, as we discuss in more detail below, our analysis will focus on differences across locations in prices, quantities, quantity TFP, markups, etc. *within 8-digit product categories*. Third, in comparing firm outcomes across space we are faced with the issue of how to deal with firms having more than one establishment. One solution, followed by Combes et al. (2012), is to consider single-establishment firms only. Despite serving the purpose, we believe this strategy is not ideal because it leaves out the group of large multi-establishment firms representing nearly half of employment. Therefore, in our analysis we adopt a different approach. More specifically, we consider firms as the unit of analysis and restrict our attention to firms whose establishments (if more than one) are all located in the same ZE so that we can uniquely associate a firm to a ZE at a given point in time. In this respect, we believe that the most natural unit of analysis for productivity,

[19] More precisely, Combes et al. (2012) use a Cobb-Douglas production function where the labour input is further split into 3 occupational/skill categories each measured in terms of time units.

demand and markups heterogeneity is the firm and not the establishment. Furthermore, inputs and outputs data are available at the level of the firm and not the establishment and so measuring productivity, demand and markups heterogeneity across establishments would necessarily involve debatable assignment procedures.

3.4 Main results

3.4.1 Analysis of the firm-level measures obtained with the MULAMA model

Table 3.4 provides estimates of the coefficients of the Cobb-Douglas production function (3.7) obtained with the DGKP procedure applied the sample of single-product firms (as in DGKP). Coefficient estimates are in line with expectations given a three inputs production function and in particular materials coefficients are larger than labour coefficients which are in turn larger than capital coefficients.[20] Overall, there seems to be evidence of slightly decreasing returns to scale while coefficients are comparable to those reported in FMMM and DGKP using quantity and revenue data for Belgian and Indian firms, respectively.

Table 3.4: DGKP procedure: Cobb-Douglas production function estimations by industry grouping (SP firms only)

Industry	13-15	16-17	18	20-22	23-24	25	26-28	29-30	31	32
Employment	0.160	0.136	0.133	0.185	0.143	0.157	0.169	0.161	0.147	0.152
	(0.021)***	(0.035)***	(0.025)***	(0.018)***	(0.019)***	(0.025)***	(0.016)***	(0.033)***	(0.023)***	(0.038)***
Materials	0.734	0.528	0.741	0.615	0.609	0.641	0.738	0.525	0.645	0.545
	(0.066)***	(0.087)***	(0.187)***	(0.053)***	(0.051)***	(0.036)***	(0.027)***	(0.107)***	(0.050)***	(0.107)***
Capital	0.053	0.060	0.055	0.062	0.026	0.056	0.023	0.065	0.010	0.067
	(0.025)**	(0.013)***	(0.021)***	(0.014)***	(0.006)***	(0.012)***	(0.011)**	(0.018)***	(0.014)	(0.031)**
Returns to scale	0.947	0.725	0.929	0.861	0.778	0.854	0.930	0.751	0.802	0.764
N	1,716	3,279	2,277	4,060	2,749	5,113	3,441	922	1,097	767

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by firm. Regressions include year dummies as well as 8-digit product dummies. See Data Section for industry groupings. Estimations are carried on SP firms only as in DGKP.

We start from the sample of single-product firms and, using materials, labour and capital coefficients from Table 3.4, as well as data on quantity produced and inputs used, we compute quantity TFP a_{it} as a residual from (3.7). Further using the coefficient of materials, as well as the revenue share of materials, we get markups μ_{it} from (3.6). The marginal cost MC_{it} is obtained from (3.9) using prices and markups while demand heterogeneity is computed from (3.2) using markups as well as log quantity and log revenue. Finally, revenue TFP and its components are derived from (3.12) and (3.13). We subsequently

[20] Capital coefficients are on the low side, as it is usually the case in the literature, likely due to measurement error particularly plaguing this variable as discussed in Griliches and Mairesse (1995).

apply the inputs assignment procedure described in Appendix C to allocate inputs across the different products of multi-product firms and use the above equations to obtain quantity TFP, markups, marginal costs, demand heterogeneity, as well as revenue TFP and its components, for each firm-product-year combination. The combined sample (that we label 'SP+MP firms') comprises both single-product and multi-product firms and spans a total of 189,017 firm-product-year observations corresponding to 121,004 unique firm-year combinations.

Table 3.5: DGKP procedure: summary statistics of MULAMA model measures (SP+MP firms)

	mean	sd	p50
TFP-R	2.1243	0.5481	2.0905
TFP a	3.6630	3.3983	3.1337
log revenue shifter $\widetilde{\lambda}$	-1.2898	4.0277	-0.3409
log revenue slope $1/\mu$	0.9410	0.2618	0.9301
markup μ	1.1614	0.3878	1.0751
log marginal cost	-1.6413	3.2204	-1.0117
log scale	4.3092	1.7338	4.3623
log price	-1.5388	3.2173	-0.9191
N		189,017	

Notes: Summary statistics refer to the sample of SP and MP firms. An observation is a firm-product-year combination. For SP a firm-product-year combination corresponds to a unique firm-year combination.

Table 3.5 provides some summary statistics of the various MULAMA model measures for the SP+MP firms sample. For most measures, averages and/or medians are of little value per se and what matters is instead data variation. For revenue TFP we find, in line with FMMM, that MULAMA TFP (TFP-R) is characterized by a standard deviation of about 0.5, which is also in line with the standard deviation of other TFP-R measures obtained from our data.[21] As for the standard deviations of quantity TFP and demand heterogeneity, they are again comparable to results reported in FMMM and much larger, for both quantity TFP and demand heterogeneity, than the standard deviation of TFP-R. Furthermore, there is actually more variation in demand heterogeneity values than quantity TFP values so suggesting that heterogeneity in demand is a key component of firm idiosyncrasies and is at least as sizeable as heterogeneity in productivity. Last but not least, the average markup across observations is 1.161 which compares to a value of

[21] For example, the standard deviation of TFP-R computed following the methodology developed in Wooldridge (2009) on our data is 0.6452.

1.158 obtained by FMMM with data on Belgian firms.

Table 3.6: Some OLS regressions involving quantity TFP, log price, markup and log marginal cost (SP+MP firms)

Dep. var.	TFP	markup	log price
log marginal cost	-0.9463	-0.1192	0.9095
	(0.0021)***	(0.0024)***	(0.0018)***
R^2	0.99	0.34	0.99
N	189,017	189,017	189,017

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by firm. Regressions include year dummies as well as 8-digit product dummies. Estimations are carried on the sample of SP and MP firms.

Table 3.7: Some OLS regressions involving the log revenue shifter $\widetilde{\lambda}$, the markup, log turnover, log marginal cost and log price (SP+MP firms)

Dep. var.	rev. shifter $\widetilde{\lambda}$	markup	log turnover	log marg. cost	log price
TFP	-0.7622	0.1423	0.8453	-0.8276	-0.8311
	(0.0124)***	(0.0014)***	(0.0134)***	(0.0030)***	(0.0031)***
rev. shifter $\widetilde{\lambda}$		0.1341	0.3994	0.0587	0.0519
		(0.0009)***	(0.0094)***	(0.0020)***	(0.0020)***
rev. slope $1/\mu$			5.3231	1.3919	0.1901
			(0.0879)***	(0.0183)***	(0.0183)***
log capital		-0.0249			
		(0.0010)***			
R^2	0.78	0.72	0.58	0.99	0.99
N	189,017	189,017	189,017	189,017	189,017

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by firm. Regressions include year dummies as well as 8-digit product dummies. Estimations are carried on the sample of SP and MP firms.

Tables 3.6 and 3.7 provide a number of OLS regressions suggesting the correlations between the various elements of the MULAMA model are coherent with both intuition and economic theory. For example, column (1) of Table 3.6 provides results of a regression where quantity TFP is regressed on the marginal cost while further considering year dummies as well as 8-digit product dummies and clustering standard errors at the firm level. The coefficient is negative and highly significant, as expected, and quite close to one. Column (2) of Table 3.6 displays results of a similar regression where the dependent variable in now the markup. The coefficient is negative and significant indicating that firms with a lower marginal cost charge a higher markup. In this respect, note that a negative relationship between markups and marginal costs is not a property of any well-behaved preferences structure: it points in the direction of preferences featuring increasing relative love for variety or sub-convexity from which pro-competitive effects come from.[22]

Moving to column (3) of Table 3.6 one can appreciate that prices are increasing with the marginal cost with a pass-through elasticity of about 0.9, which is again in line with results from FMMM. Related to this point, FMMM note that a 0.9 average cost pass-through elasticity might seem too high compared to existing macro evidence (Campa and Goldberg, 2005). However, by looking at detailed product-destination level price and quantity data on French exporters, Berman et al. (2012) provide evidence that standard macro/aggregate measures of pass-through elasticity mask substantial heterogeneity across firms with many firms actually being characterized by a very high pass-through elasticity. More specifically, they show that the pass-through elasticity is decreasing in firm size and productivity with the un-weighted average across firms standing at 0.83 and a near complete pass-through elasticity for smaller and less productive exporters.[23]

In Table 3.7, column (1) provides results of a regression where demand heterogeneity (the revenue shifter $\tilde{\lambda}$) is regressed on quantity TFP while further considering year dummies as well as 8-digit product dummies and clustering standard errors at the firm level. The coefficient is negative and highly significant, as in FMMM, and is suggestive of a trade-off between the appeal/perceived quality of a firm's products and their production cost as indicated in the demand system literature (Ackerberg et al., 2007). Column (2) further indicates that markups are increasing in quantity TFP (again pointing in the direction of preferences featuring increasing relative love for variety or sub-convexity) as well as in the revenue shifter $\tilde{\lambda}$. At the same time, firms with larger investments, i.e., firms with a higher log capital in our regression, tend to charge (for given quantity TFP and demand heterogeneity) lower markups, which is consistent with these firms maximising their profits by selling higher quantities and so facing a more elastic portion of the demand curve. Moving to column (3), one can appreciate that firm revenue is increasing, as it should be, with respect to quantity TFP as well as with the revenue function shifter $\tilde{\lambda}$ and the revenue function slope $1/\mu$. In terms of marginal costs, column (4) indicates that they are, as intuition would suggest, negatively related to TFP also when controlling for

[22] This is also associated to the presence of market distortions such that the market leads to too little selection with respect to the social optimum. See Zhelobodko et al. (2012), Mrázová and Neary (2017) and Dhingra and Morrow (2019) for further details.

[23] Using similar data for Belgium, Amiti et al. (2014) find an un-weighted average pass-through elasticity of 0.80 for Belgian exporters with small exporters displaying a near complete pass-through.

the intercept $\tilde{\lambda}$ and the slope $1/\mu$ of the revenue function. Furthermore, marginal costs are increasing in both $\tilde{\lambda}$ and $1/\mu$ suggesting that firms facing a higher demand curve (because of higher $\tilde{\lambda}$ and/or higher $1/\mu$) do spend more resources to produce their products. Such products are thus likely to be higher quality products from a production point of view and not simply from the view point of consumers' perception. Finally, column (5) shows that prices decrease with quantity TFP while increasing in both $\tilde{\lambda}$ and $1/\mu$, which is what one would expect if our measures capture well what they are supposed to measure.

3.4.2 On the revenue productivity advantage of denser areas: aggregation and product composition

Table 3.8 provides a number of OLS regressions where standard revenue productivity measures at the firm level are regressed on the log of population density of the ZE where firms are located using various samples. More specifically, we use three measures of revenue productivity and four different samples. The three revenue productivity measures are: 1) log value added per worker; 2) revenue TFP obtained as a residual of a three inputs Cobb-Douglas production function estimation where output is measured by revenue and coefficients are estimated via OLS (OLS TFP-R); 3) revenue TFP obtained as a residual of a three inputs Cobb-Douglas production function estimation where output is measured by revenue and coefficients are estimated using the insights provided in Wooldridge (2009) (Wooldridge TFP-R). In terms of samples we use: 1) the FARE sample; 2) the Prodcom sample; 3) the SP+MP firms sample; 4) The SP firms sample. In all regressions, we add time and industry (4-digit) dummies while standard errors are clustered at the ZE level.

Table 3.8: OLS regressions of standard revenue productivity measures on ZE population density (various samples)

Dep. var.	log VA per worker				OLS TFP-R				Wooldridge TFP-R			
	Fare sample	Prodcom sample	SP+MP sample	SP sample	Fare sample	Prodcom sample	SP+MP sample	SP sample	Fare sample	Prodcom sample	SP+MP sample	SP sample
log density	0.0159	0.0313	0.0320	0.0312	0.0062	0.0113	0.0089	0.0107	0.0194	0.0120	0.0088	0.0068
	(0.0039)***	(0.0068)***	(0.0066)***	(0.0066)***	(0.0008)***	(0.0020)***	(0.0016)***	(0.0024)***	(0.0025)***	(0.0031)***	(0.0029)***	(0.0021)***
R^2	0.11	0.10	0.11	0.15	0.82	0.81	0.70	0.67	0.82	0.91	0.84	0.92
N	628,940	201,261	189,017	55,432	628,940	201,261	189,017	55,432	628,940	201,261	189,017	55,432

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the ZE level. Regressions include time and industry (4-digit) dummies. The Fare sample includes firms with complete balance sheet data in NACE 2 industries 10-32 that remain after an initial cleaning of the data. The Prodcom sample includes the subset of such firms that are in the Prodcom dataset. In both samples, an observation is a firm-year combination. SP and MP refer to single-product and multi-product firms in the Prodcom sample that have been subject to further data cleaning. We consider two samples: 1) the sample of SP and MP; 2) the sample of SP. In both samples an observation is a firm-product-year combination. For SP a firm-product-year combination corresponds to a unique firm-year combination.

As one can appreciate, the density elasticity estimate varies depending on which revenue TFP measure is considered. In particular, value added per worker is characterized by somewhat higher coefficients. However, coefficients remain rather stable across samples for a given measure suggesting that focusing, as we do below, on the SP+MP sample or the SP sample does not appear to be particularly at odds with the relationship between revenue TFP and density in wider samples.

In Tables 3.9 and 3.10 we thus focus on the SP+MP sample and run very similar regressions to those performed in Table 3.8. Again we consider the same three revenue productivity measures employed for Table 3.8 while also adding MULAMA revenue TFP (TFP-R). At the same time, we always add year dummies but consider either 2-digit or 8-digit product dummies in order to highlight the importance of product composition in measuring the elasticity of revenue TFP with respect to density. Furthermore, while in Table 3.9 we perform weighted regressions giving equal weight to all firms located in the same ZE (what we label as number of firms weighted),[24] in Table 3.10 we perform weighted regressions giving different weights to firms located in the same ZE depending on their revenue (what we label as revenue weighted).[25] In both cases we shift, by means of regression weighting, the unit of analysis from firms (Table 3.8) to ZEs (Tables 3.9 and 3.10). However, in doing so we either give the same importance to all observations corresponding to a ZE, which means we ultimately compare the average firm across ZEs in the regressions, or we give an importance that is proportional to the revenue share within a ZE, which means our regressions at the firm level should be more comparable to macro/aggregate regressions run at the regional level. Finally, in all regressions we cluster standard errors at the ZE level.

By looking at Tables 3.8 and 3.9 one can draw three conclusions. First, coefficient values are very similar between the two Tables suggesting that whether the unit of analysis is the firm or the average firm in a location does not matter much for the measurement of the relationship between revenue TFP and density. Second, coefficients reported in Table 3.9 and obtained using either 2-digit or 8-digit dummies are very similar suggesting that product composition effects do not play a big role here. Third, coefficients corresponding to the MULAMA revenue TFP (TFP-R) are very much in line with other measures of

[24]In the number of firms weighted case, each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

[25]In the revenue weighted case, each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

revenue TFP (OLS and Wooldridge).

Table 3.9: Revenue productivity, density and product composition effects (number of firms weighted, SP+MP sample)

Dep. var.	log VA per worker		OLS TFP-R		Wooldridge TFP-R		TFP-R	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log density	0.0442	0.0374	0.0120	0.0102	0.0221	0.0174	0.0145	0.0140
	(0.0061)***	(0.0057)***	(0.0015)***	(0.0015)***	(0.0052)***	(0.0037)***	(0.0046)***	(0.0036)***
R^2	0.07	0.20	0.69	0.65	0.81	0.75	0.52	0.70
N	189,017	189,017	189,017	189,017	189,017	189,017	189,017	189,017
2-digit dummies	Yes	No	Yes	No	Yes	No	Yes	No
8-digit dummies	No	Yes	No	Yes	No	Yes	No	Yes

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by ZE. Regressions are weighted and include year dummies as well as either 2-digit or 8-digit product dummies. Estimations are carried on the sample of SP and MP firms. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r. Note that, since regressions use weights, the R^2 does not necessarily improves when considering 8-digit dummies instead of 2-digit dummies.

Table 3.10: Revenue productivity, density and product composition effects (revenue weighted, SP+MP sample)

Dep. var.	log VA per worker		OLS TFP-R		Wooldridge TFP-R		TFP-R	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log density	0.0762	0.0442	0.0171	0.0116	0.0755	0.0369	0.0670	0.0292
	(0.0137)***	(0.0075)***	(0.0027)***	(0.0019)***	(0.0150)***	(0.0077)***	(0.0148)***	(0.0073)***
R^2	0.13	0.45	0.77	0.80	0.88	0.90	0.80	0.92
N	189,017	189,017	189,017	189,017	189,017	189,017	189,017	189,017
2-digit dummies	Yes	No	Yes	No	Yes	No	Yes	No
8-digit dummies	No	Yes	No	Yes	No	Yes	No	Yes

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by ZE. Regressions are weighted and include year dummies as well as either 2-digit or 8-digit product dummies. Estimations are carried on the sample of SP and MP firms. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r. Note that, since regressions use weights, the R^2 does not necessarily improves when considering 8-digit dummies instead of 2-digit dummies.

The comparison of Tables 3.9 and 3.10 is even more interesting and reveals two important results we highlight below:

Result 1: *Weighting impacts the measurement of the elasticity of revenue productivity with respect to density.*

Result 2: *A substantial portion of the aggregate revenue productivity advantage of denser areas stems from product composition effects.*

Regarding Result 1, by simply comparing Table 3.9 and Table 3.10 it appears prominently that coefficients in the latter are larger and this is particularly the case when considering 8-digit dummies. The reason for this behavior lies in the relationship between revenue TFP and revenue. In spatial models à la Melitz (2003) like, for example, Behrens et al. (2017) there is a one to one mapping between firm TFP, as well as revenue TFP, and firm revenue within each location: a firm with higher TFP/revenue TFP will have a higher revenue and so a higher revenue share within a location. However, while the correlation between firm revenue TFP and firm revenue in our data is positive in each and every ZE (ranging between 0.050 and 0.788), it is far from one and systematically related to density. In particular, in denser areas the linear relationship is stronger meaning that firms with higher (lower) TFP-R account for a larger (smaller) share of total revenue in denser regions. One way of interpreting this is that the market better allocates market shares across firms with heterogeneous productivities in denser areas so amplifying in aggregate revenue-weighted figures any firm-level differences in productivity across space.

Regarding Result 2, estimates obtained using 2-digit product dummies are systematically larger, sometimes close to a factor of two, than estimates obtained using 8-digit product dummies and this is particularly the case when considering revenue weighting. This suggests that a considerable portion of the observed aggregate revenue productivity advantage of denser areas comes from these areas being specialised in 8-digit products generating a higher revenue TFP as opposed to denser areas generating a higher revenue TFP for a given 8-digit product.

Tables B.1 to B.4 in Appendix B-4 provide additional evidence of Results 1 and 2 by further looking at other samples: FARE, Prodcom and SP firms. More specifically, Tables B.1 and B.2 perform the very same analysis of Tables 3.9 and 3.10 for SP firms. Table B.3 displays the same regressions reported in Table 3.8 with 2-digit industry dummies and using revenue weighting across all firm samples. In the same vein, Table B.4 covers all firm samples while using 6-digit industry dummies and revenue weighting.[26]

3.4.3 On the revenue productivity advantage of denser areas: demand matters

From now onwards we systematically control for 8-digit product dummies, and so concentrate on the revenue productivity advantage stemming from denser areas generating a higher revenue TFP for a given 8-digit product, while providing both revenue weighted

[26]We use 6-digit industry dummies for all samples instead of 8-digit product dummies because the latter information is not available for firms that are not in Prodcom.

and number of firms weighted results. In particular, we now exploit the valuable information provide by the Prodcom database: quantities and prices. In doing so we more directly move the center of the analysis from firms to locations by aggregating firm-level variables, or more precisely firm-product-year variables, at the ZE level.[27] However, before doing any aggregation, we first demean these variables by 8-digit product and year. For the aggregation, we use either revenue weights or number of firms weights as in the previous Section while using robust standard errors in all ZE level regressions.

More specifically, in order to construct the unique log price measure corresponding to the ZE r, we first subtract from the raw log price information of firms located in the ZE r the corresponding, with respect to the specific product of the firm and the year, mean log price across all locations. We then aggregate up these deviations from 8-digit product and year averages across all firm-product-year observations corresponding to the ZE r using, for example in the case of revenue weights, the revenue share within ZE r corresponding to each observation as weight.[28]

In doing so we thus end up with a unique measure of prices, quantities, and revenues, for each ZE that is consistent across ZEs. We then regress these measures on the log of population density corresponding to each ZE while clustering standard errors at the ZE level. Furthermore, in order to give a more causal flavor to our results, we instrument for current density building on an approach that is standard in the literature: using long-lagged historical densities as instruments for current densities (Combes and Gobillon, 2015). In particular, we use population density in 1831, 1861 and 1891 as our instruments. The corresponding under-identification and weak-identification tests are reported in Tables 3.11 and 3.12 and strongly support the use of such instruments.

The first three columns of Tables 3.11 and 3.12 provide results for log quantity, log revenue and log price, respectively. Note that this part of our analysis simply makes use of raw data and so it is not affected in any way by the possible limitations and restrictions of the MULAMA model. Furthermore also note that, because of the way we constructed variables and the properties of linear estimators, the density coefficient corresponding to

[27]Clearly, given that we use linear models, parameters' estimates and standard errors would be identical if we were to run the same regressions at the firm-product-year level while clustering standard errors at the ZE level.

[28]Formally, our measure of log price is $p_r = \sum_{ipt \in r} (p_{ipt} - \bar{p}_{pt}) w_{ipt}$, where the weight w_{ipt} is either $1/N_r$ or R_{ipt}/R_r.

Table 3.11: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.1454	0.1868	0.0414	0.0462	-0.0048
	(0.0602)**	(0.0556)***	(0.0230)*	(0.0232)**	(0.0035)
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table 3.12: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (number of firms weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.0483	0.0652	0.0169	0.0251	-0.0082
	(0.0273)*	(0.0246)***	(0.0080)**	(0.0082)***	(0.0028)***
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

log revenue is equal to the sum of the density coefficients corresponding to log quantity and log price. In this respect, inspection of Table 3.11 for revenue weighted results and Table 3.12 for number of firms weighted results, reveals another important result:

Result 3: *Prices are higher in denser areas. At the same time, quantities sold at this higher prices are higher too and so are revenues.*

Regarding Result 3, this evidence is present in both revenue weighted results and number of firms weighted results, while being quantitatively stronger in the former. Furthermore, Tables B.10 and B.11 discussed in the next Section also show these patterns hold in the SP firms sample data. Result 3 is consistent with the idea that firms located in denser areas face, on average, higher demand curves than firms located in less dense areas so being able to sell higher quantities even though charging higher prices. Result 3 also has clear and strong implications for the revenue productivity advantage of denser areas. Indeed, from the definition of revenue productivity we have $TFP_{it}^R \equiv r_{it} - \bar{q}_{it} = p_{it} + q_{it} - \bar{q}_{it} = p_{it} + a_{it}$, i.e., revenue TFP is quantity TFP plus the log price. Therefore, even if quantity TFP was on average the same across locations, the fact that firms in denser areas are able to charge higher prices will boost their revenue TFP.

The fact that demand and prices are higher for goods produced in denser regions does not necessarily mean that firms located in such areas sell higher (actual and/or perceived) quality products. For example, in the extreme case where demand is fully local and products are only horizontally differentiated, demand and prices could be higher in denser regions because of the high concentration of service sectors (driven by agglomeration economies) consuming manufacturing products and boosting local wages and consumption. In order to shed light on this issue we provide below two additional pieces of information.

First, in columns (4) and (5) of Tables 3.11 and 3.12 we push the analysis forward by making use of some of the measures obtained from the MULAMA model: log marginal costs and log markups. For an individual firm log price is equal to log marginal cost plus log markup. In our aggregate regressions, because of the way we constructed variables and the properties of linear estimators, the sum of the density coefficients of log marginal cost and log markup equals the density coefficient of log price. In this respect, results provided in Tables 3.11 and 3.12 strongly suggest that the single most important reason why prices are higher in denser areas is because marginal costs are higher. Furthermore, given we use the number of full-time equivalent employees to measure the labour input rather than the wage bill, the fact that marginal costs are higher is not mechanically due to wages being higher in denser areas. As far as log markups are concerned, they are lower in denser areas but significantly so only in the case of number of firms weighted regressions.

The fact that marginal costs are higher in denser areas is in line with the idea that products sold there are of higher actual quality, and so they require more inputs to be produced, but it is not yet a proof. For example, marginal costs could be higher in denser areas simply because firms located there move along an increasing marginal cost

Table 3.13: OLS regression of log marginal costs on TFP and log quantity (SP+MP firms)

Dep. Var.	log marg. cost	log marg. cost
TFP	-1.1765	-1.2098
	(0.0022)***	(0.0048)***
log quantity	0.2116	0.2279
	(0.0016)***	(0.0033)***
Weighting	un-weighted	revenue
R^2	0.9965	0.9986
N	189,017	189,017

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by firm. Regressions include year dummies as well as 8-digit product dummies. Estimations are carried on the sample of SP and MP firms. The first column reports results of an un-weighted OLS regression while column two provides results of a weighted OLS regression where each firm-product-year observation is weighted by R_{ipt} where R_{ipt} is firm i revenue corresponding to product p at time t.

curve in order to meet the requirements of a higher demand rather than having their marginal cost curve upward shifted because of a more expensive and higher quality product being produced. In this respect, Table 3.13 shows the results of a simple OLS regression across firm which is meant to give an idea of by how much marginal costs should be higher in denser areas given the additional quantity sold. More specifically, in order to reconstruct the shape of the log marginal cost curve, in Table 3.13 we regress the log marginal cost corresponding to a firm-product-year observation in the SP+MP sample on the corresponding quantity TFP and log quantity. To show that coefficients are not much affected by firm weighting and/or sample choice we report in column 1 (2) of Table 3.13 simple un-weighted (firm-weighted) results while reporting in Table B.5 of Appendix B-4 both un-weighted and firm-weighted results for the SP firms sample. Turning to Table 3.13, the coefficient of quantity TFP is around -1 and strongly significant which makes sense. As for the coefficient of log quantity, it is around 0.2 indicating that, for example, a 10% higher quantity for given TFP would imply 2% higher marginal costs. In this respect, column 1 of Table 3.11 indicates that doubling density increases quantity sold by 14.54% which should translate, for given TFP and marginal cost curve, into about 3% higher marginal costs. Yet the same Table 3.11 indicates in column 4 that doubling density is associated with a 4.62% higher marginal cost. Repeating the same exercise with Table 3.12 provides an expected, from Table 3.13 and column 1 of Table 3.12, higher marginal cost of about 1% compared to a 2.51% coming from column 4 of Table 3.12. These findings are somewhat supportive of the idea that marginal costs are higher in denser areas compared with what they would be if quantities sold were the same, i.e., that products sold by firms located in denser areas cost more and are of a higher actual quality.

Table 3.14: Exports quantity, revenue and price analysis (export values weighted by ZE, various samples)

	FARE sample			PRODCOM sample			SP+MP sample			SP sample		
Dep. var.	log quantity	log revenue	log price	log quantity	log revenue	log price	log quantity	log revenue	log price	log quantity	log revenue	log price
log density	0.0819	0.1127	0.0308	0.0772	0.1138	0.0366	0.0851	0.1220	0.0369	0.0431	0.1026	0.0595
	(0.0494)*	(0.0492)**	(0.0184)*	(0.0391)**	(0.0466)**	(0.0211)*	(0.0505)*	(0.0493)**	(0.0220)*	(0.0430)	(0.0607)*	(0.0405)*
R^2	0.87	0.77	0.92	0.88	0.80	0.92	0.90	0.82	0.94	0.94	0.89	0.96
N	2,220,803	2,220,803	2,220,803	1,705,905	1,705,905	1,705,905	1,361,002	1,361,002	1,361,002	442,301	442,301	442,301

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the ZE level. We use exports data provided by the French customs. We first match exports data over the period 2008-2014 to the relevant sample data (FARE, Prodcom, SP+MP and SP) and so discard multi-ZE firms from the analysis. We further eliminate observations with missing prices and trim the data based on the top and bottom 1% of the distribution of the demeaned (by HS 8-digit product-country-year) log prices. We also apply a trimming based on the top 3% of the value of exports by ZE. We then use as y variables firm-product-country-year log quantity, log revenue and log price and regress those variables on the log density of the location of the firm along with product-destination-year dummies using the Stata command areg.

The second and more substantial piece of evidence to support the claim that products of firms located in denser regions are of higher perceived/actual quality comes from exports data. Exports represent a substantial portion of French manufacturing firms sales. For example, the overall 2015 goods exports to manufacturing production ratio was 0.7727 while using the sum of manufacturing production and goods imports as denominator delivers a ratio of 0.4187 for 2015. We thus match firm-product-country-year level data on French exporters over the period 2008-2014[29] to our samples and compare quantities, revenues and prices of the same product sold in the same destination and year by firms located in more or less dense areas. We do so for all of the four firm samples we consider in our analysis and overall find a consistent message provided in Table 3.14. More specifically, in Table 3.14 we regress log export quantity, log export revenue and log export price (unit value) on the log density of the location of the firm along with product-destination-year dummies and using revenue weights. Evidence across samples is consistently supportive of products coming from denser areas being sold in higher quantities, despite higher prices, *in the same market.*

Considering all of the above evidence we draw Result 4:

Result 4: *Marginal costs are higher and markups are lower in denser areas. At the same time, marginal costs are higher in denser areas also because of a higher product*

[29]French exporters data for the year 2015 is not available to us. A product is an HS 8-digit code. There are roughly 10,000 such codes.

quality.

3.4.4 On the revenue productivity advantage of denser areas: it is all about demand

Tables 3.15 and 3.16 provide additional insights into the productivity advantage of denser areas by exploiting more measures obtained from the MULAMA model. In particular, columns (1) to (3) report MULAMA revenue TFP (TFP-R), quantity TFP (TFP) and the log price. For an individual firm, revenue TFP is quantity TFP plus the log price. In our aggregate regressions, because of the way we constructed variables and the properties of linear estimators, the sum of the density coefficients of quantity TFP and log price equals the density coefficient of revenue TFP. Results in Tables 3.15 and 3.16 point to the same direction, with findings referring to revenue weights being stronger in magnitude as in the rest of the analysis, and allow establishing a further important result:

Result 5: *The revenue productivity advantage of denser areas is driven by higher prices with no overall significant differences in quantity TFP.*

The picture emerging by combining Results 3 to 5 can be summarized as follows. Manufacturing firms located in denser areas are not necessarily characterized by a significantly higher quantity TFP. They do, however, enjoy a revenue TFP advantage due to their capacity to produce and sell higher demand products at higher prices and in larger quantities compared to firms located in less dense areas. Furthermore, their products are characterized by lower markups and higher marginal costs and part of this higher cost reflects an actual higher product quality.

Additional insights are provided in columns (5) and (7) of Tables 3.15 and 3.16. More specifically, looking at the density coefficients related to the log revenue function intercept $\tilde{\lambda}$ and slope $1/\mu$ reveals that only the latter is significantly and positively related to density suggesting that firms in denser areas face a higher revenue function, i.e, face a higher demand curve, mainly because of a higher slope. Finally, columns (4) to (6) provide results of the revenue TFP decomposition of equation (3.13) with density coefficients of columns (4) to (6) adding up to the density coefficient of revenue TFP in column (1). We already discussed that $\tilde{\lambda}$ is not significantly increasing with density and column (4) points

Table 3.15: 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\bar{q}}$	rev. slope $1/\mu$
log density	0.0418	0.0004	0.0414	0.0365	-0.1079	0.1132	0.0075
	(0.0156)***	(0.0270)	(0.0230)*	(0.0455)	(0.0652)*	(0.0396)***	(0.0037)**
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table 3.16: 2SLS regressions of firm TFP-R and Mulama measures on log density (number of firms weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\bar{q}}$	rev. slope $1/\mu$
log density	0.0149	-0.0020	0.0169	0.0222	-0.0363	0.0290	0.0076
	(0.0042)***	(0.0094)	(0.0080)**	(0.0159)	(0.0231)	(0.0093)***	(0.0025)***
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

to a similar result for markups-adjusted TFP $\tilde{a}$. It is markups-adjusted scale $\tilde{\bar{q}}_{it} = \frac{(1-\mu_{it})\bar{q}_{it}}{\mu_{it}}$ that is significantly higher in denser areas because of firms selling higher quantities and using more inputs, and so having a larger scale, coupled with a higher revenue function slope.[30]

3.4.5 Two examples

Results 1 to 5 refer to the aggregate of manufacturing products. Therefore, it might well be the case that, for some specific products, there is a positive and significant relationship between TFP and density. In this respect, we provide here one such example: 'Ready mixed concrete' (NACE code 2363).[31] Indeed, this particular industry/product has been the object of a number of studies[32] also suggesting that there are significant differences in TFP across space. At the same time, we also provide an example, among many others,

[30] Note that $\frac{(1-\mu_{it})}{\mu_{it}} = \frac{1}{\mu_{it}} - 1$ and so the higher the revenue function slope $\frac{1}{\mu_{it}}$ the higher is markups-adjusted scale.

[31] 'Ready mixed concrete' corresponds to a unique 8-digit Prodcom code.

[32] See, for example, Syverson (2004) and Syverson (2008).

of a particular industry ('Manufacture of other parts and accessories for motor vehicles'; NACE code 293) behaving as the aggregate of manufacturing products.

Table 3.17: 'Ready-mixed concrete' industry (NACE code 2363): 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.5202	0.4959	-0.0243	-0.0080	-0.0163
	(0.2083)**	(0.2091)**	(0.0225)	(0.0253)	(0.0180)
N	123	123	123	123	123
LM stat under-identif.	10.8	10.8	10.8	10.8	10.8
Under-identif. p-value	0.013	0.013	0.013	0.013	0.013
Wald F stat weak identif.	88.2	88.2	88.2	88.2	88.2

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Focusing on the sample of SP and MP firms producing products belonging to the 'Ready-mixed concrete' industry (NACE code 2363), firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations (belonging to the industry 'Ready-mixed concrete') corresponding to the ZE r.

Tables 3.17 and 3.18 provide the same type of information contained in Tables 3.12 and 3.15, but refer to the sub-sample of firm-product-year observations corresponding to the production of 'Ready mixed concrete'.[33] At the same time, Tables 3.19 and 3.20 refer to the sub-sample of firm-product-year observations corresponding to the production of 'Manufacture of other parts and accessories for motor vehicles'.[34] Table 3.12 indicates that, within the 'Ready mixed concrete' sample, firms located in denser areas sell higher quantities and generate higher revenues but do not charge significantly higher or lower prices, while having overall similar marginal costs and markups with respect to firms located in less dense areas. Furthermore, Table 3.15 reveals that 'Ready mixed concrete' firms located in denser areas are characterized by a higher revenue TFP and that this is entirely driven by a higher TFP. At the same time Table 3.19 indicates that, within the 'Manufacture of other parts and accessories for motor vehicles' sample, firms located in denser areas sell higher quantities and generate higher revenues while charging significantly higher prices and having higher marginal costs and lower markups than firms located in less dense areas. Table 3.20 further shows that 'Manufacture of other parts and accessories for motor vehicles' firms located in denser areas are characterized by a higher revenue TFP

[33] There are 726 firm-product-year observations corresponding to 'Ready mixed concrete' distributed across 123 ZEs.

[34] There are 2,036 firm-product-year observations corresponding to 'Manufacture of other parts and accessories for motor vehicles' distributed across 184 ZEs.

and that this is entirely driven by higher prices.

Table 3.18: 'Ready-mixed concrete' industry (NACE code 2363): 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\widetilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.1151	0.1394	-0.0243	0.3195	-0.4780	0.2737	0.0241
	(0.0450)**	(0.0483)***	(0.0225)	(0.1486)**	(0.2794)*	(0.1604)*	(0.0205)
N	123	123	123	123	123	123	123
LM stat under-identif.	10.8	10.8	10.8	10.8	10.8	10.8	10.8
Under-identif. p-value	0.013	0.013	0.013	0.013	0.013	0.013	0.013
Wald F stat weak identif.	88.2	88.2	88.2	88.2	88.2	88.2	88.2

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Focusing on the sample of SP and MP firms producing products belonging to the 'Ready-mixed concrete' industry (NACE code 2363), firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations (belonging to the industry 'Ready-mixed concrete') corresponding to the ZE r.

Table 3.19: 'Manufacture of other parts and accessories for motor vehicles' industry (NACE code 293): 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.3797	0.4772	0.0975	0.1370	-0.0396
	(0.1625)**	(0.1581)***	(0.0552)*	(0.0613)**	(0.0169)**
N	184	184	184	184	184
LM stat under-identif.	21.1	21.1	21.1	21.1	21.1
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	169.1	169.1	169.1	169.1	169.1

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Focusing on the sample of SP and MP firms producing products belonging to the 'Manufacture of other parts and accessories for motor vehicles' industry (NACE code 293), firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations (belonging to the industry 'Manufacture of other parts and accessories for motor vehicles') corresponding to the ZE r.

Table 3.20: 'Manufacture of other parts and accessories for motor vehicles' industry (NACE code 293): 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\widetilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\widetilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0991	0.0016	0.0975	0.1493	-0.4006	0.3504	0.0492
	(0.0437)**	(0.0691)	(0.0552)*	(0.1412)	(0.2073)*	(0.1180)***	(0.0200)**
N	184	184	184	184	184	184	184
LM stat under-identif.	21.1	21.1	21.1	21.1	21.1	21.1	21.1
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	169.1	169.1	169.1	169.1	169.1	169.1	169.1

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Focusing on the sample of SP and MP firms producing products belonging to the 'Manufacture of other parts and accessories for motor vehicles' industry (NACE code 293), firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations (belonging to the industry 'Manufacture of other parts and accessories for motor vehicles') corresponding to the ZE r.

3.5 Robustness checks

Results 1, 2 and 3 do not depend on the Mulama model assumptions and limitations because they are either shown to be consistent across several methodologies (Results 1 and 2) or they come straight from the raw data (Result 3) while holding across several samples and weighting approaches. As for Results 4 and 5, they are instead more reliant on the Mulama model and in this Section we provide a number of additional results showing that Results 4 and 5 are little affected by whether we use the DGKP or the FMMM estimation procedure, by whether we use the single-product firms sample or the larger sample of single-product and multi-product firms, by whether we employ the number of full-time equivalent employees or the total wage bill to measure the labour input, by whether we consider firm revenue or firm wage bill to weigh observations, by whether we include or not the Paris area (specifically the Île de France region), as well as by whether we use a Cobb-Douglas or a Translog production function.[35]

FMMM estimation procedure. Two shortcomings of the DGKP procedure are related to its implicit assumptions and the amount of identifying variation. More specifically, existence and invertibility of a suitable conditional input demand for materials implies making implicit assumptions about demand and market structure that are nor readily verifiable. Furthermore, in the estimation procedure described in DGKP firm market share (de facto firm revenue) and price in $t-1$ are, among other things, added as covariates in

[35]The data samples used below are sometimes slightly different from the one used in the main analysis because of data cleaning and particularly trimming on markups.

a regression where quantity at time t in on the left-hand side. Therefore, there might be little variation left to precisely identify technology parameters.

In an attempt to address these two issues FMMM develop an alternative estimation methodology that does not rely on the proxy variable approach. More specifically, FMMM use both the first-order approximation of the log revenue function (3.4) and the production function equation to recover technology parameters. Indeed, FMMM are sufficiently explicit about demand to be able to explicitly write the log revenue function in terms of observables and heterogeneities and use both this and the production function equation to estimate technology parameters. The key disadvantage of this methodology is that one has to be explicit about the process governing the evolution of product appeal λ_{it} and in particular we, as FMMM, assume it follows an AR(1) process.

Tables B.6 to B.9 in Appendix B-4 provide supporting evidence of Results 4 and 5 obtained using the FMMM procedure.

Single-product firms. The key advantage of using multi-product firms is coverage. Multi-product firms are large and account for the lion's share of manufacturing production. However, their technology needs to be inferred from information on single-product firms (the production function is actually estimated using data on single-product firms only), and assumptions need to be made about how to split inputs across the different products of a multi-product firm.

In order to side-step these limitations, Tables B.10 to B.13 in Appendix B-4 report results referring to the smaller sample of single product firms. Again, evidence is in line with Results 4 and 5.

Using firm wage bill to measure the labour input. Some spatial productivity studies use the firm wage bill instead of the number of full time workers to measure the labour input on the grounds that this controls in some way for the ability of workers. However, our aim is not to establish what share of the productivity advantage of denser areas is related to workers' skills and abilities (possibly due to sorting of better workers across space), but rather to establish how much of the observed revenue-based productivity advantage of firms located in denser areas is due to actual TFP differences as opposed to demand and markups differences. In this light, we prefer to use a measure of the labour input allowing our firm-level revenue TFP and quantity TFP to incorporate differences in workers' skills and abilities across locations. Furthermore, as discussed in Section 3.4, using the number of full-time equivalent employees allows us to more clearly establish whether products sold by firms located in denser locations actually require more inputs to be produced as opposed to more expensive inputs.

We nevertheless provide evidence in Tables B.14 to B.17 in Appendix B-4 that Results 4 and 5 are qualitatively, and to a large extent also quantitatively, unaffected by using the wage bill to measure the labour input.

Using firm wage bill to weigh observations. Using firm revenue to weigh observations is simple and straightforward. However, given that firms generating a similar revenue might generate a very different value added over inputs, statistical offices often prefer to use other approaches when aggregating firm-level data. The most common approach is to consider either the number of employees or the wage bill. In Tables B.18 to B.19 in Appendix B-4 we use the firm wage bill instead of firm revenue to weight observations and in doing so we confirm Results 4 and 5.

Eliminating Paris. When considering the spatial distribution of economic activities and/or regional differences in productivity and wages in France the elephant in the room is the Paris area. To check whether or not our findings are driven by some particular patterns arising in the Paris area we provide in Tables B.20 to B.23 in Appendix B-4 results obtained eliminating firms located in the Île de France region. Again, findings are strongly supportive of Results 4 and 5.

Translog production function. The Cobb-Douglas production function is widely used in productivity analyses including the spatial productivity investigation of Combes et al. (2012). However, the Translog production function is more general albeit more demanding in terms of number of parameters to estimate and degree of analytical complication. In Tables B.24 and B.25 in Appendix B-4 we provide results obtained employing a Translog production function while using product revenue shares in order to assign inputs to the different products of a multi-product firm. Reassuringly, Results 4 and 5 find again strong support.

3.6 Conclusions

We make use of detailed quantity, prices and revenue data on products produced by French manufacturing firms and, building upon FMMM, we quantify heterogeneity in TFP, demand and markups across firms while further providing an exact decomposition of revenue TFP. We measure these heterogeneities at the firm level and subsequently aggregate them at the regional level to analyze differences in TFP, demand and markups across space. We find a number of robust results providing fresh insights on agglomeration economies that have implications for both economic theory and regional policy.

For example, the current policy approach is based on the presumption that firms in lagging regions are characterized by a lower TFP and so interventions are directed towards increasing their technical efficiency. In this respect, our evidence suggests that interventions should rather promote firms' product quality and marketing capabilities in order to increase revenue TFP in lagging regions. Furthermore, our findings suggest that achieving regional convergence has a lot to do with increasing the relative size of the most productive firms in lagging regions which might be hindered more than in other regions by factors like inputs misallocation.

On a concluding note, while our analysis provides a number of fresh insights on agglomeration economies it does not address the old question of what micro-channels generate the observed advantages of denser areas and how important they are individually. However, our analysis does suggest that micro-channels related to product quality and demand are key to understand differences in revenue TFP across space while at the same time highlighting the importance of the largely understudied links between firm revenue TFP, firm size and density in generating aggregate regional-level outcomes. In terms of avenues for future research, we believe the analysis could be fruitfully pushed forward by exploring, along the lines of Combes et al. (2012), if and how much the distribution of each component of revenue TFP is subject to left-truncation (as a measure of the importance of selection) and/or right-shifting and dilation (as a measure of the importance of agglomeration economies).

Chapter 4

The UK's Great Demand Recession

Abstract

We revisit the UK's poor productivity performance since the Great Recession by means of both a suitable theoretical framework and firm-level prices and quantities data for detailed products. This allows us to both measure demand and its changes over time and distinguish between quantity total factor productivity (TFP-Q), i.e., the capacity to turn inputs into more physical output (number of shirts, liters of beer), and what we call revenue total factor productivity (TFP-R), i.e., productivity calculated using (price-index deflated) revenue or value-added as a measure of output and so the capacity to turn inputs into more revenue/value-added. This in turn allows us to measure how changes in TFP-Q, demand and markups ultimately affected revenue TFP, as well as labour productivity, over the Great Recession. Our findings suggest that UK firms' poor productivity performance post-recession is due to both a weakening of demand and a decreasing TFP-Q pushing down sales, markups, revenue TFP and labour productivity.

4.1 Introduction

Nine years after the end of the Great Recession in the UK, labour productivity had barely returned to the level it reached on the eve of the downturn at the end of 2007 (Office for National Statistics, 2018b). Output per hour worked grew just 1.8% between the start of 2008 and the end of 2017: had it grown at its 1994-2007 trend, it would have been 19.6% higher.[1] This poor performance is a puzzle. A sustained period of little to no labour productivity growth following a recession is indeed rare in the UK's historical record.[2] Furthermore, the productivity slowdown has occurred despite a buoyant labour market,[3] and the UK's experience is widely judged to have been worse than most of its EU and OECD peers.[4] Despite many complementary explanations that have so far been put forward, little attention has been paid to the role of demand and markups as well as to the crucial distinction between quantity total factor productivity (TFP-Q), i.e., the capacity to turn inputs into more physical output (number of shirts, litres of beer), and what we call revenue total factor productivity (TFP-R), i.e., productivity calculated using (price-index deflated) revenue or value-added as a measure of output and so the capacity to turn inputs into more revenue/value-added.

Regarding the role of demand, the problem is that without actual data on products' prices and quantities it is not possible to measure demand and its changes over time and so assess whether, and to what extent, a fall in demand might have contributed to the UK productivity slowdown. For example, if revenues increase less than the price index one might well conjecture that the underlying unobservable quantities sold have decreased, but it would not be possible to establish whether the decrease in quantities is simply due to price changing, and firms moving along the same demand curve, or the underlying demand curve firms face has changed. At the same time, the unavailability of data on products' prices and quantities does not allow properly distinguishing between quantity TFP and revenue TFP. To be more specific, both macro and micro productivity studies use price

[1] Our period of analysis ends in 2013 and at that point, output per hour was 1.7% lower than at the end of 2007. Had it grown at its 1994-2007 trend rate, it would have been 14.4% higher.

[2] Four years after the end of the recessions that began in 1973, 1980 and 1990, labour productivity was between 5% and 15% higher than its previous peak (Grice, 2012).

[3] See, for example, Bryson and Forth (2016) and Pessoa and Van Reenen (2014).

[4] See, for example, OECD (2018) and Office for National Statistics (2018a). There is evidence that the productivity slowdown in the US and major European economies pre-dates the financial crisis (Cette et al., 2016) but the UK experience, as documented by Office for National Statistics (2018b), shows a marked slowdown of the productivity growth trend in the recovery from the 2008-9 recession.

indices to deflate nominal sales, or value-added, in order to measure output and its changes over time. Besides the usual caveats of aggregation bias and the presence of substantial price heterogeneity across firms, the issue with this approach is that actual price changes do not correspond to price index changes due to, for example, the imputation of part of actual price changes to quality improvements to products or changes to the specification of a product. Therefore, standard price-index deflated nominal values still contain a price measure and that is why we label productivity measures obtained with this approach as revenue-TFP measures.

In this paper we provide novel evidence that UK firms' poor productivity performance post-2008 is due to both weakening demand and sluggish TFP-Q growth pushing down sales, markups and revenue TFP, as well as labour productivity. More specifically, in the first part of our analysis we focus on manufacturing firms and use information on firm-level prices and quantities for detailed products, from the UK Prodcom dataset, as well as inputs over the period 2003-2013[5] to measure firm-level quantity TFP by building upon the frameworks developed in De Loecker et al. (2016) and Forlani et al. (2016). This allows us to quantify firm-level markups, as well as firm-level demand and its changes over time and, while aggregating-up the information to the whole-of-manufacturing level, compare the evolution of TFP-Q, markups and demand before and after 2008. Finally, we exploit two exact decompositions for TFP-R and labour productivity to show how changes in TFP-Q, markups and demand have affected the two productivity measures. Our results strongly indicate that both a slowing down of demand and a decline in quantity TFP, and the related fall in markups, are behind the decline in revenue TFP and labour productivity in manufacturing. Furthermore, we show that the difference between actual price changes and price index changes, that we label 'real price changes' and that reflect more than just quality improvements to products, widened post-2008 in response to the increasing production costs generated by a lower TFP-Q, so helping to contain the fall in TFP-R.

In the second part of our analysis, we consider service industries and estimate a restricted version of the model due to the absence of reliable and meaningful information on prices and quantities. In doing so we find, for those measures that are common to both the full and restricted versions of the model, very similar patterns to those obtained for

[5]We consider the time frame 2003-2013 in our analysis for better comparability with previous studies on the UK productivity puzzle.

manufacturing. These findings, along with the absence of noticeable differences in capital stock investment patterns between manufacturing and services industries, lead us to conjecture that both demand and TFP-Q are also responsible for the poor revenue TFP and labour productivity performance of UK service industries.

We believe that our results are important for at least two reasons. First, they are informative about the long-term impacts of the Great Recession. A fall in quantity TFP, due for example to a decline in the rate of technical progress, represents a permanent loss of productive potential with substantial long-term implications for the economy. By contrast a demand downturn, due for example to a prolonged general climate of uncertainty, could have less permanent consequences. Second, our results are informative about the policies that could more effectively address the weak growth of labour productivity and revenue TFP post-2008. In particular, our findings suggest that government policies should more prominently act towards boosting demand for UK firms rather than focusing only on productivity.

Our paper is related to the literature devoted to the UK productivity puzzle.[6] This literature has so far considered many complementary reasons for the poor post-2008 performance relative to the long-term trend. Among these are: measurement errors in output (Goodridge et al., 2016; Grice, 2012); productivity losses in specific sectors (Riley et al., 2018); labour hoarding (Martin and Rowthorn, 2012); capital shallowing (Pessoa and Van Reenen, 2014; Goodridge et al., 2016; Riley et al., 2018); the impact of badly-measured intangible capital (Goodridge et al., 2013); changes to firm entry and exit behaviour in the context of an impaired financial sector (Barnett et al., 2014; Riley et al., 2013, 2014); a lengthening of the left tail of poorly performing firms in the productivity distribution (Andrews et al., 2015); and a slowdown among high-performing firms in the right tail of the distribution (Schneider, 2018). However, while there are many proposed culprits and some fit certain features of the puzzle better than others, there is lack of consensus on some key elements of the productivity downturn,[7] while we know little about to what

[6] See Bryson and Forth (2016) and McCann (2018) for a literature review.

[7] For example, in manufacturing there is not a consensus on whether the labour productivity puzzle is also a total factor productivity puzzle. More specifically, Goodridge et al. (2016) build upon an aggregate-level growth accounting approach and find that labour productivity in manufacturing has declined because of a decline in total factor productivity. By contrast, Harris and Moffat (2017) build upon a firm-level approach and find that the labour productivity puzzle in manufacturing is mainly driven by a decline in intermediates intensity while TFP growth continued.

extent the puzzle is demand- and/or supply-driven and what is the macro role played by markups (De Loecker et al., 2020).

Our paper is also related to the literature on heterogeneous markups and productivity inspired by Hall (1986) and Olley and Pakes (1996) and further developed in Ackerberg et al. (2015), De Loecker et al. (2016) and Forlani et al. (2016). More specifically, in our analysis we estimate a quantity-based production function for UK manufacturing using two estimation procedures, the one developed in De Loecker et al. (2016) (henceforth DGKP) and the one described in Forlani et al. (2016) (henceforth FMMM). These two methods are similar in their motivation to disentangle heterogeneity in revenue TFP into supply-side differences between firms, notably TFP-Q, from demand-side differences in prices which could be due to differences in input and/or output quality, demand and markups. As for services, we instead estimate revenue-based production functions by building on either the restricted version of the model introduced in FMMM or the more standard Wooldridge (2009) approach (henceforth WLD). Again, our results are largely unaffected by whether we use one or the other estimation method. For both manufacturing and services we use the TFP-R decomposition provided in FMMM to break down revenue productivity changes into underlying changes in TFP-Q and demand (for services, a composite of TFP-Q and demand) as well as markups and production scale. We further develop a labour productivity decomposition generalizing the standard formula used in growth accounting exercises to the presence of heterogeneity in demand and markups. Finally, we show that our results are robust to whether we use a Cobb-Douglas or a Translog production function, to different weighting schemes (revenue, employment, un-weighted) and samples (single-product vs. multi-product firms) as well as different estimation procedures.

The remainder of this paper is structured as follows. Section 4.2 presents key highlights of the FMMM model and related TFP-R and labour productivity decompositions. Section 4.3 describes the datasets used and provides some summary statistics and high-level patterns while Section 4.4 contains production function estimation results and develops a number of useful insights into how to interpret the results. Section 4.5 presents baseline results for both manufacturing and services while Section 4.6 contains a number of additional findings showing the robustness of our results. Finally, Section 4.7 concludes. Further details about the data and robustness findings are provided in Appendix C.

4.2 The MULAMA model

This Section follows FMMM and in particular we provide here the single-product firm version of the model. See FMMM for the multi-product firm extension of the model. The model is labelled MULAMA because of the names of the 3 heterogeneities it allows for: markups **MU**, demand **LAM**bda and quantity TFP **A**. FMMM also provide an estimation procedure to quantity markups, demand and quantity TFP that we employ in our analysis. At the same time, the estimation procedure developed in DGKP is also consistent with the MULAMA model and we employ that estimation procedure to corroborate the robustness of our findings. Furthermore, the MULAMA model allows for an exact decomposition of revenue TFP in terms of the underlying heterogeneities. In addition, we develop below a decomposition of labour productivity generalizing the standard formula used in growth accounting exercises to the presence of heterogeneity in demand and markups.

4.2.1 Measuring demand

In what follows we index firms by i and time by t and denote with lower case the log of a variable (for example r_{it} denotes the natural logarithm of revenue R_{it}). Standard profit maximization (marginal revenue equal to marginal costs) implies that the elasticity of revenue R_{it} with respect to quantity Q_{it} is one over the profit maximizing markup:

$$\frac{\partial r_{it}}{\partial q_{it}} = \underbrace{\frac{\partial R_{it}}{\partial Q_{it}}}_{\text{marginal revenue}} \frac{Q_{it}}{R_{it}} = \underbrace{\frac{\partial C_{it}}{\partial Q_{it}}}_{\text{marginal cost}} \frac{Q_{it}}{P_{it}Q_{it}} = \frac{\frac{\partial C_{it}}{\partial Q_{it}}}{P_{it}} = \frac{1}{\mu_{it}}, \tag{4.1}$$

where $\mu_{it} = \frac{P_{it}}{\frac{\partial C_{it}}{\partial Q_{it}}}$ is the profit maximizing markup. This result comes from static profit maximization and holds under different assumptions about demand (representative consumer and discrete choice models) and product market structure (monopolistic competition, monopoly and standard forms of oligopoly).

Despite the log revenue function, i.e., the function relating log revenue to log quantity, being both unknown and potentially different across firms, equation (4.1) provides us with the slope of the firm-specific log revenue function while data on the actual log revenue r_{it} and log quantity q_{it} referring to firm i provide us with a point where such log revenue function cuts through the (q, r) space. If we now linearize the log revenue function around the observed data point (q_{it}, r_{it}) with a slope given by $\frac{1}{\mu_{it}}$ we can uniquely pin down an intercept for this linearized log revenue function on the r axis. We use such intercept $\tilde{\lambda}_{it}$ as a measure of the firm-specific demand:[8]

$$\tilde{\lambda}_{it} \equiv r_{it} - \frac{\partial r_{it}}{\partial q_{it}} q_{it} = r_{it} - \frac{q_{it}}{\mu_{it}}. \tag{4.2}$$

Given our definition of $\tilde{\lambda}_{it}$ observed firm log revenue is simply

$$r_{it} = \tilde{\lambda}_{it} + \frac{1}{\mu_{it}} q_{it}, \tag{4.3}$$

and so $\tilde{\lambda}_{it}$ is a firm-specific log revenue shifter corresponding to the log price firm i would face if selling one unit of its product.[9]

While being general and intuitive, this measure of firm-specific demand also maps to more formal and explicit differences in the underlying structure of preferences. In particular, FMMM show that $\tilde{\lambda}_{it} = \frac{\lambda_{it}}{\mu_{it}}$ where λ_{it} is a parameter characterizing differences in utility derived from the consumption of products sold by different firms. More specifically, consider a representative consumer who maximises at each point in time t a differentiable utility function $U(.)$ subject to budget B_t:

$$\max_{Q} \left\{ U\left(\tilde{Q}\right) \right\} \text{ s.t. } \int_i P_{it} Q_{it} \mathrm{d}i - B_t = 0$$

where $\tilde{Q}$ is a vector of elements $\Lambda_{it} Q_{it}$ and $\lambda_{it} = \log(\Lambda_{it})$. Therefore, while the representative consumer chooses quantities Q, these quantities enter into the utility function as $\tilde{Q}$ and Λ_{it} can be interpreted as a measure of the perceived quality/appeal of a particular variety. In our analysis we employ λ_{it} as a complementary measure of the firm-specific demand and sometimes refer to $\tilde{\lambda}_{it}$ as markup-adjusted demand.[10]

[8] To simplify notation we ignore components that are constant across firms in a given time period or within a product category. Those constants will be captured in our empirical analysis by a suitable set of dummies.

[9] At the intercept point $q_{it} = 0$ and so we have $Q_{it} = 1$ from which $R_{it} = P_{it}$ and $r_{it} = p_{it} = \tilde{\lambda}_{it}$. Note this has no implications whatsoever about the presence/absence of a choke price.

[10] The interpretation of Λ_{it} as a utility shifter and its relationship with the firm log revenue function are based on a first-order linear approximation around the profit maximizing solution, i.e., $r_{it} \simeq \frac{1}{\mu_{it}}(q_{it} + \lambda_{it})$. In this respect, FMMM show that such linear approximation applies to any preferences structure that can be used to model monopolistic competition, and for which a well-behaved differentiable utility function exists, as well as to the oligopoly model developed in Atkeson and Burstein (2008) and further refined in Hottman et al. (2016). This includes standard CES preferences as well as generalized CES preferences (Spence, 1976), CARA preferences (Behrens et al., 2014), HARA preferences (Haltiwanger et al., 2018), Translog preferences (Feenstra, 2003) as well as the class of Variable Elasticity of Substitution (VES) preferences discussed in Zhelobodko et al. (2012) and Dhingra and Morrow (2019). Finally, FMMM provide examples suggesting that a log-linear approximation of the revenue function, which is behind both the construction of $\tilde{\lambda}_{it}$ and its interpretation as a markup-adjusted measure of product appeal, works well for many utility specifications.

4.2.2 Measuring Markups

As far as markups are concerned FMMM build upon a result, first highlighted in Hall (1986) and implemented in De Loecker and Warzynski (2012) and DGKP among others, based on cost-minimization of a variable input free of adjustment costs (materials in our empirical implementation) and price-taking behaviour on the inputs side (the cost of materials W_{Mit} is allowed to be firm-time specific but it is given to the firm). The proof goes as follows. Starting from the definition of marginal cost:

$$\frac{\partial C_{it}}{\partial Q_{it}} = \frac{\partial C_{it}}{\partial M_{it}} \frac{\partial M_{it}}{\partial Q_{it}} = W_{Mit} \frac{\partial M_{it}}{\partial Q_{it}}.$$

Now define the markup as:

$$\mu_{it} \equiv \frac{P_{it}}{\frac{\partial C_{it}}{\partial Q_{it}}}.$$

We thus have:

$$\frac{P_{it}}{\mu_{it}} = W_{Mit} \frac{\partial M_{it}}{\partial Q_{it}}.$$

Multiplying by Q_{it} and dividing by M_{it} on both sides we get:

$$\frac{P_{it} Q_{it}}{M_{it} \mu_{it}} = \frac{R_{it}}{M_{it} \mu_{it}} = W_{Mit} \frac{\partial M_{it}}{\partial Q_{it}} \frac{Q_{it}}{M_{it}} = W_{Mit} \frac{\partial m_{it}}{\partial q_{it}}.$$

Re-arranging we finally have:

$$\mu_{it} = \frac{\frac{\partial q_{it}}{\partial m_{it}}}{\frac{W_{Mit} M_{it}}{R_{it}}} = \frac{\frac{\partial q_{it}}{\partial m_{it}}}{s_{Mit}}. \tag{4.4}$$

The simple rule to pin-down markups is consistent with many hypotheses on product market structure (monopolistic competition, monopoly and standard forms of oligopoly) and consists in taking the ratio of the output elasticity of materials ($\frac{\partial q_{it}}{\partial m_{it}}$) to the share of materials in revenue ($s_{Mit} \equiv \frac{W_{Mit} M_{it}}{R_{it}}$). Measuring the output elasticity of materials requires estimation of the coefficients of the production function while the share of materials in revenue is directly observable in most datasets (including ours). For example, in the case of a Cobb-Douglas production function with 3 inputs (labour L, materials M and capital K) and with (log) quantity TFP being labeled as a_{it}, log quantity is:

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + a_{it}, \tag{4.5}$$

and so the output elasticity of materials is constant and equal to α_M meaning that $\mu_{it} = \frac{\alpha_M}{s_{Mit}}$. When instead considering a Translog production function log quantity is:

$$q_{it} = \sum_{x \in \{m,l,k\}} \left[\alpha_X x_{it} + \frac{1}{2} \alpha_{XX} (x_{it})^2 \right] + \alpha_{MK} m_{it} k_{it} + \alpha_{ML} m_{it} l_{it} + \alpha_{LK} l_{it} k_{it} + a_{it}, \quad (4.6)$$

and so:

$$\mu_{it} = \frac{\alpha_M + \alpha_{MM} m_{it} + \alpha_{ML} l_{it} + \alpha_{MK} k_{it}}{s_{Mit}}.$$

Therefore, with estimates of the production function coefficients at hand, (4.4) can be used to recover firm-specific markups. At the same time, with markups as well as log quantity and log revenue, (4.2) can be used to get the demand measures $\tilde{\lambda}_{it}$ and λ_{it}.

4.2.3 Quantity TFP

The last step to close the model involves estimating the parameters of the production function and so recover quantity TFP a_{it} and subsequently markups and demand as explained above. There are many different hypotheses, and related estimation procedures, one can use in order to achieve this and in what follows we describe and employ two techniques.

One readily available approach to estimate the production function, that is consistent with the MULAMA model, is provided in DGKP. This methodology relies on the popular proxy variable approach pioneered by Olley and Pakes (1996) and in particular, starting from the conditional input demand for materials, adds to such function a number of observables (prices and market shares in particular) to proxy for unobservables (markups and demand heterogeneity in our framework) while further imposing invertibility of the conditional input demand for materials. More specifically, DGKP build on the GMM approach outlined in WLD and in particular consider the leading case of an AR(1) process for productivity:

$$a_{it} = \phi_a a_{it-1} + \nu_{ait}, \quad (4.7)$$

where ν_{ait} stands for productivity shocks that are iid and represent innovations with respect to the information set of the firm in $t-1$. Therefore, *productivity shocks* ν_{ait} are uncorrelated with past values of all firm-level variables (capital, revenue, quantity, etc.) including productivity. However, the *productivity level* a_{it} is allowed to be correlated with past and present firm-level variables and in particular is a variable considered by the firm when making choices in t.

Under the (usual) additional assumption that capital is predetermined in t, i.e., capital is chosen beforehand and cannot adjust *immediately* to shocks ν_{ait} occurring in t,[11] the firm will thus consider capital as given in t and will choose the optimal amount of materials in order to minimize costs based on the given values of capital k_{it} and TFP a_{it} as well as the price of materials W_{Mit}. Such optimal amount will in general be a deterministic function $h(.)$ of k_{it}, a_{it} and W_{Mit}. Furthermore, with underlying differences in markups and demand, $h(.)$ will also depend on markups μ_{it} and demand λ_{it}. Finally, if labour has also been chosen prior to t (because it is like capital difficult to adjust in the wake of short-term shocks ν_{ait}),[12] then $h(.)$ will also contain l_{it}: $m_{it} = h(k_{it}, l_{it}, a_{it}, W_{Mit}, \mu_{it}, \lambda_{it})$. If $h(.)$ is globally invertible with respect to a_{it}, the inverse function $a_{it} = g(k_{it}, l_{it}, m_{it}, W_{Mit}, \mu_{it}, \lambda_{it})$ exists and is well behaved and so one can use a semi-parametric polynomial approximation of $g(.)$ in order to proxy for the unobservable (to the econometrician) quantity TFP a_{it}. Furthermore, given also W_{Mit}, λ_{it} and μ_{it} are unobservable (to the econometrician), DGKP suggest using the observable output price and market share of firm i as proxies for W_{Mit}, λ_{it} and μ_{it} in the semi-parametric approximation of $g(.)$,[13] that so becomes a function of observables only. Operationally, $g(.)$ is thus approximated by a polynomial function in the 3 inputs plus the output price and the market share.

Two shortcomings of the DGKP approach are related to its implicit assumptions and the amount of identifying variation. More specifically, existence and invertibility of a suitable conditional input demand for materials implies making implicit assumptions about demand and market structure that are nor readily verifiable. Furthermore, in the estimation procedure described in DGKP firm market share (de facto firm revenue) and price in $t-1$ are, among other things, added as covariates in a regression where quantity at time t in on the left-hand side. Therefore, there might be little variation left to precisely identify

[11] Capital can nonetheless adjust to shocks ν_{ait} at time t+1.

[12] Given the features of the UK labour market we do not assume, as DGKP do for India and we do in Chapter 3 for France, that labour is, like capital, predetermined for the firm in t. More specifically we consider, in both the DGKP and FMMM estimation procedures, that labour is a semi-flexible input that is chosen at time $t-b$ $(0 < b < 1)$, i.e., it can be adjusted in light of contemporaneous shocks in t but not as fully as materials. In practice, this means we consider it as an endogenous variable, and use its lags as instruments in the estimation, but do not consider it flexible enough to use in the calculation of markups. See FMMM for more details on how to deal with semi-flexible inputs.

[13] DGKP cite Kugler and Verhoogen (2011) who document how producers of more expensive products also use more expensive inputs so suggesting that observable output prices could be reasonably used to proxy for unobservable input prices.

technology parameters.

In an attempt to address these two issues FMMM develop an alternative estimation methodology that does not rely on the proxy variable approach. More specifically, FMMM use both the first-order approximation of the log revenue function and the production function to recover technology parameters. Indeed, FMMM are sufficiently explicit about demand to be able to explicitly write the log revenue function in terms of observables and heterogeneities and use both this and the production function equation to estimate technology parameters. The key disadvantage of this methodology is that one has to be explicit about the process governing the evolution of demand λ_{it} and in particular FMMM assume it follows an AR(1) process.[14] In our implementation of the FMMM procedure we use:

$$\lambda_{it} = \phi_\lambda \lambda_{it-1} + \nu_{\lambda it}, \tag{4.8}$$

where $\nu_{\lambda it}$ stands for demand shocks that are iid and represent innovations with respect to the information set of the firm in $t-1$. However, FMMM do not impose any constraints on the correlation between demand shocks $\nu_{\lambda it}$ and quantity TFP shocks ν_{ait} and so ultimately do not impose a priori any constraints on the correlation between demand λ_{it} and quantity TFP a_{it}. Indeed, our analyses confirm previous findings in FMMM of a negative correlation between λ_{it} and a_{it} irrespective of whether we use the FMMM or the DGKP procedure. This is suggestive of a trade-off between the appeal/quality of a firm's products and their production cost which in line with findings in the demand system literature (Ackerberg et al., 2007).

4.2.4 TFP-R decomposed

To appreciate how the MULAMA model is useful in linking revenue TFP and quantity TFP note that, with standard Hicks-neutral TFP, one can write the log of the production function as $q_{it} = \bar{q}_{it} + a_{it}$ where $\bar{q}_{it}$ is an index of inputs use that we label scale.[15] Revenue TFP is simply log revenue minus scale $TFP^R_{it} \equiv r_{it} - \bar{q}_{it} = a_{it} + p_{it}$, and it is also equal to

[14] λ_{it} captures consumers' perception of a firm's products quality and appeal; something that arguably does not change much from one year to another. It takes years of effort and costly investments to firms to establish their brand and build their customers' base very much like it takes years of effort and costly investments to firms to put in place and develop an efficient production process for their products. FMMM thus argue that there are profound similarities between the processes of productivity (typically modelled as an autoregressive process) and product appeal.

[15] For example, with a Cobb-Douglas production technology $\bar{q}_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it}$.

quantity TFP plus log price. Using equation (4.3) to substitute for r_{it} along with $\tilde{\lambda}_{it} = \frac{\lambda_{it}}{\mu_{it}}$ we get:

$$TFP_{it}^{R} = \frac{a_{it}}{\mu_{it}} + \frac{\lambda_{it}}{\mu_{it}} + \frac{1-\mu_{it}}{\mu_{it}}\bar{q}_{it}, \tag{4.9}$$

meaning that TFP_{it}^{R} is a (non-linear) function of quantity-based TFP a_{it}, demand λ_{it}, the markup μ_{it} and production scale $\bar{q}_{it}$. (4.9) can also be made linear by considering markup-adjusted quantity TFP and scale ($\tilde{a}_{it} = \frac{a_{it}}{\mu_{it}}$ and $\tilde{\bar{q}}_{it} = \frac{(1-\mu_{it})\bar{q}_{it}}{\mu_{it}}$):

$$TFP_{it}^{R} = \tilde{a}_{it} + \tilde{\lambda}_{it} + \tilde{\bar{q}}_{it}, \tag{4.10}$$

so that TFP_{it}^{R} differences across firms and time can be decomposed as the sum of differences in $\tilde{a}_{it}$, $\tilde{\lambda}_{it}$ and $\tilde{\bar{q}}_{it}$. In particular, using Δ to denote changes between $t-1$ and t:

$$\Delta TFP_{it}^{R} = \Delta\tilde{a}_{it} + \Delta\tilde{\lambda}_{it} + \Delta\tilde{\bar{q}}_{it}. \tag{4.11}$$

4.2.5 Labour productivity decomposed

TFP, whether of the quantity or revenue flavour, is not the only productivity measure of interest to economists and policymakers. Labour productivity measured as output per worker or per hour worked is widely used and is often more closely related to wages and living standards. In many empirical settings researchers use a simple growth accounting methodology to attribute (log) labour productivity changes to changes in the labour input, other inputs and TFP building on the Cobb-Douglas production function:

$$r_{it} = q_{it} + p_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \underbrace{a_{it} + p_{it}}_{TFP_{it}^{R}},$$

where $a_{it} + p_{it}$ is nothing else than revenue TFP. By subtracting l_{it} from both sides while rearranging and considering time changes Δ, we have the following labour productivity (LP_{it}) decomposition:

$$\Delta LP_{it} = \Delta(r_{it} - l_{it}) = (\alpha_L - 1)\Delta l_{it} + \alpha_M \Delta m_{it} + \alpha_K \Delta k_{it} + \Delta TFP_{it}^{R}. \tag{4.12}$$

The equivalent factor proportions version used in Goodridge et al. (2016), Harris and Moffat (2017) and Pessoa and Van Reenen (2014) is:

$$\Delta LP_{it} = \gamma \Delta l_{it} + \alpha_M \Delta(m_{it} - l_{it}) + \alpha_K \Delta(k_{it} - l_{it}) + \Delta TFP_{it}^{R}, \tag{4.13}$$

where $\gamma = \alpha_L + \alpha_M + \alpha_K - 1$ is a parameter measuring returns to scale.

Within the MULAMA model both decompositions can be further developed. More specifically, substituting (4.11) for ΔTFP_{it}^R in (4.12) and simplifying leads to:

$$\Delta LP_{it} = \Delta[(\frac{\alpha_L}{\mu_{it}} - 1)l_{it}] + \alpha_M \Delta(\frac{m_{it}}{\mu_{it}}) + \alpha_K \Delta(\frac{k_{it}}{\mu_{it}}) + \Delta(\frac{a_{it}}{\mu_{it}}) + \Delta(\frac{\lambda_{it}}{\mu_{it}}), \tag{4.14}$$

while, in the factor proportions version, substituting (4.11) for ΔTFP_{it}^R in (4.13) delivers:

$$\Delta LP_{it} = \Delta[(\frac{\gamma + 1}{\mu_{it}} - 1)l_{it}] + \alpha_M \Delta(\frac{m_{it} - l_{it}}{\mu_{it}}) + \alpha_K \Delta(\frac{k_{it} - l_{it}}{\mu_{it}}) + \Delta(\frac{a_{it}}{\mu_{it}}) + \Delta(\frac{\lambda_{it}}{\mu_{it}}). \tag{4.15}$$

From (4.14) and (4.15) it now appears clearly how changes in labour productivity materialize as a consequence of changes in quantity TFP, demand, markups and inputs use.

4.2.6 The restricted model and services

Quantity and price data are very often not available to researchers, almost universally for the service sectors where output measures can be particularly problematic (Office for National Statistics, 2007). In such cases, the only available option is to estimate the production function and related TFP using revenue, or value added, as a measure of output, i.e., measure revenue TFP. This raises the issue, discussed above, of the bias in the estimation of production function coefficients coming from any correlation between the underlying prices and inputs use. In this respect, FMMM provide an overall reassuring message.

More specifically, FMMM find that more standard revenue TFP measures obtained using revenue as a measure of output are reasonably well correlated with revenue TFP measures obtained using quantity as a measure of output; something we will show later on holds in our data too. In other words, the bias involving production function coefficients coming from the correlation between prices and inputs use is not a first-order issue. The key insights about this finding are as follows. In the data firms with higher TFP-Q are characterized by lower marginal costs and charge, as expected, lower prices while selling higher quantities, so using more inputs, everything else equal. Furthermore, firms with high demand λ are characterized by higher marginal costs and charge, as expected, higher prices while selling higher quantities, so using more inputs, everything else equal. At the same time, the data suggest a trade-off between producing large quantities at low marginal costs and prices (high TFP-Q and low λ firms) and producing large quantities at high marginal costs and prices (low TFP-Q and high λ firms) which in line with findings in the demand system literature (Ackerberg et al., 2007). Therefore, there is no overall strong correlation between prices and inputs use and so the bias involving production function coefficients, when using revenue as a measure of output, is present but it is such that related TFP-R measures are still informative about the 'true' TFP-R.

FMMM further show that the key disadvantage of not having price and quantity data is the fact that one cannot any more disentangle quantity TFP a from demand λ but only retrieve a composite of the two: $\omega_{it} = a_{it} + \lambda_{it}$. However, markups can still be computed from the estimated production function coefficients using (4.4) while the TFP-R and labour productivity decompositions provided above still hold by replacing the distinct a and λ terms with a unique ω term. For example, considering (4.9) we have:

$$TFP^R_{it} = \frac{\omega_{it}}{\mu_{it}} + \frac{1-\mu_{it}}{\mu_{it}}\bar{q}_{it}, \tag{4.16}$$

which provides a formula to retrieve ω_{it} from measures of revenue TFP, markups and scale; measures that only require estimates of the production function coefficients. At the same time, for example, (4.15) becomes:

$$\Delta LP_{it} = \Delta[(\frac{\gamma+1}{\mu_{it}} - 1)l_{it}] + \alpha_M \Delta(\frac{m_{it}-l_{it}}{\mu_{it}}) + \alpha_K \Delta(\frac{k_{it}-l_{it}}{\mu_{it}}) + \Delta(\frac{\omega_{it}}{\mu_{it}}). \tag{4.17}$$

FMMM label this restricted version of the model MUOMEGA in reference to the two heterogeneities it allows for, markups (MU) and a composite of TFP-Q and λ (OMEGA). FMMM also develop an estimation procedure for the restricted model.

In what follows we consider, in order to provide evidence of the robustness of our results, revenue-based production function estimations for UK service sectors obtained using either the estimation procedure provided in FMMM or the more standard WLD approach.

4.2.7 Some additional remarks

Manufacturing. As far as manufacturing is concerned, we consider as baseline the implementation of the MULAMA model and related decompositions based on the DGKP estimation method applied to the Cobb-Douglas production function (4.5) on single-product firms. However, we also present results based on the FMMM estimation method, also in Cobb-Douglas form, as well as findings obtained from the Translog production function (4.6) and the sample of multi-product firms for robustness. Our key findings are little affected by whether we use the DGKP or the FMMM estimation procedure, by whether we use a Cobb-Douglas or a Translog production function and whether we use the single-product firms sample or the multi-product firms sample.

As is customary in productivity analyses, we correct (in all estimations) for the presence of measurement error in output (quantity and revenue) and/or unanticipated (to the firm) shocks using the methodology described in DGKP and FMMM. We also consider a full battery of 8-digit product dummies, as well as year dummies in our production function estimations. Indeed, quantity in the data is measured in units (kilograms, litres, number of items, etc.) that are specific to each 8-digit product and so quantity TFP a_{it} can be reasonably compared across firms and time only within an 8-digit product category. For similar reasons, also λ_{it} can be reasonably compared across firms and time only within an 8-digit product category.

In terms of the number of production function estimations, we are forced by sample constraints to run a single estimation across the whole manufacturing firms sample instead of by two-digit industry groupings. As explained more in detail below, we need firms to be in both the Prodcom and ARDx datasets, which are described below, while also requiring information on one and two period lags for all variables. In this respect, results obtained using the more flexible Translog production function should reassure about the issue of heterogeneity in output elasticities across firms and industries in manufacturing. At the same time, we show later on that patterns of various TFP-R measures, as well as of value added per worker and output per worker, are very similar when comparing our estimation sample to the full set of manufacturing firms available in the ARDx dataset.

Services. As far as service industries are concerned, we consider as baseline the revenue TFP estimations, and related MUOMEGA model decompositions, based on the WLD approach applied to the Cobb-Douglas production function (4.5). We also present very similar results based on the FMMM estimation method for the MUOMEGA model, also in Cobb-Douglas form.

Again, as is customary in productivity analyses, we correct (in all estimations) for the presence of measurement error in revenue and/or unanticipated (to the firm) shocks using the methodology described in DGKP and FMMM. Production function estimations are run separately for each NACE Section (11 in total) and include a full battery of 2-digit industry dummies as well as year dummies.

Composition effects and weighting. There are reasonable concerns about composition effects as the ARDx firm sample changes over time, and particularly so from 2007-8 when the ONS switched from producing the Annual Respondents Database to the Annual Business Survey. Therefore we present, as our baseline, results for what we label the 'within sample' which compares the mean of within-firm changes between $t-1$ and t. The within sample is thus comprised of firms present in the data in both $t-1$ and t (and if manufacturing firms, also producing the same product in both years) and it allows us to minimise the impact of sample composition effects, including those related to different units of measurement for the products of manufacturing firms. We show below that the within sample accounts for the lion's share of overall firm revenue.

Finally, we choose to present our baseline results using revenue weights, given that our research question is more closely aligned to understanding aggregate changes in productivity rather than for the average firm. We also present robustness results based on employment weights as well as on equal weights, i.e., un-weighted.

Operationally, we calculate an index for each variable of interest after averaging within-firm changes between $t-1$ and t:

$$\Delta \bar{y}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} (y_{it} - y_{it-1}) w_{it}, \tag{4.18}$$

where y_{it} is a variable of interest (TFP-Q, TFP-R, λ, μ, etc.), $\Delta \bar{y}_t$ is the weighted average of within firm changes in y_{it}, $w_{it} = \frac{1}{2}(R_{it} + R_{it-1})$ are the weights computed using the average firm revenue between $t-1$ and t,[16] and N_t is the number of firms present in the data in both $t-1$ and t. We use this formula to construct the index of changes, setting the base year to 2008 for manufacturing and 2007 for services.

[16]In terms of weighting we thus choose a compromise between Laspeyres and Paasche weights. In unreported results we also experimented with Laspeyres and Paasche weights obtaining very similar results.

4.3 Data and Descriptives

4.3.1 Data

The core data required to estimate firm-level revenue TFP using standard methodologies comprise revenue, employment costs, intermediate inputs and capital stock. For these variables we turn to the Office for National Statistics (ONS) Annual Respondents Database X[17] (ARDx). The ARDx is a recently-created dataset for researchers using ONS secure access via the Virtual Microdata Library and the Secure Data Service. It combines and standardizes data and variables across the period 1998-2017 from two surveys, the Annual Business Survey (ABS) which has been carried out since 2009, and its predecessor, the Annual Business Inquiry (ABI), which was carried out 1998-2008 and used to create the Annual Respondents Database. These are the largest business surveys in the UK and have been used by many UK productivity researchers including Barnett et al. (2014), Harris and Moffat (2017) and Riley et al. (2013). The ABI and ABS are similar in sampling method, structure and questions, and the ARDx was created to provide researchers with a consistent dataset across time.

The ARDx covers around two-thirds of UK economic activity, comprising most SIC 2007 sections, except parts of sections A (agriculture) and K (finance), and all of O (public administration and defence), T (activities of households) and U (extraterritorial organisations). The sample frame of the ABS is the Inter-Departmental Business Register (IDBR), a register of firms created from HM Revenue and Customs data on VAT and PAYE details. The sample is stratified by SIC 2007 activities (at the 4-digit level), employment size and country (England & Wales, and Scotland). A sample of 62,000 of the 2.1m firms on the IDBR is drawn annually. All firms in the largest employment categories in each cell are selected. Firms in each of the cells including smaller businesses are drawn for two consecutive years only, and then not re-selected for at least two years afterwards. For the smallest (0-9) employment category, firms are only selected in a single year, and then not again for at least three years afterwards in respect of the Osmotherly rules to ensure that the compliance burden on firms is proportionate. Because we require lagged values of variables in our estimations, we drop these firms and focus on firms with at least 10

[17] Office for National Statistics. Virtual Microdata Laboratory, University of the West of England, Bristol (2017)

employees.

Table 4.1: Examples of 8 digit PRODCOM products within 6 digit CPA categories

PRODCOM	Description
	COTTON YARN (OTHER THAN SEWING THREAD)
13.10.61.32	Yarn of uncombed cotton, not per retail sale, for woven fabrics (excluding for carpets and floor coverings)
13.10.61.33	Yarn of uncombed cotton, not per retail sale, for knitted fabrics and hosiery
13.10.61.35	Yarn of uncombed cotton, not per retail sale, for other uses (including carpets and floor coverings)
13.10.61.52	Yarn of combed cotton, not per retail sale, for woven fabrics (excluding for carpets and floor coverings)
13.10.61.53	Yarn of combed cotton, not per retail sale, for knitted fabrics and hosiery
13.10.61.55	Yarn of combed cotton, not per retail sale, for other uses (including carpets and floor coverings)
	INPUT OR OUTPUT UNITS, WHETHER OR NOT CONTAINING STORAGE UNITS IN THE SAME HOUSING
26.20.16.40	Printers, copying machines and facsimile machines, capable of connecting to an automatic data processing machine or to a network (excluding printing machinery used for printing by means of plates, cylinders and other components, and machines performing two or more of the functions of printing, copying or facsimile transmission)
26.20.16.50	Keyboards
26.20.16.40	Other input or output units, whether or not containing storage units in the same housing

Source: EC RAMON Database (2009 Prodcom List)

Estimation of quantity TFP and demand for the manufacturing sector requires data on quantities sold and prices, information that is available in the Products of the European Community (Prodcom) dataset.[18] Prodcom is a standardized survey of production across the European Union, collected by national statistical agencies using a 3,500 product list in an 8-digit nomenclature established by Eurostat. The first four digits correspond to the Nomenclature Statistique des Activités Economiques dans la Communauté Européenne (NACE) using revision 1.1 up to 2007 and revision 2 from 2008, and the first six digits to the Classification of Products by Activity (CPA) with the last 2 digits adding further detail. It covers SIC 2007 sections B (mining) and C (manufacturing) sectors. We exclude section B to focus on manufacturing. The survey captures at least 90% of production in all the four digit industries covered by the survey.

[18] Office for National Statistics (2018c)

Illustrating the advantages of highly disaggregated data, Table 4.1 shows an extract from the 2009 Prodcom list for the six-digit codes 13.10.61: 'Cotton yarn (other than sewing thread)' and 26.20.16: 'Input or output units, whether or not containing storage units in the same housing'. The latter example highlights how it can often be necessary to work with 8-digit data rather than the already quite detailed 6-digit level, in order to be confident to compare reasonably similar items. The former example, highlights instead how a 8-digit product breakdown can be very precise in terms of narrowing down products and so working at this level of disaggregation allows us to take into account rich differences in technology, demand and degree of competition across finely defined products.

Around 20,000 firms a year, representing at least 90% of the value of production in each 4-digit industry, are surveyed to construct the Prodcom dataset using the IDBR as the sample frame. The sample is stratified by employment size and SIC 2007 4-digit industry. There are 3 employment band thresholds above which all firms are surveyed (20, 50 and 100), where the cut-off varies between industries. Below that firms are rotated through the sample.

The quantity and value of sales are recorded for each 8-digit product produced by a firm annually.[19] We measure firm-product-year specific prices as the ratio of the value of sales to the quantity and apply a small trimming on the distribution of prices by 8-digit product to get rid of outliers. Prodcom product codes change occasionally over time and we employ the methodology described in Van Beveren et al. (2012) to obtain a time-consistent product classification. Appendix C-1 provides more details on the product concordance procedure. We also ensure that the units of measure used to record quantities are consistent over time. Metadata provided by the ONS for Prodcom links each product-year with a unit of measure and where these units change over time within a product we define a new product, leaving us a total of 5,028 product-units. Some products are reported within Prodcom without quantity data and we drop these products, leaving 3,239 consistent product-units with non-missing quantity data. Our unit of observation is strictly firm-product-unit-year but for ease of exposition we refer to firm-product-year throughout the analysis.

[19]This introduces a discrepancy with the ARDx. In Prodcom, firms report calendar-year product sales and quantities, while in the ARDx firms can report either calendar year or financial year revenue figures. We deal with this by dropping firms that report values for ARDx sales that are outside a range of +/- 30% of total Prodcom sales. This also has the effect of removing manufacturing firms with a high proportion of services in revenues.

Both datasets cover Great Britain while data for Northern Ireland are held separately and are excluded from our analysis. Our analysis focuses on the period 2003-2013 in order to both gain insights into the pre- and post-crisis productivity performance and provide evidence comparable to previous studies. We deflate, as standard, both output and inputs values from the ARDx using information provided by the ONS. Appendix C-1 provides more details on the datasets, the construction of capital stocks and the deflators used in the analysis while Table 4.2 describes the main variables used in our estimations.

Table 4.2: Description of main variables

Variable	Description
Output £000s	Approximate output at basic prices, calculated (following Ayoubkhani (2014)) as 'total turnover excluding VAT' less 'goods bought for resale' plus 'changes in stocks and work-in-progress' less 'changes in stocks of materials, storage and fuels' plus 'work of a capital nature' less 'total net taxes' plus 'business rates' plus 'net taxes on production'.
Intermediates £000s	Approximate intermediate consumption at purchaser's prices, calculated as 'total purchases of goods and services' less 'value of insurance claims received' less 'goods and services bought for resale' less 'changes in stocks of materials, storage and fuels'.
Value added £000s	Approximate GVA at basic prices, calculated as output less intermediates.
Capital stock £000s	Plant and machinery capital stock, see Appendix C-1 for details of calculation
Employment	From Inter-departmental Business Register at time of sample selection, reported in ARDx
Wage bill £000s	Total employment costs from ARDx
Quantity	Volume of production sold from Prodcom, measured in product-specific units
Sales	Value of production sold from Prodcom

4.3.2 Descriptives

We merge the ARDx, capital stock and Prodcom data using a unique identifier for what the ONS refers to as a 'reporting unit'[20] and we refer to as a firm, our unit of analysis.[21]

Manufacturing

Although the ARDx is a representative sample of private-sector firms and Prodcom is designed to cover 90% of manufacturing output, there is not perfect overlap between the two datasets, a problem compounded by the requirements of the DGKP and FMMM estima-

[20] Large businesses ("enterprises" to the ONS, and the legal entity of the business) may be split into a number of reporting units, while reporting units can comprise a number of local units which are separated geographically. Data in the ARDx are collected at the reporting unit level

[21] Some authors, e.g. Harris and Moffat (2017), argue that the local unit (plant) is the preferred unit of analysis because it provides cleaner estimates of capital stock due to plant entry and exit within a firm, but this requires apportioning firm-level inputs and outputs to plants, while information on production from Prodcom is only available at the firm level.

tion procedures. More specifically, to estimate the production function for manufacturing firms while using quantity as a measure of output we require/impose:

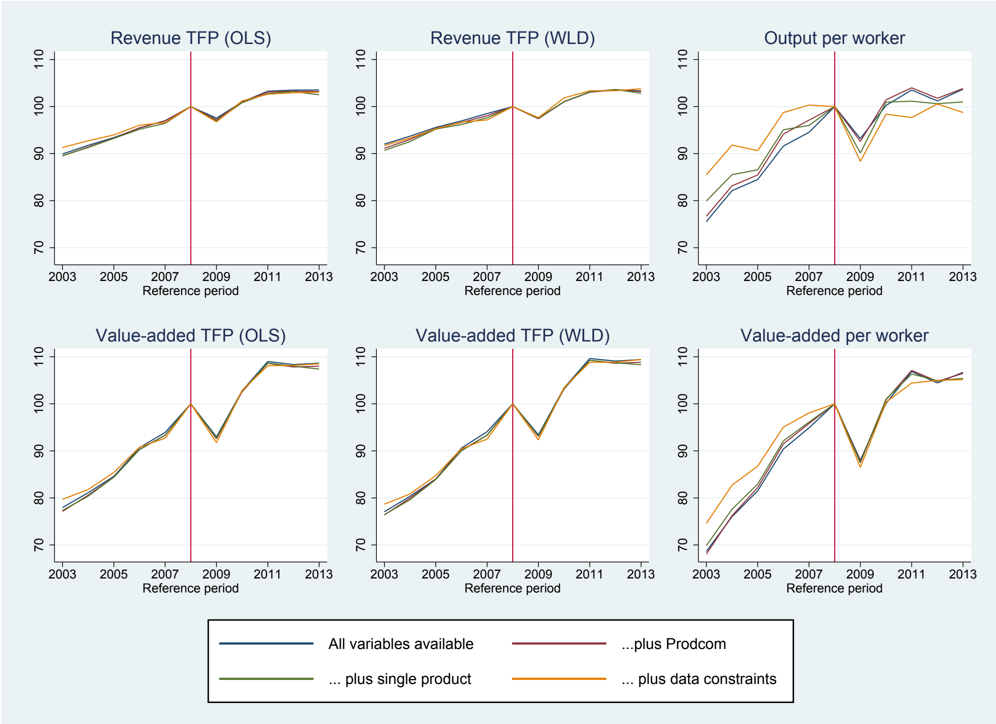
- non-missing, positive values for employment, total turnover ex. VAT, purchases of goods and materials, capital stock, total wages and salaries from the ARDx
- the firm is in the Prodcom survey
- that Prodcom records a non-zero value for quantity of this product
- the firm produces only 1 product in any given year
- and that firm revenues reported in Prodcom are within 30% of the output calculated using ARDx data

These demands sharply reduce the available sample size and in Appendix C-1 we provide more details on the merging process and the various constraints.[22] We label 'final sample' the combined ARDx and Prodcom sample satisfying all our requirements.

We do not think of the final sample as being representative of the broader populations of manufacturing firms and/or single-product firms as selection into the estimation sample is unlikely to be random, particularly reflecting industry characteristics (for the availability of physical quantity data in Prodcom) and firm size (for selection separately into the Prodcom and ARDx surveys with differing stratification bands). Rather, we seek to show that time trends in productivity measures within the group of firms for which we are able to estimate the full MULAMA model (final sample) are similar to those in broader samples. Our first step to do this is to estimate standard two-factor (value-added based) and three-factor (revenue-based) production functions, using both the method of ordinary least squares and the WLD approach, on the whole sample of manufacturing firms in the ARDx for which revenue productivity can be estimated ('all variables available'). We report production function coefficients in Table C.4 in Appendix C-2, and use these to calculate mean revenue-based and value-added-based TFP over time for different samples going from the largest ('all variables available') to the final sample we use in our estimations ('...plus data constraints'). We graph the results, along with output per worker and value-added per worker, in Figure 4.1. All the productivity measures we consider indeed display

[22]In our estimations we also apply some small trimming (top/bottom percentile) of unit prices by product, the capital-to-labour ratio and labour-output ratio by 2-digit industry, and drop firm-year observations where the share of materials in output is less than 0.1 or greater than 0.95.

Figure 4.1: Revenue TFP and labour productivity measures by sample, Manufacturing, 2003-2013



All revenue TFP and labour productivity measures are indexed: 2008=100. 'All variables available' refer to manufacturing firms in the ARDx that have a) at least 10 employees and b) have the following variables available: employment, total turnover ex. VAT, purchases of goods and materials, capital stock, total wages and salaries. '...plus Prodcom' adds the requirement that the firm-year observation is also in the Prodcom dataset. '...plus single product' adds the requirement that at least 90% of a firm's output at basic prices is accounted for by sales of a single product . '...plus data constraints' adds the requirement that Prodcom measures a non-zero quantity of production, and that firm revenues reported by Prodcom are within 30% of the output calculated from the ARDx. Revenue TFP estimated using ordinary least squares (OLS) or the Wooldridge procedure (WLD) in revenue and value-added forms.

a very similar behavior across time for the four samples. This builds confidence that the trends in quantity TFP, demand and markups for the wider sample of firms in the ARDx are similar to those we uncover in the final sample.

One final issue we address is related to the ARDx sampling frame changing over time – notably in 2008/9 when SIC 2008 replaced SIC 2004 – leading to concern that comparing firm averages over time, whether weighted on un-weighted, will be biased by changes in the sample composition. Because of this, in our results we focus, as already discussed above, on within-firm changes over time using an unbalanced panel of firms for which we have observations in both year t and year $t-1$: the within sample.

Table 4.3 shows the number of observations by year corresponding to the final sample and the within sample for manufacturing along with the share of revenue (combined rev-

Table 4.3: Comparison of the final sample and the within sample: manufacturing firms.

Year	Final sample N	Within sample N	Within sample revenue share
2003	1,193	641	–
2004	1,134	641	0.6838
2005	1,274	564	0.5686
2006	1,148	617	0.6742
2007	1,213	594	0.6556
2008	1,048	401	0.4981
2009	1,047	443	0.6315
2010	1,056	496	0.6991
2011	1,014	450	0.6870
2012	1,016	478	0.6711
2013	968	477	0.7401

The last column reports the share of revenue (combined revenue in $t-1$ and t) accounted for by firms in the within sample.

enue in $t-1$ and t) accounted for by firms in the within sample. While within sample observations account for fewer than half the available firm-year observations, they account for roughly two-thirds of full-sample revenues. Summary statistics for key variables and MULAMA model estimates (obtained with the DGKP estimation procedure) referring to both samples are provided in Table 4.4.

Table 4.4: Summary statistics for the final sample and the within sample: manufacturing firms.

	Final sample		Within sample	
	mean	sd.	mean	sd.
log output	9.351	1.438	9.815	1.352
log wage bill	7.928	1.295	8.346	1.214
log intermediates	8.798	1.596	9.322	1.468
log capital stock	7.605	1.772	8.160	1.622
log price	-2.380	3.368	-2.309	3.381
TFP-R (DGKP)	0.859	0.174	0.832	0.136
TFP-R (OLS)	1.050	0.134	1.048	0.129
TFP-R (WLD)	0.722	0.131	0.697	0.106
log value-add/worker	3.668	0.520	3.683	0.493
log revenue/worker	4.717	0.648	4.794	0.629
a	3.238	3.377	3.141	3.387
λ	-0.932	5.859	-1.674	4.692
ω	2.307	4.875	1.467	3.276
μ	1.183	0.561	1.081	0.348
scale	8.492	1.463	8.983	1.353
adjusted a	3.060	3.461	3.120	3.564
adjusted λ	-1.935	4.488	-2.332	4.433
adjusted ω	1.124	2.165	0.788	2.013
adjusted scale	-0.265	2.080	0.045	1.969
N	12,111		5,802	

Table 4.5: Comparison of the final sample and the within sample: services firms.

Year	Final sample N	Within sample N	Within sample revenue share
2003	18,941	8,387	–
2004	18,063	8,387	0.71
2005	17,174	7,813	0.73
2006	14,024	6,438	0.73
2007	15,206	6,266	0.70
2008	12,945	5,674	0.68
2009	12,795	6,493	0.80
2010	12,086	6,079	0.82
2011	12,829	5,966	0.83
2012	12,983	6,404	0.82
2013	12,802	6,626	0.82

The last column reports the share of revenue (combined revenue in $t-1$ and t) accounted for by firms in the within sample.

Table 4.6: Summary statistics for the final sample and the within sample: services firms.

	Final sample		Within sample	
	mean	sd.	mean	sd.
log ouput	8.396	1.775	9.028	1.810
log wage bill	7.432	1.743	8.072	1.776
log intermediates	7.360	1.960	8.026	1.979
log capital stock	5.461	2.426	6.280	2.369
TFP-R (OLS)	1.165	0.164	1.166	0.169
TFP-R (WLD)	0.629	0.588	0.597	0.633
μ	1.720	0.878	1.631	0.748
ω	6.361	7.158	6.014	6.591
scale	7.767	2.017	8.431	2.063
adjusted ω	2.737	2.420	2.766	2.508
adjusted scale	-2.108	2.417	-2.168	2.508
N	159,848		74,533	

Services

For services firms we merge the ARDx and capital stock data, using the unique identifier reporting unit, and impose non-missing, positive values for employment, total turnover ex. VAT, purchases of goods and materials, capital stock and total wages and salaries. This delivers the final sample for services firms.[23] Again, because of concerns about sampling frame changes, we focus on within-firm changes over time using an unbalanced panel of firms for which we have observations in both year t and year $t-1$: the within sample for services. We also provide (very similar) results using both un-weighted and weighted results.

Table 4.5 shows the number of observations by year corresponding to the final sample

[23]We also apply, in line with the manufacturing analysis, a small trimming based on the top/bottom percentiles of the capital-to-labour ratio and labour-output ratio by 2-digit industry while dropping firm-year observations where the share of materials in output is less than 0.1 or greater than 0.95.

and the within sample for services along with the share of revenue (combined revenue in $t-1$ and t) accounted for by firms in the within sample. Again, while within sample observations account for fewer than half the available firm-year observations, they account for more than two-thirds of full-sample revenues. Summary statistics for key variables and MUOMEGA model estimates (obtained with the WLD estimation procedure) referring to both samples are provided in Table 4.6.

4.4 Estimates and insights

4.4.1 Manufacturing

Estimates of the Cobb-Douglas production function coefficients for manufacturing firms belonging to the final sample using the DGKP procedure are shown in Table 4.7 while descriptive statistics on the various MULAMA model components are displayed in Table 4.4 in Section 4.3. The coefficients are in line with those reported in DGKP for India, FMMM for Belgium and in Chapter 3 for France using quantity as a measure of output and follow the expected pattern; materials coefficient larger than that of labour, which is larger than that of capital and evidence of roughly constant returns to scale. The coefficients – representing estimates of quantity elasticities rather than revenue elasticities – are reasonably close to those obtained in the revenue production function estimations reported in Table C.4 in Appendix C-2, particularly the WLD method.

Table 4.7: Manufacturing. Production function coefficients from the DGKP estimation procedure using quantity as a measure of output

	log quantity
log wage bill	0.342
	(0.015)***
log intermediates	0.636
	(0.036)***
log capital stock	0.025
	(0.012)**
Obs	3,002

Cobb-Douglas production function coefficients estimated using the DGKP method on the final sample of manufacturing firms. The regression includes 8-digit product dummies and year dummies. The number of observations refers to firms with at least 2-period lags for all variables required in the estimation procedure. See FMMM for more details. Robust standard errors clustered by firm. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Production function estimation allows us to calculate quantity TFP a_{it} along with markups μ_{it} and demand λ_{it}, as well as revenue total factor productivity TFP^R_{it}, from the corresponding equations in Section 4.2. With reference to revenue TFP, and in line with evidence provided in FMMM discussed above, Table 4.8 shows that this measure – TFP-R (DGKP) – is reasonably well correlated with alternative measures of revenue productivity. More specifically, the correlation is stronger the more sophisticated the alternative measure: low with output per worker, then increasing through value-added per worker, OLS revenue TFP and most strongly with WLD revenue TFP. In a similar spirit we show, building on the linear revenue TFP decomposition provided by (4.10), that markup-adjusted TFP-Q $\tilde{a}_{it}$, markup-adjusted demand $\tilde{\lambda}_{it}$ and markup-adjusted scale $\tilde{\tilde{q}}_{it}$ explain a large share of the variation in OLS and WLD revenue TFP in both levels and changes. This is reported in Table 4.9 along with coefficients (equal to one) and R^2 (equal to one) related to the DGKP revenue TFP. Coefficients for OLS and WLD revenue TFP indicate that an increase in markup-adjusted TFP-Q and/or demand and/or scale are associated with an increase in firm revenue TFP. As discussed above, the evidence provided in Tables 4.8 and 4.9 suggests that the bias involving production function coefficients in revenue TFP estimations coming from the correlation between prices and inputs use is not a first-order issue. In this respect, the real advantage of having price and quantity data is the capacity to disentangle quantity TFP from demand.

Table 4.8: Pairwise correlation coefficients between revenue productivity measures.

	TFP-R (OLS)	TFP-R (WLD)	TFP-R (DGKP)	log value-add per worker	log revenue per worker
TFP-R (OLS)	1				
TFP-R (WLD)	0.791	1			
TFP-R (DGKP)	0.638	0.783	1		
log value-add per worker	0.425	0.262	0.506	1	
log revenue per worker	0.542	0.0468	0.167	0.724	1
Obs	12,111				

Correlations are across observations in the final sample for manufacturing firms.

In order to provide useful insights for our analysis, while confirming previous findings in FMMM and in Chapter 3, Table 4.10 reports results of some simple regressions looking at the relationship of prices and markups with the underlying driving variables of the MULAMA model; namely quantity TFP, demand and the capital stock.[24] More specifically,

[24]In the MULAMA model the capital stock is predetermined for the firm in t and so, along with quantity TFP and demand, determines inputs choices and pricing.

Table 4.9: Regressions of TFP-R measures on $\tilde{a}$, $\tilde{\lambda}$ and $\tilde{\tilde{q}}$

	Levels			Changes		
	OLS	WLD	DGKP	OLS	WLD	DGKP
	(1)	(2)	(3)	(4)	(5)	(6)
adjusted a	0.713	0.698	1.000	0.699	0.660	1.000
	(0.013)***	(0.013)***	(0.000)***	(0.018)***	(0.019)***	(0.000)***
adjusted λ	0.710	0.696	1.000	0.697	0.659	1.000
	(0.013)***	(0.013)***	(0.000)***	(0.017)***	(0.018)***	(0.000)***
adjusted scale	0.744	0.698	1.000	0.732	0.667	1.000
	(0.013)***	(0.014)***	(0.000)***	(0.018)***	(0.019)***	(0.000)***
R^2	0.88	0.86	1.00	0.79	0.80	1.00
N	12,111	12,111	12,111	5,161	5,161	5,161

Regressions in columns 1 to 3 use the manufacturing firms final sample and the levels of all variables, while columns 4 to 6 use the manufacturing firms within sample and firm-level changes between $t-1$ and t. All regressions include a set of year dummies and 8-digit product dummies. Robust standard errors clustered by firm. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

we regress log price p_{it} and log markup $\log(\mu_{it})$ on quantity TFP a_{it}, demand λ_{it} and log capital stock k_{it}. We estimate those regressions both in levels and in first-differences (finding very similar coefficients), and include a full set of year dummies and 8-digit product dummies while clustering standard errors at the firm-level.

Estimations referring to log prices (log markups) in columns 1 and 2 (3 and 4) of Table 4.10 indicate, in line with expectations and previous results in FMMM and 3, that more productive firms and/or an increase in productivity for a firm (higher a_{it}) are associated with both lower prices and higher markups, i.e., an incomplete pass-through. At the same time firms facing a stronger demand and/or an increase in demand for a firm (higher λ_{it}) are associated with both higher prices and higher markups, so confirming the relevance of λ_{it} as a measure of firm-specific demand. Finally, larger firms and/or an increase in size for a firm (higher k_{it}) is associated with both lower prices and lower markups, which is in line with the elasticity of demand decreasing along the demand curve.

Table 4.10: Regressions of log prices and log markups on a, λ and log capital

	log price		log markup	
	Levels	Changes	Levels	Changes
	(1)	(2)	(3)	(4)
a	-0.968	-0.944	0.118	0.109
	(0.001)***	(0.005)***	(0.002)***	(0.003)***
λ	0.026	0.027	0.115	0.107
	(0.000)***	(0.001)***	(0.001)***	(0.001)***
log capital	-0.010	-0.013	-0.013	-0.005
	(0.001)***	(0.006)**	(0.001)***	(0.002)**
R^2	1.00	0.96	0.98	0.98
N	12,111	5,161	12,111	5,161

Regressions in columns 1 and 3 use the manufacturing firms final sample and the levels of all variables, while columns 2 and 4 use the manufacturing firms within sample and firm-level changes between $t-1$ and t. All regressions include a set of year dummies and 8-digit product dummies. Robust standard errors clustered by firm. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

4.4.2 Services

We estimate revenue-based production functions using the WLD approach by SIC section and use the estimated coefficients to measure revenue TFP (TFP-R (WLD)) and further employ the corresponding equations of the MUOMEGA model from Section 4.2 to quantify markups, the composite of quantity TFP and demand (ω), and production scale.

Production function estimates are reported in Table 4.11 while descriptive statistics on the various MUOMEGA model components are displayed in Table 4.6 in Section 4.3. Coefficients are, as expected, on average higher for the labour input in services as compared to manufacturing (a weighted average of 0.446 in services compared to the 0.342 for manufacturing). The labour input coefficients is particularly high in sections one might expect, for instance, 'Q: Human health and social work', and to a lesser extent in 'I: Accommodation and food service', and low in 'F: Construction' and 'H: Transport and storage'. The opposite is true for the intermediates coefficient (a weighted average of 0.603 in services compared to 0.636 in manufacturing). Sections like construction and transport are both at or above 0.8 but most others are well below 0.6. Capital stock coefficients are small, as typically found in 3-factors production function estimations, and somewhat noisy, which is likely related to the well-known issue of measurement error particularly plaguing the capital stock (Griliches and Mairesse, 1995). Finally, there is evidence of roughly constant or slightly increasing returns to scale.

Table 4.11: Service industries: Production function coefficients from the WLD estimation procedure using revenue as a measure of output.

SIC Section	log wage bill	log intermediates	log capital	Ret. to scale	N
F	0.329 (0.019)***	0.862 (0.138)***	-0.013 (0.017)	1.178	16,340
G	0.499 (0.006)***	0.485 (0.016)***	0.025 (0.008)***	1.009	50,758
H	0.369 (0.011)***	0.748 (0.077)***	0.028 (0.016)*	1.145	10,089
I	0.529 (0.019)***	0.503 (0.054)***	0.037 (0.012)***	1.069	11,948
J	0.463 (0.013)***	0.596 (0.048)***	0.033 (0.020)*	1.093	9,043
L	0.430 (0.012)***	0.533 (0.045)***	0.032 (0.010)***	0.995	4,064
M	0.506 (0.015)***	0.546 (0.057)***	0.015 (0.012)	1.068	18,336
N	0.474 (0.011)***	0.553 (0.085)***	0.002 (0.015)	1.029	13,169
P	0.387 (0.030)***	0.508 (0.060)***	0.006 (0.018)	0.901	7,326
Q	0.690 (0.009)***	0.292 (0.024)***	0.046 (0.010)***	1.028	8,539
RSTU	0.369 (0.016)***	0.660 (0.055)***	-0.001 (0.018)	1.028	10,236

Estimations of Cobb-Douglas revenue production functions by SIC 2007 sector using the WLD method applied to the final sample of services firms. Regressions include 2-digit industry and year dummies. The number of observations refers to firms with at least 2-period lags for all variables required in the estimation procedure. See FMMM for more details. Robust standard errors clustered by firm. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

4.5 Results

4.5.1 Manufacturing

We next turn to our main results and start by focusing on two features that our quantity and price data allow us to analyze for manufacturing firms. For manufacturing, we find more relevant to use the year 2008 to define the pre- and post-crisis periods.

The first feature is about prices, and in particular about the difference between actual price changes and price index changes and how this impacts the measurement of productivity. Producer price indices, and in particular the output price index we have used to deflate sales in our data, should measure average changes in prices for the same product over time. Actual price changes, that are typically not observable to researchers but are observable in our data, instead reflect any changes in prices made by an individual firm and so include, for example, quality improvements to products or changes to the specification of a product. The procedure followed by statistical offices to translate actual price changes into average price changes for the same product, and so deliver official prices indices, is complex and beyond the scope of this paper.[25] However, all we want to highlight here is that actual price changes and price index changes are two different things and, in principle, the difference between the former and the latter, that we label 'real prices changes', will partly reflect product quality improvements. At the same time, real prices changes are included in standard micro and macro productivity analyses because nominal values are deflated using prices indices and so what is left does contain real prices changes. This is the reason why we label as revenue TFP those productivity measures obtained deflating nominal values with a price index. By contrast, quantity TFP levels and changes are in our framework derived from actual quantity data expressed in physical units that are specific to each product (number of items, kilograms, litres, etc.) and, by construction, the revenue TFP of a firm is the sum of the quantity TFP and the log real price.

Table 4.12, and the companion Figure C.1 in Appendix B, provide information on changes in actual prices, the official price index and real prices for the manufacturing firm within sample. More specifically, we start from the firm-product-specific actual prices changes, and the industry/product-specific output price index changes, between $t-1$ to t and aggregate both changes across firms using revenue weights. At the same time, we also construct real price changes starting from the differences between actual prices changes and output price index changes and aggregating them up using again revenue weights. Table 4.12 indicates that, on average, the official price index rose by 2.8% per year for manufacturing products over the period 2003-2008 and by 2.3% per year over the period 2008-2013. Meanwhile, actual prices increased more than the official price index and particularly so in the post-2008 period. Indeed, real price changes averaged 0.7% per year over 2003-2008 and a considerably higher 1.5% per year over 2008-2013. Therefore, irrespective of whether real price changes only reflect product quality improvements or also other elements, and in this respect the 2.2% real prices drop in 2008 does suggest that real prices measure more than just quality improvements, these changes actually helped mitigate the measured productivity slowdown. Indeed, everything else equal, revenue TFP growth over 2008-2013 would have been 0.8=1.5-0.7 per cent lower had real prices increased at the same rate as in the period 2003-2008. Tables C.5 and C.6 in Appendix B further show that this result is robust to considering simple un-weighted changes as well

[25]See, for example, Office for National Statistics (2014).

as to using the multi-products firms sample.

Table 4.12: Real, actual and official price index changes

	Δ real prices	Δactual prices	Δ official price index	Obs
2004	0.014	0.022	0.008	641
2005	0.010	0.026	0.015	564
2006	0.010	0.034	0.024	617
2007	0.018	0.044	0.027	594
2008	-0.022	0.046	0.068	401
2009	0.042	0.069	0.028	443
2010	0.017	0.036	0.019	496
2011	0.001	0.041	0.040	450
2012	-0.001	0.011	0.011	478
2013	0.016	0.031	0.016	477
2003-2008	0.007	0.034	0.028	2,817
2008-2013	0.015	0.037	0.023	2,344

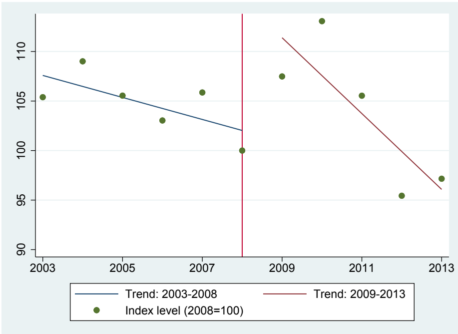
The Table shows mean revenue-weighted changes from $t-1$ to t, for the manufacturing firms within sample, of real and actual prices as well as of the official price index. The final two rows show the mean of changes over the two periods using all the annual observations.

The second feature we highlight here is that, the availability of both price and quantity data, allows us to measure demand and in particular demand changes over time. The demand measures $\tilde{\lambda}_{it}$ and λ_{it} obtained from the MULAMA model, that we use below to analyze the evolution of demand over the time frame 2003-2013, are quite general in that they allow for Variable Elasticity of Substitution (VES) preferences and for markups to be different both across firms and across time. In terms of findings, both of our demand measures $\tilde{\lambda}_{it}$ and λ_{it} indicate a significant slowdown of demand post-2008. In this respect, Figure 4.2 provides complementary evidence of such a slowdown building on a more restrictive, albeit more widely used, framework to measure demand and demand changes over time. More specifically, building on CES preferences and monopolistic competition, and so constant markups across firms and time, we computed, for a given value of the elasticity of substitution σ, a firm-product-time-specific demand measure d_{it}^{CES} matching actual (log) quantity and (log) real price data: $q_{it} = d_{it}^{CES} - \sigma p_{it}$.[26] We then computed within-firm (and product) time changes in d_{it}^{CES} and aggregate them up using revenue weights to construct a demand measure index based on the year 2008 (2008=100). Figure 4.2 also provides, for each panel, two regression lines obtained by fitting index yearly

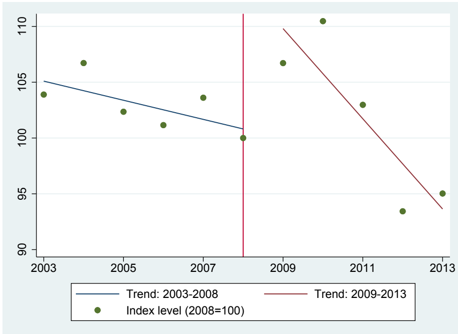
[26]This measure of demand $d_{it}^{CES} = q_{it} + \sigma p_{it}$ is, for example, used by Khandelwal et al. (2013) and Stiebale and Vencappa (2018).

changes between 2003 and 2008 (left regression line) and changes between 2009 and 2013 (right regression line). Figure 4.2 shows two CES demand indices, one constructed using $\sigma = 5$ (left panel) and one constructed using $\sigma = 10$ (right panel),[27] both pointing towards a significant slowdown of demand particularly from 2011 onwards. Figure C.2 Appendix C further show that this result is robust to considering simple un-weighted changes.

Figure 4.2: Manufacturing. CES demand measure for different values of σ



(a) CES demand measure $\sigma = 5$



(b) CES demand measure $\sigma = 10$

Indices of changes in CES demand measures (2008=100). Indices are computed for the manufacturing firms within sample using revenue-weighted changes from $t-1$ to t and two alternative values for the elasticity of substitution: $\sigma = 5$ and $\sigma = 10$.

We now move to the results obtained from the MULAMA model and report in Table 4.13 mean revenue-weighted changes from $t-1$ to t, referring to the manufacturing firms within sample, in DGKP revenue TFP and its components building on the non-linear decomposition of equation (4.9). We also report (column 1) mean revenue-weighted changes in WLD revenue TFP as well as the mean revenue-weighted changes in real prices (column 4). First, we see a similar trend for the DGKP revenue TFP, reported in column 2, to that illustrated across larger samples in Figure 4.1. The DGKP measure rises at a rate of 1.6 percentage points a year from 2003 to 2008, then falls by 2.2 points in 2008/9, averaging across the 2008 to 2013 post-crisis period at a rate of 0.6pp per year, and leaving it 5 points below the pre-crisis trend by 2013. The WLD measure displays a similar pattern: growth of 1.5pp a year through 2008, a drop of 2.3pp in 2008/9 and growing over the 2008 to 2013 post-crisis period at a rate of 0.8 points per year. These results are in line with the evidence that the dismal post-crisis productivity performance of UK manufacturing contains a strong productivity, and in particular revenue TFP, component. However, this paper's contribution is to disentangle the underlying causes of this revenue TFP drop and in particular assess whether and how changes in quantity TFP, demand and markups have

[27] A value of σ between 5 and 10 is what the literature would typically suggest (Anderson and Van Wincoop, 2004).

generated the fall.

In this respect, columns 3 and 4 of Table 4.13 provide evidence that quantity TFP a actually slowed more than revenue TFP in the post-crisis period. The average pre-crisis TFP-Q growth rate of 0.9% turned into a -0.9% growth rate post-crisis leading to a 9 point gap with respect to the pre-crisis trend by 2013. At the same time, and as already discussed above, real prices increased substantially more post-2008 switching from a 0.7% growth rate pre-crisis to a 1.5% growth rate post-crisis. Revenue TFP changes (column 2), are the sum of quantity TFP changes (column 3) and real price changes (column 4), and so the stronger real prices increase post-2008 helped to contain the fall in revenue TFP to a 5 point gap with respect to the pre-crisis trend.

Table 4.13: Manufacturing. Changes of revenue TFP and its components over the period 2003-2013.

	Δ TFP-R (WLD)	Δ TFP-R (DGKP)	Δa	Δp	$\Delta\lambda$	$\Delta\omega$	$\Delta\mu$	Δ scale	Obs
2004	0.007	0.021	0.008	0.013	0.232	0.240	0.020	0.040	641
2005	0.011	-0.004	-0.014	0.010	0.029	0.014	0.002	0.021	564
2006	0.020	0.016	0.006	0.010	0.094	0.100	0.007	0.021	617
2007	0.012	0.020	0.003	0.017	0.087	0.090	0.006	0.029	594
2008	0.028	0.022	0.044	-0.022	0.280	0.324	0.028	-0.037	401
2009	-0.023	-0.022	-0.064	0.042	-0.221	-0.284	-0.023	-0.071	443
2010	0.033	0.028	0.012	0.016	-0.009	0.003	-0.000	0.042	496
2011	0.026	0.024	0.023	0.001	0.305	0.328	0.027	0.019	450
2012	-0.008	-0.015	-0.014	-0.001	-0.043	-0.057	-0.003	0.017	478
2013	0.011	0.015	-0.001	0.016	0.187	0.186	0.015	0.024	477
2003-2008	0.015	0.016	0.009	0.007	0.143	0.152	0.013	0.016	2,817
2008-2013	0.008	0.006	-0.009	0.015	0.045	0.036	0.003	0.007	2,344

The Table shows mean revenue-weighted changes from $t-1$ to t, for the manufacturing firms within sample, of WLD and DGKP revenue TFP, of real prices p as well as of the various components of the revenue TFP non-linear decomposition following from equation (4.9) applied to the DGKP revenue TFP. The final two rows show the mean of changes over the two periods using all the annual observations shown above.

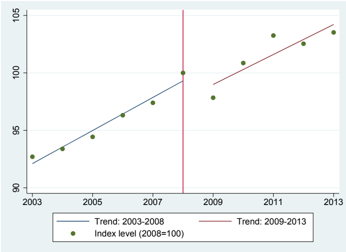
Without information on demand and markups, one would be left wondering what caused the increase in real prices and how quantity TFP and real price changes translate into firms' profit margins and scale of operations. In this respect, column 5 indicates that demand (as measured by changes in λ) also plunged in 2008-9 and overall slowed down with respect to the pre-crisis growth trend. Therefore, the increasing real prices post-2008 are likely related to firms passing to consumers the increasing production costs driven by the declining quantity TFP. The pass-through is incomplete as shown by both the difference between the TFP-Q drop and the real prices increase with respect to the trend

(1.8% drop in TFP-Q and 0.8% increase in real prices) and by the decline in markups in column 7. Indeed, markups sharply declined in 2008-9 and growth slowed thereafter. One important element to stress at this point is that the decline in demand, as measured by λ, is on top of the negative effect on sales produced by increasing real prices post-2008. Indeed, changes in λ measure changes in demand for the same price (and markup), i.e., changes in the underlying demand curve.

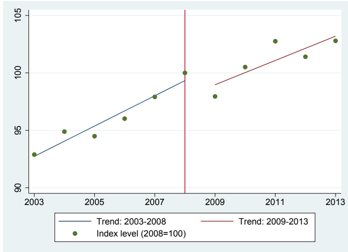
Therefore the depth of the crisis, and in particular its overall impact on sales, production and inputs use, has been particularly severe due to both a supply (TFP-Q a) and a demand (λ) downturn. This is, for example, reflected in the use of inputs by firms in column 8 (scale). More specifically, the growth rate of the average inputs bundle turned from a 1.6% pre-crisis growth rate to a 0.7% post-crisis growth rate, leading to a 4.5% negative gap with respect to the pre-crisis trend. This is reflected also in the combined TFP-Q and demand MULAMA component $\omega = a + \lambda$ (column 6) summarizing the negative supply and demand shocks. In terms of broader implications, the fact that in 2013 scale was up by 3.1% and quantity TFP was down by 4.4% with respect to their 2008 levels implies, given that (log) quantity is equal to TFP-Q plus scale at the firm-level as well as in our aggregation, that quantities sold in 2013 were still 1.3% below their levels back in 2008.

Figure 4.3 presents the results graphically. We construct an index for each variable, setting 2008 as the reference year with a value of 100 so that the graph shows the percentage deviation in the index. Figure 4.3 also provides two regression lines obtained by fitting index yearly changes between 2003 and 2008 (left regression line) and changes between 2009 and 2013 (right regression line). Panels (a) and (b) show quite neatly the break in WLD and DGKP revenue productivity growth before and after 2008 while panels (c) and (d) highlight the more severe downturn in quantity TFP and the mitigating effect of real prices. At the same time, panels (e) and (f) show the downturn in demand and the overall combined change in the pattern of ω. Furthermore, panels (f) and (g) display the post-2008 decline in the evolution of markups and production scale. Finally, Table C.7 in Appendix C-2 shows formal Chow test results regarding the presence of a structural break for some key variables. As can be appreciated from Table C.7, there is indeed strong support for the presence of a structural break in 2008.

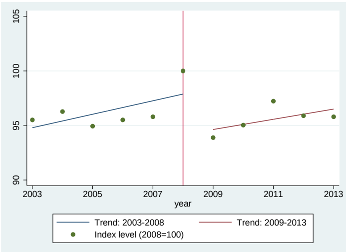
Figure 4.3: Manufacturing. Evolution of revenue TFP and its components over the period 2003-2013.



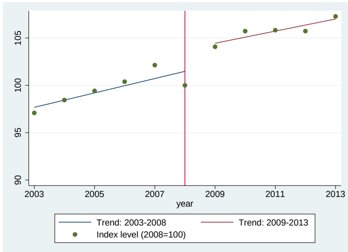
(a) WLD rev TFP



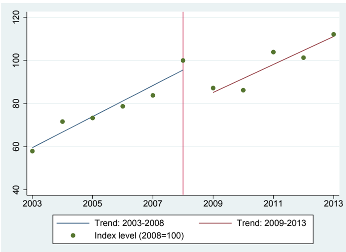
(b) DGKP rev TFP



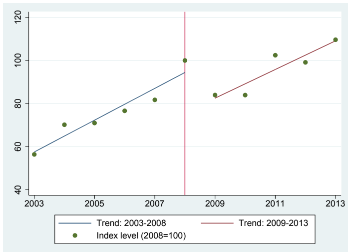
(c) TFPQ a



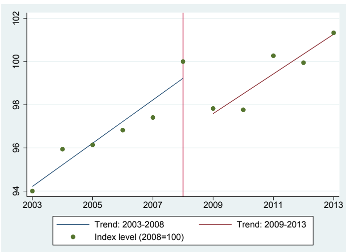
(d) Real price p



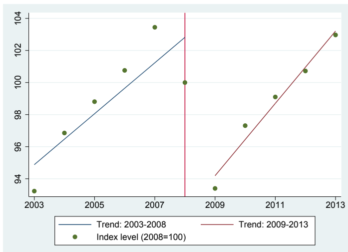
(e) Demand λ



(f) Composite ω



(g) Markups μ



(h) Production scale

Indices (2008=100) calculated using revenue-weighted changes between $t-1$ and t for the within sample of manufacturing firms. Panels (a) and (b) refer to revenue TFP computed using the WLD method and DGKP method, respectively. Panels (c)-(h) show real prices and components of the DGKP revenue TFP following from equation (4.9).

A common approach in the literature on the UK productivity puzzle is to decompose the shortfall in labour productivity into contributions of changes in factor inputs and TFP or, to be more precise, factor inputs and revenue TFP as highlighted by equation (4.13) in Section 4.2. Growth accounting exercises using sectoral national accounts data find that the labour productivity puzzle turns out to be a revenue TFP puzzle (Goodridge et al., 2013). That is, after accounting for changes to capital and labour inputs, the bulk of the 'lost' growth over 2008-2013 was due to a slower rate of revenue TFP growth. More similar to our paper, is a bottom-up econometric approach that estimates revenue-based production functions to obtain inputs parameters and firm-level TFP in Harris and Moffat (2017).[28] Harris and Moffat (2017) find that while in services the decline in labour productivity growth is mostly the result of a decline in revenue TFP growth, in manufacturing there is no revenue TFP puzzle: weighted plant-level revenue TFP barely changed or grew in the 2008-12 period. Instead, a measured 19% decline over 2007-12 in labour productivity is entirely due to changes in the intensity of inputs and in particular the log intermediates per labour ratio $\Delta(m_{it} - l_{it})$ in (4.13).

Table 4.14: Manufacturing. Standard labour productivity decomposition (factor proportions version) over the period 2003-2013.

	$\Delta(r-l)$	Δ TFP-R (DGKP)	$\gamma\Delta l$	$(\alpha_M)\Delta(m-l)$	$(\alpha_K)\Delta(k-l)$	Obs
2004	0.062	0.035	-0.000	0.026	0.001	641
2005	0.012	0.004	0.000	0.007	0.001	564
2006	0.044	0.023	-0.000	0.024	-0.002	617
2007	0.040	0.019	0.000	0.021	0.000	594
2008	-0.035	0.006	0.000	-0.038	-0.002	401
2009	-0.092	-0.046	-0.000	-0.042	-0.004	443
2010	0.118	0.042	-0.000	0.074	0.002	496
2011	0.038	0.030	0.000	0.009	-0.001	450
2012	-0.012	-0.008	0.000	-0.004	0.000	478
2013	0.007	0.011	0.000	-0.005	0.000	477
2003-2008	0.027	0.018	0.000	0.009	-0.001	2,817
2008-2013	0.012	0.006	0.000	0.006	-0.000	2,344

See equation (4.13) in Section 4.2 for the derivation of this standard labour productivity decomposition. Final two rows show the mean of changes over the two periods using all the annual observations shown above. Computations refer to revenue-weighted changes for the within sample of manufacturing firms.

Table 4.14 shows this standard labour productivity decomposition over the period 2003-2013 for our revenue-weighted manufacturing firms within sample.[29] Results indicate that manufacturing labour productivity growth in the period 2003-2008 was around 2.7pp a year and, while experiencing significant drops in 2008 and 2009, increased on average by

[28]It builds on previous work that estimates plant-level TFP (Harris and Drinkwater, 2000)

[29]In our analysis we use the firm wage bill, not the number of workers, as a measure of the labour input because we do not want potential changes in worker quality to affect the results. However, to show how changes in demand, TFP-Q and markups are important for labour productivity, it makes little sense to consider the (log of the) ratio between revenue and the wage bill on the left hand side of the decomposition. Instead we use, in our labour productivity decompositions, the (log) number of full-time-equivalent employees as a measure of the labour input and, in order to make sure that the decomposition goes through, we borrow our estimate of the output elasticity of labour, $\hat{\alpha}_L$, and recompute TFP-Q, TFP-R and scale accordingly. As can be appreciated from Table 4.14, this makes little difference in terms of the patterns of TFP-Q, TFP-R and scale so far discussed.

Table 4.15: Manufacturing. More detailed labour productivity decomposition (factor proportions version) over the period 2003-2013.

	$\Delta(r-l)$	$\Delta(a/\mu)$	$\Delta(\lambda/\mu)$	$\Delta[(\frac{\gamma+1}{\mu}-1)l]$	$\alpha_M\Delta(\frac{m-l}{\mu})$	$\alpha_K\Delta(\frac{k-l}{\mu})$	Obs
2004	0.062	-0.056	0.300	-0.139	-0.042	-0.001	641
2005	0.012	-0.012	0.004	0.007	0.012	0.001	564
2006	0.044	0.044	0.050	-0.053	0.005	-0.002	617
2007	0.040	-0.060	0.138	-0.035	-0.003	-0.001	594
2008	-0.035	-0.124	0.357	-0.151	-0.112	-0.005	401
2009	-0.092	0.011	-0.312	0.171	0.040	-0.002	443
2010	0.118	-0.020	0.082	-0.012	0.066	0.002	496
2011	0.038	-0.103	0.354	-0.150	-0.061	-0.003	450
2012	-0.012	0.002	-0.071	0.039	0.017	0.001	478
2013	0.007	-0.082	0.234	-0.091	-0.054	-0.001	477
2003-2008	0.027	-0.040	0.169	-0.074	-0.027	-0.002	2,817
2008-2013	0.012	-0.038	0.059	-0.009	0.001	-0.001	2,344

See equation (4.15) in Section 4.2 for the derivation of this more detailed labour productivity decomposition. Final two rows show the mean of changes over the two periods using all the annual observations shown above. Computations refer to revenue-weighted changes for the within sample of manufacturing firms.

only 1.2pp a year in the period 2008-2013 ending up almost 8pp below its per-crisis trend (more if considering 2003-2007 as the baseline period). Column 2 of Table 4.14 indicates that the main culprit of this under-performance is the drop in revenue TFP growth which changed from about 1.8pp a year in the period 2003-2008 to 0.6pp per year post-2008, i.e., $[(1.8-0.6)/(2.7-1.2)]$=80% of the labour productivity growth slowdown.[30] The remaining 20% is almost entirely accounted for by a reduction of the log intermediates to labour ratio, i.e., the term $\Delta(m-l)$.

Table 4.15, based on the more involved decomposition provided by equation (4.15) in Section 4.2, provides deeper insights on the decline in labour productivity by highlighting the role of demand, quantity TFP and markups. Markup-adjusted TFP-Q barely changed its growth rate in the two periods 2003-2008 and 2008-2013 while the largest growth rate drop is related to markup-adjusted demand $\tilde{\lambda}$ experiencing a decline from the 16.9pp per year average over 2003-2008 to only 5.9pp post-2008. At the same time, the related slowdown of markups seen in Table 4.13 helped to contain the fall in labour productivity through an increase in the average yearly growth rate of the markup-adjusted labour term ($\Delta[(\frac{\gamma+1}{\mu}-1)l]$) and intermediates over labour term ($\alpha_M\Delta(\frac{m-l}{\mu})$).[31] Markup-adjusted

[30]Numbers for DGKP revenue TFP in column 2 of Tables 4.13 and 4.14 are slightly different because in the latter case we use, as indicated in a previous footnote, the number of employees rather than the wage bill as a measure of the labour input.

[31]From the expressions of these two terms it appears clearly how a reduction in markups μ increases both.

capital over labour changes only weakly contributed throughout.

4.5.2 Services

Using the within sample for services, and weighting observations by revenues, we report in Table 4.16 mean changes from $t-1$ to t in WLD revenue TFP and its components building on the decomposition of equation (4.16). For services, we find more relevant to use the year 2007 to define the pre- and post-crisis periods.

First, revenue TFP growth in services was already rather weak before the crisis with an average of 0.2pp per year in the period 2003-2007. This further weakened after the crisis falling to 0.1pp per year over 2007-2013. In this respect, column 2 of Table 4.16 suggests (like for manufacturing) that a decline in the composite quantity TFP and demand component ω has been driving the weakening of TFP-R growth in services. At the same time, columns 3 and 4 indicate that this process has been accompanied by a decline in markups and production scale growth, which is again in line with the evidence provided above for manufacturing. Overall, this turned into a weakening of markup-adjusted ω (column 5) while the reduction in markups helped, like in manufacturing, containing the fall in revenue TFP through the increase in markup-adjusted scale (column 6), with the two adjusted components adding up to the overall change in revenue TFP as from equation (4.16).

Figure 4.4 presents the results graphically and it is constructed in the same way as Figure 4.3 for manufacturing except that we now set 2007 as the reference year. Panels (a) and (b) show quite neatly the break in revenue TFP and ω before and after 2007 while panels (c) and (d) highlight the decline in the evolution of markups and production scale. Panels (e) and (f) visualize the downturn in adjusted ω and the counter-increase in adjusted scale helping to contain the overall fall in TFP-R. Finally, Table C.8 in Appendix C-2 shows formal Chow test results confirming the presence of a structural break around 2007.

Table 4.17 shows the standard labour productivity decomposition over the period 2003-

Markups are endogenous in the MULAMA/MUOMEGA models and their equilibrium level (determined by profit maximization) increases with both TFP-Q and demand. A fall in demand and/or TFP-Q thus pushes markups to decrease and this decrease in markups helps firms to contain the fall in both profits and revenue TFP.

Table 4.16: Services. Changes of revenue TFP and its components over the period 2003-2013.

	Δ TFP-R (WLD)	$\Delta\omega$	$\Delta\mu$	Δ scale	Δ adjusted ω	Δ adjusted scale	Obs
2004	0.004	0.154	0.011	0.044	0.070	-0.066	8,387
2005	-0.006	0.028	0.001	0.044	0.002	-0.007	7,813
2006	-0.001	-0.032	-0.005	0.044	0.038	-0.039	6,438
2007	0.008	0.258	0.019	0.042	0.101	-0.092	6,266
2008	-0.020	-0.185	-0.015	0.016	-0.090	0.070	5,674
2009	-0.002	0.042	0.005	-0.041	-0.011	0.010	6,493
2010	0.011	0.124	0.009	0.014	0.068	-0.057	6,079
2011	-0.000	-0.203	-0.018	0.017	-0.088	0.088	5,966
2012	0.017	0.141	0.009	0.022	0.073	-0.056	6,404
2013	-0.002	-0.066	-0.006	0.031	0.006	-0.008	6,626
2003-2007	0.002	0.104	0.007	0.043	0.053	-0.052	28,904
2007-2013	0.001	-0.021	-0.002	0.010	-0.005	0.006	37,242

The Table shows mean revenue-weighted changes from $t-1$ to t, for the services firms within sample, of WLD revenue TFP as well as of the various MUOMEGA model components following from the linear revenue-TFP decomposition provided by equation (4.16) and applied to WLD revenue TFP. The final two rows show the mean of changes over the two periods using all the annual observations shown above.

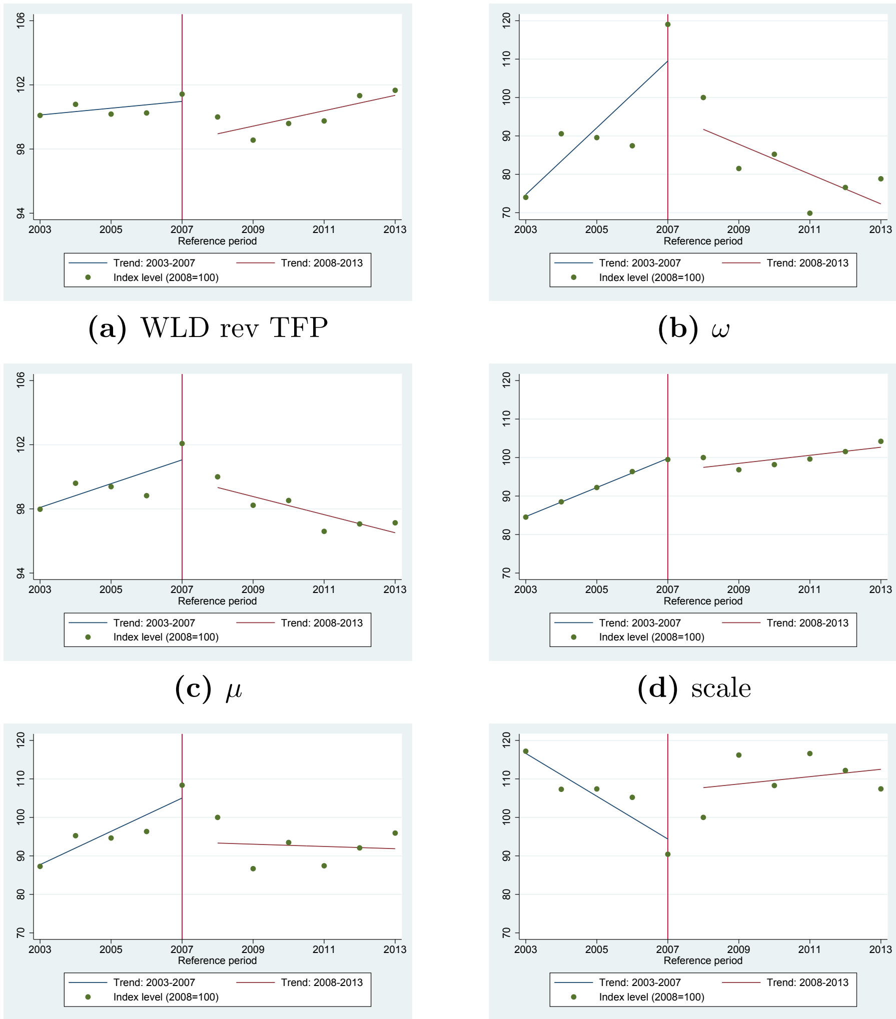
2013 for our revenue-weighted services firms within sample.[32] Results indicate that services labour productivity growth in the period 2003-2007 was about 0.8pp a year and, while experiencing large drops in 2008 and 2009, decreased on average by 1.1pp a year in the period 2007-2013 ending up around 11.4pp below its per-crisis trend. Column 2 of Table 4.17 indicates that a key culprit of this under-performance is the drop in revenue TFP growth which changed from a positive 0.3pp a year in the period 2003-2007 to a negative -0.6pp per year post-2007, i.e., $[(0.3+0.6)/(0.8+1.1)]$=47.4% of the labour productivity growth slowdown.[33] The remaining share is almost entirely accounted for by a reduction of the log intermediates to labour ratio, i.e., the term $\Delta(m-l)$.

Table 4.18 provides further insights on the decline in labour productivity by highlighting the combined role of demand and TFP-Q as well as of markups. Markup-adjusted ω experienced a strong growth decline from the 6.5pp per year average over 2003-2007 to only 0.1pp post-2007. At the same time, the related slowdown of markups helped to

[32] As in the case of manufacturing we use, in our labour productivity decompositions, the (log) number of full-time-equivalent employees as a measure of the labour input and, in order to make sure that the decomposition goes through, we borrow our estimate of the output elasticity of labour, $\hat{\alpha}_L$, and recompute ω, TFP-R and scale accordingly. As can be appreciated from Table 4.17, this makes little difference in terms of the overall patterns of ω, TFP-R and scale so far discussed.

[33] Numbers for WLD revenue TFP in column 1 of Table 4.16 and column 2 of Table 4.17 are different because in the latter case we use, as indicated in a previous footnote, the number of employees rather than the wage bill as a measure of the labour input.

Figure 4.4: Services. Evolution of revenue TFP and its components over the period 2003-2013.



(a) WLD rev TFP

(b) ω

(c) μ

(d) scale

(e) adjusted ω

(f) adjusted scale

Index of revenue TFP (2007=100) calculated using revenue-weighted changes in mean annual value and measured using the WLD method (a) for the within sample of services firms. Panels (b)-(f) show indices of components of the WLD revenue TFP following from the decomposition (4.16).

Table 4.17: Services. Standard labour productivity decomposition (factor proportions version) over the period 2003-2013.

	$\Delta(r-l)$	Δ TFP-R (WLD)	$\gamma\Delta l$	$(\alpha_M)\Delta(m-l)$	$(\alpha_K)\Delta(k-l)$	Obs
2004	0.011	0.006	0.002	0.001	0.001	8,387
2005	0.004	-0.004	0.003	0.004	0.001	7,813
2006	0.007	-0.002	0.003	0.006	-0.001	6,438
2007	0.009	0.011	0.002	-0.005	0.001	6,266
2008	-0.049	-0.036	0.008	-0.021	-0.000	5,674
2009	-0.069	-0.024	-0.001	-0.043	-0.002	6,493
2010	0.039	0.017	-0.001	0.022	0.001	6,079
2011	-0.003	-0.009	-0.000	0.007	0.000	5,966
2012	0.011	0.013	-0.002	-0.000	0.000	6,404
2013	-0.002	-0.004	0.003	-0.001	0.000	6,626
2003-2007	0.008	0.003	0.003	0.002	0.001	28,904
2007-2013	-0.011	-0.006	0.001	-0.006	-0.000	37,242

See equation (4.13) in Section 4.2 for the derivation of this standard labour productivity decomposition. Final two rows show the mean of changes over the two periods using all the annual observations shown above. Computations refer to revenue-weighted changes for the within sample of services firms.

Table 4.18: Services. More detailed labour productivity decomposition (factor proportions version) over the period 2003-2013.

	$\Delta(r-l)$	$\Delta(\omega/\mu)$	$\Delta[(\frac{\gamma+1}{\mu}-1)l]$	$\alpha_M\Delta(\frac{m-l}{\mu})$	$\alpha_K\Delta(\frac{k-l}{\mu})$	Obs
2004	0.011	0.062	-0.041	-0.011	0.000	8,387
2005	0.004	0.002	0.006	-0.004	0.001	7,813
2006	0.007	0.033	-0.020	-0.005	-0.001	6,438
2007	0.009	0.091	-0.059	-0.024	0.001	6,266
2008	-0.049	-0.092	0.051	-0.009	0.001	5,674
2009	-0.069	-0.031	0.006	-0.043	-0.002	6,493
2010	0.039	0.062	-0.033	0.009	0.001	6,079
2011	-0.003	-0.084	0.055	0.025	0.000	5,966
2012	0.011	0.065	-0.038	-0.016	0.000	6,404
2013	-0.002	0.001	-0.002	-0.001	0.000	6,626
2003-2007	0.008	0.047	-0.029	-0.011	0.000	28,904
2007-2013	-0.011	-0.011	0.005	-0.006	0.000	37,242

See equation (4.17) in Section 4.2 for the derivation of this more detailed labour productivity decomposition. Final two rows show the mean of changes over the two periods using all the annual observations shown above. Computations refer to revenue-weighted changes for the within sample of services firms.

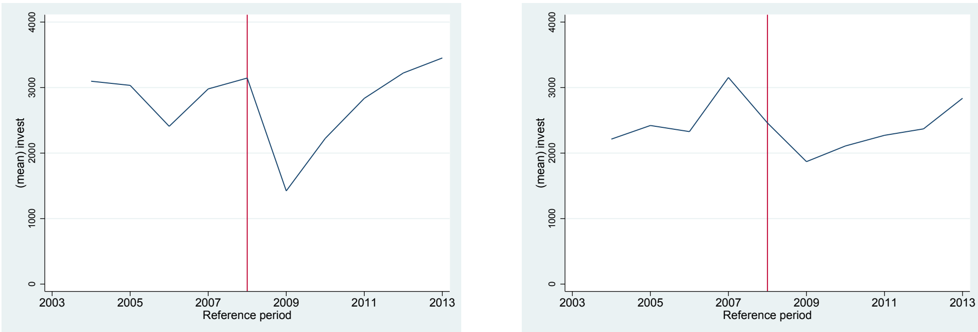
contain the fall in labour productivity through a substantial improvement in the average yearly growth rate of the markup-adjusted labour term ($\Delta[(\frac{\gamma+1}{\mu}-1)l]$) and intermediates over labour term ($\alpha_M\Delta(\frac{m-l}{\mu})$).[34] Finally, markup-adjusted capital over labour changes only weakly contributed throughout.

[34] As already highlighted above, markups are endogenous in the MULAMA/MUOMEGA models and their equilibrium level (determined by profit maximization) increases with ω. A fall in ω thus pushes markups to decrease and this decrease in markups helps firms to contain the fall in both profits and revenue TFP.

4.5.3 Comparing services with manufacturing

The evidence provided so far for manufacturing and services points to very similar patterns, although with somewhat different magnitudes, in terms of the common measures. More specifically, our results indicate that the labour productivity puzzle is to a large extent also a revenue TFP puzzle while the downturn in revenue TFP has been largely driven by a decline in the combined TFP-Q and demand component ω, to which firms have reacted by decreasing both markups and production scale. In turn, this decrease of markups and production scale has helped in containing the negative impact of the less favourable post-crisis environment on TFP-R and so also on labour productivity.

Figure 4.5: Investment patterns in manufacturing and services over the period 2004-2013.



(a) Manufacturing **(b)** Services

Mean firm annual real investments in the capital stock in manufacturing and services. Computations refer to the manufacturing and services within samples and are revenue-weighted.

In the case of manufacturing, our data allows us to go one step further and in particular assess whether the reduction in ω has been TFP-Q and/or demand driven. The answer to that question is that both a supply and a demand shock have negatively affected manufacturing revenue TFP. Our data does not allow us to directly answer the question for services. However we might conjecture, based on one key element, that supply (and possibly also demand) contributed to the overall downturn of services TFP-R. More specifically, we believe that the capital stock available to firms, which is generated by yearly investments, is the production input most closely related to the level of technology, i.e., quantity TFP used in the production process. In this respect, Figure 4.5 shows the mean, across firms and revenue-weighted, annual level of real investments in the capital stock in manufacturing (left panel) and services (right panel) over 2004-2013. The timing and depth of the decrease in investments and subsequent recovery are somewhat different between manufacturing and services but the overall picture is rather similar: a deep slump around the financial crisis and a sizeable (especially in manufacturing) recovery thereafter. Given that the fall in investments in manufacturing has turned into a sizeable post-crisis drop in TFP-Q, we can conjecture that a sizeable post-crisis drop in TFP-Q for services is also likely.

4.6 Robustness

In what follows we provide evidence supporting the robustness of our results by using different samples, different weighting schemes, different estimation techniques as well as a Translog production function.

Using the multi-product firms sample for manufacturing. In our main analysis for manufacturing we focus on single-product firms because dealing with multi-product firms requires a number of additional assumptions. However, multi-product firms account for a large share of production and revenue in manufacturing. In order to analyse multi-product firms we proceed as in FMMM and DGKP, i.e., we break them down into several single-product firms by using a procedure to assign firm-level inputs to the different products produced by a multi-product firm (inputs assignment problem). In doing so, we then consider again within firm and product changes between $t-1$ and t and weigh observations based on the corresponding firm-product-specific revenue. Results displayed in Table C.9 and Figure C.3 indicate that our key insights apply to the sample of multi-product manufacturing firms too.

Using un-weighted or employment-weighted values. So far we have always considered firm revenue in order to weight observations because we want our results to be representative of aggregate rather than average-firm outcomes. However, in Tables C.10 and C.11, and corresponding Figures C.4 and C.5, we provide results obtained using un-weighted changes. As can be appreciated from the two Tables and Figures, our baseline results are virtually unaffected. At the same time, Tables C.12 and C.13, and corresponding Figures C.6 and C.7, show results obtained using firm employment to weight observations and still confirm the robustness of our findings.

Using alternative estimation procedures. In our baseline results, we use the DGKP estimation procedure to estimate the production function and recover the different components of the MULAMA model for manufacturing, while for services we use the WLD estimation procedure to estimate the production function and recover the different components of the restricted MUOMEGA model. In order to assess the robustness of our results to the specific estimation technique employed we provide in Tables C.14 and C.15, as well as in Figures C.8 and C.9, results obtained using the FMMM estimation procedures for the MULAMA model (manufacturing) and MUOMEGA model (services). This new set of results is again in line with our baseline findings.

Using the Translog production function for manufacturing. The limited overlap between the Prodcom and ARDx datasets force us to estimate a unique production function for manufacturing firms rather than estimate different production functions for different 2-digit industries. In this respect, results provided in Table C.16 and Figure C.10, and obtained using the more flexible Translog production function, allay concerns about the issue of heterogeneity in output elasticities across firms and industries in manufacturing.

4.7 Conclusions

In this paper we provide novel evidence that UK firms' poor productivity performance post-2008 is due to a downturn in both quantity TFP and demand pushing down sales, markups and revenue TFP, as well as labour productivity. More specifically, in the first part of our analysis we focus on manufacturing firms and use information on firm-level prices and quantities to measure firm-level quantity TFP by building upon the frameworks developed in De Loecker et al. (2016) and Forlani et al. (2016). This allows us to further quantify firm-level demand and markups and, while aggregating-up the information at the manufacturing industry-level, compare the evolution of TFP-Q, markups and demand before and after 2008. Finally, we exploit two exact decompositions for TFP-R and labour productivity to show how changes in TFP-Q, markups and demand have affected the two productivity measures. Our results suggest that both a slowing down of demand and a decline in quantity TFP, and the related markups fall, are behind the decline in revenue TFP and labour productivity in manufacturing.

In the second part of our analysis, we instead consider service industries and estimate a restricted version of the model due to the absence of reliable and meaningful information on prices. In doing so we find, for those measures that are common to both the full and restricted versions of the model, very similar patterns to those obtained for manufacturing. These findings, along with the absence of noticeable differences in capital investments patterns between manufacturing and services industries, lead us to conjecture that both supply and demand also contributed to the poor revenue TFP and labour productivity performance of UK service industries.

We believe that our results are important for at least two reasons. First, they are informative about the long-term impacts of the Great Recession. A fall in quantity TFP, due for example to a decline in the rate of technical progress, represents a permanent loss of productive potential with substantial long-term implications for the economy. By contrast a demand downturn, due for example to a prolonged general climate of uncertainty, could have less permanent consequences. Second, they are informative about the policies that could more effectively address the weak growth of labour productivity and revenue TFP post-crisis. In particular, our findings suggest that government policies should more prominently act towards boosting demand for UK firms rather than focusing only on productivity.

Chapter 5

Concluding remarks

In this thesis I have exploited recent advances in estimating firm-level quantity-based production functions to bring a new perspective on three long-standing questions of interest to economists. First, I have asked whether improvements in product quality have a role to play in a rise in price-marginal cost markups that have been observed following tariff reductions during the Indian trade liberalisation of the 1990s. Then, together with a co-author, I have asked whether the well-known finding that (revenue) productivity rises with city size is due to firms being better able to turn a basket of physical inputs into output (quantity TFP), or does it arise because large-city based firms are able to charge higher prices? And finally, again with my co-author, I have explored UK firms' post-crisis productivity performance to ask if the fall in the growth rate of labour productivity and revenue TFP is due to supply-side or demand-side factors.

To answer these questions I have made use of detailed quantity, prices and revenue data for India, France and the UK respectively, and used it to first estimate quantity-based production functions and then, building on FMMM, quantified heterogeneity in TFP, demand and markups across firms.

In particular, Chapter 2 provided evidence that the impacts of the tariff liberalization on prices, marginal costs and markups are driven by the reductions in input tariffs that allowed firms to raise product quality. While both DGKP and my paper show factory gate prices falling by less than marginal costs, implying that firms received most of the benefits of the trade liberalization, the finding in my paper that product quality rose along with markups is evidence that consumers benefited alongside firms. Chapter 2 also provided evidence consistent with two findings in the literature otherwise difficult to reconcile: that the Indian trade reform raised within-firm productivity; and that it was not associated with high product churn, i.e., the dropping by firms of badly-performing products. Instead, product quality and quantity TFP gains were concentrated in previously laggard products as lower input tariffs relaxed constraints.

In Chapter 3 we aggregated measures of TFP, demand and markups for French manufacturing firms to the regional level. Our results indicate that first, that market shares across firms with heterogeneous productivities are better allocated in denser areas than less-dense areas, so amplifying in aggregate revenue-weighted figures firm-level differences in productivity across space. Second, a substantial portion of the aggregate revenue productivity advantage of denser areas stems from product composition effects. And third, that manufacturing firms located in denser areas are not necessarily characterized by a significantly higher quantity TFP, but they do have a revenue TFP advantage due to their capacity to produce and sell higher demand products at higher prices and in larger quantities with lower markups compared to firms located in less dense areas. They also produce products with higher marginal costs and part of these higher costs reflects higher product quality.

Chapter 4 provides novel evidence that UK firms' poor productivity performance post-2008 was due to a downturn in both quantity TFP and demand pushing down sales, markups and revenue TFP, as well as labour productivity. Our results suggest that both a slowing down of demand and a decline in quantity TFP, and the related markups fall, are behind the decline in revenue TFP and labour productivity in manufacturing. In the service industries we find, for those measures that are common to both the full and restricted versions of the model, very similar patterns to those obtained for manufacturing.

What is common to the results presented in all three chapters is the evidence they show of an important role for factors other than quantity TFP in measures of firm performance: product quality/appeal in the response of firms to lower tariffs in India (Chapter 2), and the revenue productivity of firms in France (Chapter 3); or demand-side conditions to firms' labour productivity and revenue TFP after the financial crisis in the UK (Chapter 4); and the allocation of resources between firms to regional differences in revenue productivity again in France.

With detailed production datasets becoming more widely available, being able to disentangle price changes from revenue productivity, demand and markups promises many possible avenues for future research on a variety of topics linked to firm productivity.

Bibliography

Ackerberg, D., Benkard, C. L., Berry, S., and Pakes, A. (2007). Econometric tools for analyzing market outcomes. *Handbook of Econometrics*, 6:4171–4276. 61, 67, 95, 98

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, 83(6):2411–2451. 2, 49, 89, X, XVII

Amiti, M., Itskhoki, O., and Konings, J. (2014). Importers, exporters, and exchange rate disconnect. *American Economic Review*, 104(7):1942–1978. 67

Amiti, M. and Khandelwal, A. K. (2013). Import Competition and Quality Upgrading. *The Review of Economics and Statistics*, 95(2):476–490. 10, 11

Amiti, M. and Konings, J. (2007). Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. *American Economic Review*, 97(5):1611–1638. 9, VI

Anderson, J. E. and Van Wincoop, E. (2004). Trade costs. *Journal of Economic Literature*, 42(3):691–751. 116

Andrews, D., Criscuolo, C., and Gal, P. N. (2015). Frontier Firms, Technology Diffusion and Public Policy: Micro Evidence from OECD Countries. *OECD Productivity Working Papers*, 2. 88

Antoniades, A. (2015). Heterogeneous Firms, Quality, and Trade. *Journal of International Economics*, 95(2):263–273. 9, 10, 30, 32, 37

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, 98(5):1998–2031. 20, 56, 91

Ayoubkhani, D. (2014). A Comparison between Annual Business Survey and National Accounts Measures of Value Added. Technical report, Office for National Statistics, Newport. 104, XXXI

Barnett, A., Chiu, A., Franklin, J., and Sebastiá-Barriel, M. (2014). The productivity puzzle: a firm-level investigation into employment behaviour and resource allocation over the crisis. *Bank of England Staff Working Papers*, 495. 88, 101

Behrens, K., Mion, G., Murata, Y., and Südekum, J. (2014). Trade, wages, and productivity. *International Economic Review*, 55(4):1305–1348. 20, 56, 91

Behrens, K., Mion, G., Murata, Y., and Suedekum, J. (2017). Spatial frictions. *Journal of Urban Economics*, 97:40–70. 71

Berman, N., Martin, P., and Mayer, T. (2012). How do different exporters react to exchange rate changes? *The Quarterly Journal of Economics*, 127(1):437–492. 67

Bernard, A. B., Eaton, J., Jensen, J. B., and Kortum, S. (2003). Plants and Productivity in International Trade. *The American Economic Review*, 93(4):1268–1290. 8

Bernard, A. B., Redding, S. J., and Schott, P. K. (2010). Multiple-product firms and product switching. *American Economic Review*, 100(1):70–97. 42

Bollard, A., Klenow, P. J., and Sharma, G. (2013). Indias mysterious manufacturing miracle. *Review of Economic Dynamics*, 16(1):59–85. 40, 42

Broda, C. and Weinstein, D. E. (2006). Globalization and the Gains From Variety. *The Quarterly Journal of Economics*, 121(2):541–585. 11

Bryson, A. and Forth, J. (2016). The UK 's Productivity Puzzle. In Askenazy, P., Bellman, L., Bryson, A., and Moreno Galbis, E., editors, *Productivity Puzzles Across Europe*, chapter 5, pages 129–173. Oxford University Press. 86, 88

Campa, J. M. and Goldberg, L. S. (2005). Exchange rate pass-through into import prices. *The Review of Economics and Statistics*, 87(4):679–690. 67

Cette, G., Fernald, J., and Mojon, B. (2016). The pre-Great Recession slowdown in productivity. *European Economic Review*, 88:3–20. 86

Chakraborty, P. and Raveh, O. (2018). Input-trade liberalization and the demand for managers: Evidence from India. *Journal of International Economics*, 111:159–176. 21

Ciccone, A. (2002). Agglomeration effects in Europe. *European Economic Review*, 46(2):213–227. 48

Ciccone, A. and Hall, R. E. (1996). Productivity and the Density of Economic Activity. *The American Economic Review*, 86(1):54–70. 48

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742. 45, 48

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection.

Econometrica, 80(6):2543–2594. 48, 63, 83, 84

Combes, P.-P. and Gobillon, L. (2015). The Empirics of Agglomeration Economies. In *Handbook of Regional and Urban Economics*, chapter 5, pages 247–348. Elsevier. 45, 48, 50, 72

De La Roca, J. and Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, 84(1):106–142. 48

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications. *The Quarterly Journal of Economics*, 135(2):561–644. 89

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, Markups, and Trade Reform. *Econometrica*, 84(2):445–510. 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 21, 22, 24, 25, 26, 28, 29, 31, 32, 34, 35, 36, 37, 39, 42, 48, 49, 87, 89, 128, 130, II, III, XVII

De Loecker, J. and Warzynski, F. (2012). Markups and Firm-Level Export Status. *American Economic Review*, 102(6):2437–2471. 12, 16, 57, 92

Debroy, B. and Santhanam, A. (1993). Matching Trade Codes with Industrial Codes. *Foreign Trade Bulletin*, 24(1):5–27. 24

Dhingra, S. and Morrow, J. (2019). Monopolistic competition and optimum product diversity under firm heterogeneity. *Journal of Political Economy*, 127(1):196–232. 20, 56, 67, 91

Duranton, G. and Puga, D. (2004). Micro-Foundations of Urban Agglomeration Economies. In *Handbook of Regional and Urban Economics*, volume 4, chapter 48, pages 2063–2117. Elsevier. 45

Fan, H., Li, Y. A., and Yeaple, S. R. (2015). Trade Liberalization, Quality, and Export Prices. *The Review of Economics and Statistics*, 97(5):1033–1051. 10, 37

Fan, H., Li, Y. A., and Yeaple, S. R. (2018). On the relationship between quality and productivity: Evidence from China's accession to the WTO. *Journal of International Economics*, 110:28–49. 10, 37, 39

Feenstra, R. C. (2003). A homothetic utility function for monopolistic competition models, without constant price elasticity. *Economics Letters*, 78(1):79–86. 20, 56, 91

Feenstra, R. C. (2018). Restoring the product variety and pro-competitive gains from trade with heterogeneous firms and bounded productivity. *Journal of International Economics*, 110:16–27. 11

Forlani, E., Martin, R., Mion, G., and Muûls, M. (2016). Unraveling Firms: Demand, Productivity and Markups Heterogeneity. *CEPR Discussion Paper*, 11058. 2, 3, 4, 5, 7, 9, 11, 12, 13, 18, 19, 20, 21, 30, 42, 46, 49, 87, 89, 128, 130

Goldberg, P., Khandelwal, A., Pavcnik, N., and Topalova, P. (2009). Trade Liberalization and New Imported Inputs. *American Economic Review: Papers & Proceedings*, 99(2):494–500. 8, 9

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., and Topalova, P. (2010a). Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. *The Quarterly Journal of Economics*, 125(4):1727–1767. 8

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., and Topalova, P. (2010b). Multiproduct firms and product turnover in the developing world: Evidence from India. *Review of Economics and Statistics*, 92(4):1042–1049. 9, 21, 40, 42

Goodridge, P., Haskel, J., and Wallis, G. (2013). Can Intangible Investment Explain the UK Productivity Puzzle? *National Institute Economic Review*, 224(1):48–58. 88, 119

Goodridge, P., Haskel, J., and Wallis, G. (2016). Accounting for the UK Productivity Puzzle: A Decomposition and Predictions. *Economica*, 85:581–605. 88, 96

Grice, J. (2012). The Productivity Conundrum, Interpreting the Recent Behaviour of the Economy. Technical report, Office for National Statistics, Newport. 86, 88

Griliches, Z. and Mairesse, J. (1995). Production Functions: The Search for Identification. *NBER Working Paper Series*, (5067). 64, 112

Hall, R. E. (1986). Market Structure and Macroeconomic Fluctuations. *Brookings Papers on Economic Activity*, 1986(2):285. 2, 12, 16, 57, 89, 92

Haltiwanger, J., Kulick, R., and Syverson, C. (2018). Misallocation measures: The distortion that ate the residual. *NBER Working Paper*, 14491. 20, 56, 91

Harris, R. and Moffat, J. (2017). The UK productivity puzzle, 2008 to 2012: evidence using plant-level estimates of total factor productivity. *Oxford Economic Papers*, 69(3):529–549. 88, 96, 101, 104, 120, XXXIII

Harris, R. I. D. and Drinkwater, S. (2000). UK Plant and Machinery Capital Stocks and Plant Closures. *Oxford Bulletin of Economics and Statistics*, 62(2):243–265. 120, XXXIII

Harrison, A. E., Martin, L. A., and Nataraj, S. (2013). Learning versus Stealing: How Important Are Market-Share Reallocations to India's Productivity Growth? *The World*

Bank Economic Review, 27(2):202–228. 8, 40, 42

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364. 20, 56, 91, XVIII

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., and Yesiltas, S. (2015). How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts and Aggregate Implications. NBER Working Papers 21558, National Bureau of Economic Research, Inc. 23

Khandelwal, A. (2010). The Long and Short (of) Quality Ladders. *The Review of Economic Studies*, 77(4):1450–1476. 10, 11, 18, 37

Khandelwal, A. K., Schott, P. K., and Wei, S.-J. (2013). Trade liberalization and embedded institutional reform: Evidence from Chinese exporters. *American Economic Review*, 103(6):2169–2195. 115

Kremer, M. (1993). The O-Ring Theory of Economic Development. *The Quarterly Journal of Economics*, 108(3):551–575. 31

Kugler, M. and Verhoogen, E. (2011). Prices, plant size, and product quality. *The Review of Economic Studies*, 79(1):307–339. 9, 10, 37, 38, 60, 94

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies*, 70(70):317–341. 12, 13, 14, 49

Manova, K. and Zhang, Z. (2012). Export prices across firms and destinations. *The Quarterly Journal of Economics*, 127(1):379–436. 37, 38

Marshall, A. (1890). *The Principles of Economics.* McMaster University Archive for the History of Economic Thought. 44, 45

Martin, B. and Rowthorn, R. (2012). Is the British economy supply constrained II? A renewed critique of productivity pessimism. *Centre for Business Research, University of Cambridge*. 88

McCann, P. (2018). Productivity Perspectives Synthesis. Productivity insights network evidence review 07, ESRC Productivity Insights Network. 88

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725. 71

Melitz, M. J. and Ottaviano, G. I. P. (2008). Market size, trade, and productivity. *The Review of Economic Studies*, 75(1):295–316. 8

Mion, G. and Naticchioni, P. (2009). The spatial sorting and matching of skills and firms. *Canadian Journal of Economics/Revue canadienne d'{é}conomique*, 42(1):28–55. 48

Mrázová, M. and Neary, J. P. (2017). Not so demanding: Demand structure and firm behavior. *American Economic Review*, 107(12):3835–3874. 67

Nataraj, S. (2011). The impact of trade liberalization on productivity: Evidence from India's formal and informal manufacturing sectors. *Journal of International Economics*, 85(2):292–301. 8, 39

OECD (2018). *OECD Compendium of Productivity Indicators 2018*. OECD Publishing, Paris. 86

Office for National Statistics (2007). *The ONS Productivity Handbook*. Palgrave-Macmillan. 97

Office for National Statistics (2014). Producer Price Indices: Methods and Guidance. Technical report. 114

Office for National Statistics (2018a). International comparisons of productivity. Technical report. 86

Office for National Statistics (2018b). Labour productivity, UK: October to December 2017. Statistical bulletin. 86

Office for National Statistics (2018c). Products of the European Community, 1997-2016: Secure Access. (8th Edition). [data collection]. 102

Office for National Statistics. Virtual Microdata Laboratory, University of the West of England, Bristol (2017). Annual Respondents Database X, 1998-2014: Secure Access (4th Edition). [data collection]. 101

Olley, G. S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–1297. 2, 12, 13, 15, 48, 59, 89, 93

Pessoa, J. P. and Van Reenen, J. (2014). The UK productivity and jobs puzzle: Does the answer lie in wage flexibility? *Economic Journal*, 124(576):433–452. 86, 88, 96

Riley, R., Bondibene, C. R., and Young, G. (2013). Productivity Dynamics in the Great Stagnation: Evidence from British businesses. CFM Discussion Papers 1407, Centre for Macroeconomics. 88, 101

Riley, R., Rincon-Aznar, A., and Samek, L. (2018). Below the Aggregate: A Sectoral Account of the UK Productivity Puzzle. ESCoE Discussion Papers 2018-06, Economic

Statistics Centre of Excellence. 88

Riley, R., Rosazza-Bondibene, C., and Young, G. (2014). The Financial Crisis, Bank Lending and UK Productivity: Sectoral and Firm-Level Evidence. *National Institute Economic Review*, 228(1):17–34. 88

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. In *Handbook of Regional and Urban Economics*, volume 4, chapter 49, pages 2119–2171. Elsevier. 45, 46, 48

Schneider, P. (2018). Decomposing differences in productivity distributions. Bank of England working papers 740, Bank of England. 88

Sivadasan, J. (2009). Barriers to Competition and Productivity: Evidence from India. *The B.E. Journal of Economic Analysis & Policy*, 9(1 (Advances)):Article 42. 8

Spence, M. (1976). Product selection, fixed costs, and monopolistic competition. *The Review of Economic Studies*, 43(2):217–235. 20, 56, 91

Stiebale, J. and Vencappa, D. (2018). Acquisitions, markups, efficiency, and product quality: Evidence from India. *Journal of International Economics*, 112:70–87. 11, 21, 115

Stock J.H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. *In D.W.K. Andrews and J.H. Stock (Eds.), Identification and Inference for Econometric Models, Essays in Honor of Thomas Rothenberg,*, 80-108. 73, 78, 79, 80, 81, XXIII, XXIV, XXV, XXVI, XXVII, XXVIII, XXIX

Syverson, C. (2004). Market Structure and Productivity: A Concrete Example. *Journal of Political Economy*, 112(6):1181–1222. 78

Syverson, C. (2008). Markets: Ready-mixed concrete. *Journal of Economic Perspectives*, 22(1):217–234. 78

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49(2):326–365. 45

Topalova, P. (2010). Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India. *American Economic Journal: Applied Economics*, 2(4):1–41. 9, 24

Topalova, P. and Khandelwal, A. (2011). Trade liberalization and firm productivity: The case of India. *Review of Economics and Statistics*, 93(3):995–1009. 8, 12, 24, 25

Van Beveren, I., Bernard, A., and Vandenbussche, H. (2012). Concording EU Trade and

Production Data over Time. *NBER Working Papers*, 18604. 51, 103, XXXIII

Verhoogen, E. A. (2008). Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector. *The Quarterly Journal of Economics*, 123(2):489–530. 9, 10, 14, 18, 37, 38

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114. 15, 49, 59, 65, 68, 89

Zhelobodko, E., Kokovin, S., Parenti, M., and Thisse, J.-F. (2012). Monopolistic competition: Beyond the constant elasticity of substitution. *Econometrica*, 80(6):2765–2784. 20, 56, 67, 91

Appendices

Appendix A

Appendix to Chapter 2: Trade Reform Redux: Prices, Markups and Product Quality

A-1 Replication of DGKP results

This Appendix describes how my dataset replicates the main results presented in DGKP. As described in Section 2.5, the preparation of my estimation sample delivers a sample size of 27,504 firm-product-year observations, substantially more than the 21,246 in DGKP, however this delivers quantitatively similar and qualitatively the same main results. I start by showing in Table A.1 that although there are some large differences in mean industry-level markups between DGKP and my data this is likely driven by outliers since the medians are in almost all cases quite close. For instance, in NIC industry 31, "Electrical machinery & communications", DGKP report a mean markup of 5.66 while I find a mean of 2.77, but both datasets deliver a median of 1.43.

DGKP present four key empirical results. Exploiting variation within firm-product pairs they show (1) the pass-through of marginal cost to prices is incomplete; (2) that the large fall in output tariffs over the 1989-1997 period is associated with only a small fall in prices; (3) that the fall in input tariffs reduced marginal costs – and hence via (1), generated higher markups; and (4) that when controlling for marginal cost in a regression of markups on output tariffs so as to estimate the pure pro-competitive effect of trade reform, that these effects helped to moderate the rise in markups.

Table A.1: Markups by sector

Sector		Original		New	
		Mean	Median	Mean	Median
15	Food products & beverages	1.78	1.15	1.68	1.29
17	Textiles, apparel	1.57	1.33	1.96	1.48
21	Paper & paper products	1.22	1.21	1.86	1.54
24	Chemicals	2.25	1.36	2.03	1.26
25	Rubber & plastic	4.52	1.37	1.62	1.29
26	Nonmetallic mineral products	4.57	2.27	3.24	1.57
27	Basic metals	2.54	1.20	1.69	1.21
28	Fabricated metal products	3.70	1.36	1.75	1.14
29	Machinery & equipment	2.48	1.34	2.48	1.33
31	Electrical machinery & communications	5.66	1.43	2.77	1.43
34	Motor vehicles, trailers	4.64	1.39	2.39	1.20
	Average	2.70	1.34	2.01	1.34

Notes: Table shows mean and median markups by sector for the 1989-2003 sample. 'Original' shows results from DGKP, Table VI, page 483, while 'New' shows the dataset from this paper. Both use a +/- 3% trimming of the markup distribution by sector.

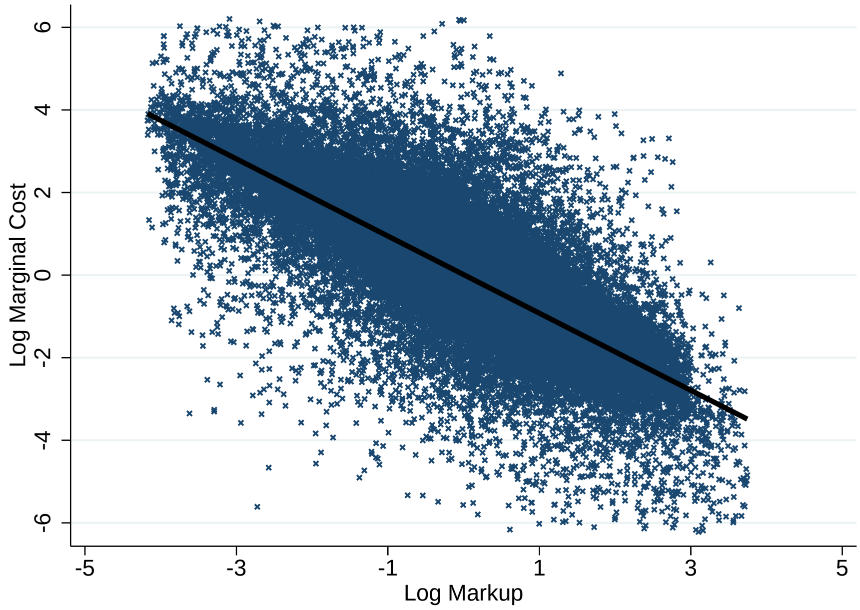
To show incomplete pass-through, DGKP start by re-writing the identity $p_{fjt} = \ln mc_{fjt} + \ln \mu_{fjt}$ by splitting markups into a time-invariant firm-product component and the time-varying deviation:

$$p_{fjt} = \ln \mu_{fj} + \ln mc_{fjt} + (\ln \mu_{fjt} - \ln \mu_{fj}) \tag{A-1}$$

If markups are variable then DGKP argue the deviation of the markup from its average will be correlated with marginal costs. When firms face a demand curve that features a price elasticity of demand that is increasing in price, an increase in marginal cost that raises price will be associated with a higher elasticity and a lower profit-maximising markup. Marginal costs will then be negatively correlated with markups, so pass-through of changes in marginal cost to changes in price will be incomplete. DGKP argue that any demand system that delivers variable markups will also deliver incomplete pass-through. In the data, there indeed appears to be a strong correlation between markups and marginal cost. Figure A.1 uses the same within firm-product pair variation used in the main analysis, demeaning log markups and log marginal cost by firm-product fixed effects, to show a strong negative relationship between markups and marginal costs.

To estimate the degree of pass-through DGKP run an OLS regression of (log) prices on (log) marginal cost and (log) markups at the firm-product-year level, together with firm-product fixed effects:

Figure A.1: Markups vs. Marginal cost



Notes: Markups and marginal costs demeaned by firm-product-year fixed effects. Outliers of +/- 3% of markup and marginal cost trimmed.

$$p_{fjt} = \alpha_{fj} + \zeta \ln mc_{fjt} + \epsilon_{fjt} \tag{A-2}$$

If markups were constant we would find full pass-through $\hat{\zeta} = 1$ and a perfect fit of the regression line ($\epsilon_{fjt} = 0$), but with variable markups we would find $\hat{\zeta} < 1$, i.e. incomplete pass-through. However, estimation of (A-2) via OLS would only yield unbiased estimates of ζ if marginal cost could be directly observed. Using $\ln \hat{mc}_{fjt}$ introduces measurement error, leading to downward biased $\hat{\zeta}$ and risking erroneously finding incomplete pass-through. To deal with this issue, DGKP instrument for marginal costs using input tariffs (which vary only at the industry level) and lagged marginal costs (which vary at the firm-product level). They do not require these lagged marginal costs to be uncorrelated with the error term from (A-2), only uncorrelated with the measurement error, since the correlation between markups and marginal costs is exactly what ζ should measure. In some specifications they add a second lag of marginal cost in order to assuage concerns about serial correlation in the errors.

DGKP report results from running the pass-through regression (A-2) using OLS (finding a coefficient of 0.337, significant at the 1% level), and with the IV approach that instruments for marginal cost with input tariffs and once lagged marginal cost (coefficient of 0.305, significant at 1% level) and that instruments for marginal cost using input tariffs and the second lag of marginal costs (coefficient of 0.406, significant at the 10.1% level). However it is not clear from these results whether it is the use of the instrumental variables to deal with measurement error that is driving the change in estimates, or whether it is the smaller sample sizes associated with using one or two lags of data. I use the replication data file to confirm that repeating the OLS regression using the first-lag IV sample delivers a coefficient of 0.264 (significant at 1% level), while using the second-lag IV sample with OLS delivers a coefficient of 0.240 (significant at 1% level). These OLS coefficients are both smaller than the comparable IV estimates using the same samples, indicating that the use of instruments does appear to remove the downward bias in the pass-through coefficient that stems from measurement error. However, as the sample size reduces as the lag structure increases, so does the estimated coefficient, meaning it is not clear that DGKP and this paper do not continue to underestimate the degree of pass-through.

Table A.2: Passthrough of marginal costs to prices

	ln *price*		
	(1)	(2)	(3)
$\ln mc$	0.413	0.303	0.370
	(0.029)***	(0.055)***	(0.044)***
R^2	0.35		0.29
N	27,504	15,848	15,848
Firm-product FEs	yes	yes	yes
Instruments	-	yes	-
First-stage F-stat		64.8	

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Corresponds to DGKP, Table VII, p.488. Standard errors clustered by firm. All regressions include a constant and firm-product fixed effects. Column 1 reports OLS results, column 2 instruments log of marginal cost with input tariffs and the first lag of marginal costs, column 3 reports OLS results using the same sample as column 2. First stage F-stat is the F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. R^2 is the Within R-squared.

Table A.2 shows that results using my dataset are qualitatively similar to DGKP with a low degree of pass-through, a coefficient of 0.413 in column (1). As discussed above, I may be underestimating due to measurement error in marginal costs. Column (2) repeats the IV approach of DGKP and finds that the coefficient falls to 0.303, from a sample of around half the size due to the need for lagged values as instruments. In column (3), I use the same, smaller, sample to run the OLS regression and find that the smaller sample by itself delivers a lower value of $\hat{\zeta}$, making conclusions about the true value imprecise but indicating that pass-through is far from complete.

Having established this basic feature of the data, DGKP's second key result is the finding that the 62 percentage point lowering of output tarrifs during the Indian trade liberalisation from 1989-1997 was associated with a reduction of (real) prices of just 10.4%. To do so, they estimate (A-3), using variation within firm-product pairs.

$$p_{fjt} = \alpha_{fj} + \alpha_{st} + \beta_1 \tau_{it}^{out} + \epsilon_{fjt} \tag{A-3}$$

They report a β_1 of 0.136 (significant at 5%) when controlling for year fixed effects and a β_1 of 0.167 (significant at 1%) when controlling for sector-year fixed effects (so controlling for industry-specific cycles). This estimate implies that a 10pp fall in output tariffs reduces prices by 1.67%, and by multiplying the coefficient by the average tariff change over the period, the overall impact is calculated as a 10.4% decline in prices, with a standard error of 3.3 (significant at the 1% level).

My estimates of β_1, shown in Table A.3, are slightly higher but qualitatively identical. The output average tariff change in my data is 68.3 percentage points, which leads to an overall impact on prices of -12.1% (significant at 1%) when using sector-year fixed effects.

Table A.3: Prices and output tariffs

	$\ln price$	$\ln price$
τ^{output}	0.145	0.177
	(0.069)**	(0.050)***
R^2	0.01	0.03
N	27,504	27,504
Firm-product FEs	Yes	Yes
Year FEs	Yes	No
Sector-Year FEs	No	Yes
Overall impact	-9.9	-12.1
Standard error	4.7	3.4

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Corresponds to DGKP, Table VIII, p.490. Standard errors clustered by industry. All regressions include a constant and firm-product fixed effects. Column 1 reports OLS results with year fixed effects and column 2 reports OLS results with sector×year fixed effects. The last two rows use the average decline in output tariffs of 68pp to compute the mean and standard error of the impact on prices. R^2 is the Within R-squared.

The third key result is to build on evidence from Amiti and Konings (2007) and others, that firms' responses to the fall in input tariffs associated with trade reform are an important part of the adjustment. By estimating (A-4), separately including output and input tariffs, DGKP show the independent effects of output and input tariffs. Results for my data are shown in Column (1) of Table A.4.

$$p_{fjt} = \alpha_{fj} + \alpha_{st} + \beta_1 \tau_{it}^{out} + \beta_2 \tau_{it}^{in} + \epsilon_{fjt} \tag{A-4}$$

Controlling for input tariffs barely changes the coefficient on output tariffs estimated in Table A.3 column (2), while the coefficient on input tariffs itself is not different from zero at any reasonable level of significance. This suggests firms adjust prices to demand-side conditions but not to changes to supply-side inputs-market conditions. This result is similar to that found by DGKP – a 0.156 output tariff coefficient, significant at 1% level and a 0.352 coefficient on input tariffs, not significant at any reasonable level. For an overall effect of trade reform on prices, DGKP find an 18.1% fall, significant at 5%, larger and more precise than the 11.2% fall in my data, largely due to the influence of a larger, if still imprecise, effect of input tariffs.

Table A.4: Regressions of prices, marginal costs and markups on output and input tariffs

	$\ln price$	$\ln mc$	$\ln \mu$
τ^{output}	0.1788	0.0664	0.1125
	(0.0590)***	(0.0591)	(0.0647)*
τ^{input}	-0.0385	0.7862	-0.8247
	(0.5066)	(0.4408)*	(0.2952)***
R^2	0.03	0.02	0.01
N	27,504	27,504	27,504
Firm-product FEs	Yes	Yes	Yes
Sector-Year FEs	Yes	Yes	Yes
Overall impact	-11.2	-23.9	12.7
Standard error	10.6	9.0	6.2

Notes: Corresponds to DGKP, Table IX, p.491. Standard errors clustered by industry. All regressions include a constant, firm-product fixed effects and sector×year fixed effects. The last two rows use the average decline in output tariffs of 68pp and the averge decline of 24pp in input tariffs to compute the mean and standard error of the overall impact on prices. R^2 is the Within R-squared. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Since log prices are the sum of log marginal costs and log markups, running (A-4) with first $\ln mc_{fjt}$ and then $\ln \mu_{fjt}$ as the left hand side variable, provides estimates of coefficients that sum to the change in prices, allowing a deeper understanding of the change in prices.

Column (2) shows results for marginal costs: a change in output tariffs has no significant effect, while a 10pp fall in input tariffs is associated with a 7.86% fall in marginal costs. Over the course of the liberalisation, the fall in both tariffs reduced firms' marginal costs by 23.9% – far more than 11.2% reduction in factory-gate prices. Column (3) shows how the effects of price and marginal cost changes associated with tariff changes led to changes in markups. The fall in output tariffs, leading to lower prices but largely unchanged marginal costs weakly acts to reduce markups. However, it is the fall in input tariffs that sharply reduces marginal costs without much affecting pricing that strongly acts to raise markups. The combined impact of the tariff changes is to raise markups by an estimated 12.7%.

These results are again similar to those found by DGKP. They report a somewhat larger effect of input tariffs on marginal cost (and consequently, a larger overall effect of trade reform of -30.7%), but similar effects on markups (overall effect of 12.6% although less-precisely estimated with a standard error of 11.9).

The fourth key result presented by DGKP is to disentangle the direct pro-competitive effect of output tariff liberalisation on markups from any sort of X-efficiencies made by firms that change costs and lead to markup adjustment. To do so, DGKP re-run the markup regression while controlling for marginal costs, (A-5)

$$\mu_{fjt} = \alpha_{fj} + \alpha_{st} + \beta_3 \tau_{it}^{out} + \beta_{mc} \ln mc_{fjt} + \beta_{mc^2} (\ln mc_{fjt})^2 + \epsilon_{fjt} \tag{A-5}$$

Results are shown in Table A.5, with the coefficients on marginal cost suppressed. Column (1) shows that a firm experiencing a 10pp fall in output tariffs will reduce markups by 1.7% conditional on any impact of trade reform on marginal costs. Column (2) instruments marginal costs to deal with measurement error as above. These coefficients are somewhat higher and more precise than those reported by DGKP (with OLS, a β_3 of 0.143 with a standard error of 0.05 and with the IV a β_3 of 0.150 with a standard error of 0.062), but qualitatively similar. Both show that output tariff reductions served to moderate the rise in markups that were due firms being able to resist cutting prices despite trade lowering their marginal costs.

Table A.5: Pro-competitive effects of output tariffs on markups

	$\ln \mu$	$\ln \mu$
τ^{output}	0.141	0.180
	(0.020)***	(0.046)***
N	27,504	15,848
R^2	0.53	0.53
Controls for marginal cost	yes	yes
Firm-product FEs	yes	yes
Sector-year FEs	yes	yes
Instruments	no	yes
First-stage F-test		9.3

Notes: Corresponds to DGKP, Table X, p.494. Standard errors clustered by industry. All regressions include a constant, firm-product fixed effects and sector×year fixed effects. Column 1 controls for marginal costs with a 2nd-order polynomial of log marginal cost, column 2 instruments for the polynomial with current period input tariffs and a lagged 2nd-order polynomial of marginal costs. First stage F-stat is the F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. R^2 is the Within R-squared. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

A-2 Additional table

Table A.6: OLS regressions of firm-product characteristics on tariffs with market share interactions

	(1) $\ln quantity$	(2) $\ln price$	(3) $\ln mc$	(4) $\ln \mu$	(5) a	(6) λ/μ	(7) $1/\mu$
τ^{output}	-0.005	0.209	0.001	0.209	-0.033	7.807	-2.301
	(0.082)	(0.104)**	(0.059)	(0.110)*	(0.068)	(4.614)*	(1.520)
$\tau^{output} * top$	0.049	-0.064	0.135	-0.199	-0.124	-6.780	2.230
	(0.077)	(0.109)	(0.060)**	(0.142)	(0.053)**	(5.701)	(1.918)
τ^{input}	-1.297	-0.150	1.176	-1.326	-1.003	-43.985	15.227
	(0.395)***	(0.662)	(0.424)***	(0.481)***	(0.410)**	(21.613)**	(7.886)*
$\tau^{input} * top$	0.821	0.230	-0.829	1.059	0.798	22.718	-7.961
	(0.268)***	(0.360)	(0.216)***	(0.474)**	(0.205)***	(18.338)	(6.348)
R^2	0.07	0.03	0.02	0.02	0.03	0.02	0.02
N	27,504	27,504	27,504	27,504	27,504	27,504	27,504
Firm-product FEs	yes	yes	yes	yes	yes	yes	yes
Sector-year FEs	yes	yes	yes	yes	yes	yes	yes

Notes: Table provides regression coefficients used to construct Table 2.11. mc is marginal cost, μ is markup, a is TFP, λ/μ is the demand shifter and $1/\mu$ is the demand slope. top is a time-invariant dummy variable that takes the value 1 if on the first appearance of a firm-product in the data it has an above median market share within a 4-digit industry, and 0 otherwise. All regressions include firm-product fixed effects and sector×year dummies. R^2 is the Within R-squared. Significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix B

Appendix to Chapter 3: On The Productivity Advantage of Cities

B-1 Closing the model: the DGKP and FMMM estimation procedures

B-1.1 Closing the model

We index firms by i and time by t. In what follows we consider a Cobb-Douglas production technology with 3 production factors: labour (L), materials (M) and capital (K). In line with the existing literature we assume capital to be a dynamic input that is predetermined in the short-run, i.e., current capital has been chosen in the past and cannot immediately adjust to current period shocks.[I] We further assume, as standard in the literature, that materials are a variable input free of adjustment costs. Concerning labor we could assume it is a variable input free of adjustment costs, or we could assume it is, very much like capital, predetermined in the short-run as in DGKP, or we could also assume, following Ackerberg et al. (2015), it is a semi-flexible input.[II] In light of the features of the French labor market we opt for the predetermined case.

[I]As described in Ackerberg et al. (2015) capital is often assumed to be a dynamic input subject to an investment process with the period t capital stock of the firm actually determined at period t-1. Intuitively, the restriction behind this assumption is that it takes a full period for new capital to be ordered, delivered, and installed.

[II]More precisely, in the semi-flexible case L_{it} is chosen by firm i at time $t-b$ ($0 < b < 1$), after K_{it} being chosen at $t-1$ but prior to M_{it} being chosen at t. In this case, one should expect L_{it} to be correlated with productivity shocks in t. Yet labour would not adjust fully to such shocks as materials do. The choice between predetermined and semi-flexible for L_{it} does not change the structure of the model and estimation procedure we provide below but only affects the set of moments used in the estimation. We highlight any differences later on.

We further assume firms are single-product, while relaxing this assumption in Appendix C, and minimize costs while taking the price of materials W_{Mit}, which is allowed to be firm-time specific, as given. Consequently, at any given point in time, each firm i is dealing with the following short-run cost minimization problem:[III]

$$\min_{M_{it}} \{M_{it} W_{Mit}\} \text{ s.t. } Q_{it} = A_{it} L_{it}^{\alpha_L} M_{it}^{\alpha_M} K_{it}^{\alpha_K},$$

where A_{it} is quantity TFP which is observable to the firm (and influences her choices) but not to the econometrician. In what follows we refer to the Cobb-Douglas production technology as the quantity equation and denote with lower case the log of a variable (for example a_{it} denotes the natural logarithm of A_{it}). The quantity equation can thus be written as:

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + a_{it}. \tag{B-1}$$

First order conditions to the firm's cost minimization problem imply that:

$$W_{Mit} = \chi_{it} \frac{Q_{it}}{M_{it}} \alpha_M \tag{B-2}$$

where χ_{it} is a Lagrange multiplier.[IV]

We can thus write the short-run cost function as:

$$C_{it} = M_{it} W_{Mit} = \chi_{it} Q_{it} \alpha_M = W_{Mit} \left(\frac{Q_{it}}{A_{it}}\right)^{\frac{1}{\alpha_M}} L_{it}^{-\frac{\alpha_L}{\alpha_M}} K_{it}^{-\frac{\alpha_K}{\alpha_M}}. \tag{B-3}$$

Marginal cost thus satisfies the following property:

$$\frac{\partial C_{it}}{\partial Q_{it}} = \frac{1}{\alpha_M} \frac{C_{it}}{Q_{it}}. \tag{B-4}$$

By combining equations (B-2), (B-3) and (B-4) one obtains the result provided in Section 3.3.2 that the markup can be computed as the ratio of the output elasticity of

[III] To simplify notation we ignore components that are constant across firms in a given time period as they will be controlled for by suitable dummies.

[IV] $\chi_{it} = \frac{W_{Mit}}{\alpha_M} Q_{it}^{\frac{1}{\alpha_M}-1} A_{it}^{-\frac{1}{\alpha_M}} L_{it}^{-\frac{\alpha_L}{\alpha_M}} K_{it}^{-\frac{\alpha_K}{\alpha_M}}.$

material to the share of materials' expenditure in revenue:

$$\mu_{it} = \frac{\alpha_M}{s_{Mit}}. \tag{B-5}$$

Moving to the time process of quantity TFP a_{it} we assume, as standard, it can be characterized by a Markov process and in particular we consider the leading AR(1) case. More specifically we assume:

$$a_{it} = \phi_a a_{it-1} + G_{ar} + \nu_{ait}, \tag{B-6}$$

where G_{ar} represents geographical factors affecting productivity (like the density of economic activities), and ν_{ait} stands for productivity shocks that are iid and represent innovations with respect to the information set of the firm in $t-1$.

B-1.2 The DGKP estimation procedure

From the above equations, the optimal expenditure on materials (B-3) is a function of labour, capital, the unit cost of materials and quantity TFP (which are known and given to the firm in t) as well as of the optimal quantity produced. The latter is obtained by equalizing the marginal cost and the marginal revenue and will thus depend upon the same 4 variables (labour, capital, the unit cost of materials and quantity TFP) plus factors characterizing the specific demand facing firm i. DGKP suggest to proxy for the unobservable unit cost of materials and firm-specific demand factors with the observable price and marker share of firm i as well as regional variables G_r. Operationally, they thus assume the conditional (log) input demand for materials can be expressed as a function $h(.)$ of k_{it}, l_{it}, a_{it}, p_{it}, G_r, and the market share MS_{it}. If $h(.)$ is globally invertible with respect to a_{it}, the inverse function $a_{it} = g(k_{it}, l_{it}, m_{it}, p_{it}, MS_{it}, G_r)$ exists and is well behaved and so one can use a semi-parametric polynomial approximation of $g(.)$ in order to proxy for the unobservable (to the econometrician) quantity TFP a_{it}. Operationally, we use a second order polynomial in the arguments of $g(.)$ to proxy this function. By labeling this polynomial $Poly_{it}$ we thus have $a_{it} = Poly_{it}$.

Using $a_{it-1} = Poly_{it-1}$ in (B-6) we have:

$$a_{it} = \phi_a Poly_{it-1} + G_{ar} + \nu_{ait},$$

while substituting this into the production function one gets:

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \phi_a Poly_{it-1} + G_{ar} + \nu_{ait}. \tag{B-7}$$

Note that in (B-7) one does not need to identify the parameter ϕ_a nor separately identify G_{ar} form the G_r contained in $Poly_{it-1}$. Therefore, one can write:

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + Poly^{'}_{it-1} + \nu_{ait}. \tag{B-8}$$

where $Poly^{'}_{it-1}$ is simply a second order polynomial in k_{it-1}, l_{it-1}, m_{it-1}, p_{it-1}, G_r and MS_{it-1}.[V] Given the assumption that productivity shocks ν_{ait} are innovations with respect to the information set of the firm in $t-1$, ν_{ait} is uncorrelated with $Poly^{'}_{it-1}$ in (B-8). Furthermore, labour and capital are predetermined and so uncorrelated with ν_{ait} too. Therefore, the only endogenous variable in (B-8) is materials m_{it} and parameters can be estimated by exploiting additional moments conditions. More specifically, we use materials, labour and capital at time $t-2$ as instruments for materials in t. This ultimately allows us to get estimates of the production function parameters $\hat{\alpha}_L$, $\hat{\alpha}_M$ and $\hat{\alpha}_K$ as well as productivity $\hat{a}_{it}$=$q_{it} - \hat{\alpha}_L l_{it} - \hat{\alpha}_M m_{it} - \hat{\alpha}_K k_{it}$. (3.6) and (3.9) can then be used to recover the firm-specific markup and marginal cost while (3.2) delivers demand heterogeneity $\tilde{\lambda}_{it}$. We perform estimations of (B-8) separately for each two-digit industry (NACE Sections) and consider a full battery of 8-digit product dummies, as well as year dummies.

B-1.3 The FMMM estimation procedure

As in FMMM we assume that product appeal follows an AR(1) process and in particular:

$$\lambda_{it} = \phi_\lambda \lambda_{it-1} + G_{\lambda r} + \nu_{\lambda it}, \tag{B-9}$$

where $G_{\lambda r}$ represents geographical factors affecting demand (like the density of economic activities) and $\nu_{\lambda it}$ stands for product appeal shocks that are iid and represent innovations

[V] There is, however, an identification issue with both (B-7) and (B-8) in the DGKP procedure which does not apply to the FMMM procedure. Firm market share and price in $t-1$ are, among other things, present as covariates in (B-7) and (B-8) where quantity at time t in on the left-hand side. In this respect using the market share of firm i as a proxy is equivalent to using the revenue of firm i as a proxy. Indeed, market share is firm revenue divided by industry-level sales and the denominator can be basically considered a constant across firms when in the regression there is a set of industry or product dummies. Lagged revenue and price obviously perfectly predict lagged quantity which is a powerful predictor of current quantity. Therefore, there might be little variation left to precisely identify technology parameters. In our analyses we actually encountered such problems and ultimately decided to drop firm market share from the polynomial approximation.

with respect to the information set of the firm in $t-1$. Furthermore, we make use of the result that the log revenue function can be approximated (up to a constant across firms that will be controlled for by using suitable dummies) by a linear function of quantity and product appeal and, to avoid burdening notation, we use $=$ instead of $\simeq$:

$$r_{it} = \frac{1}{\mu_{it}}(q_{it} + \lambda_{it}). \tag{B-10}$$

We label (B-10) the revenue equation.

This estimation procedure builds upon (B-10) and uses both the revenue and quantity equations to estimate technology parameters. The two-steps procedure described below is not the only one that can be used to recover technology parameters under our set of assumptions but has the advantage of being simple to implement and linear. In what follows, it is convenient to rewrite the Cobb-Douglas production function as:

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + (\gamma - \alpha_L - \alpha_M)k_{it} + a_{it}, \tag{B-11}$$

where γ characterizes returns to scale. By substituting q_{it} with the formula of the Cobb-Douglas we can transform (B-10) further as:

$$r_{it} = \frac{\alpha_L}{\mu_{it}}\left(l_{it} - k_{it}\right) + \frac{\alpha_M}{\mu_{it}}\left(m_{it} - k_{it}\right) + \frac{\gamma}{\mu_{it}}k_{it} + \frac{1}{\mu_{it}}\left(a_{it} + \lambda_{it}\right).$$

Furthermore, by using (B-5), we get:

$$LHS_{it} \equiv \frac{r_{it} - s_{Mit}\left(m_{it} - k_{it}\right)}{s_{Mit}} = \frac{\alpha_L}{\alpha_M}\left(l_{it} - k_{it}\right) + \frac{\gamma}{\alpha_M}k_{it} + \frac{1}{\alpha_M}\left(a_{it} + \lambda_{it}\right). \tag{B-12}$$

where LHS_{it} is made out of observables only.

We then build upon our assumptions on the time process for a_{it} and λ_{it}: (B-6) and (B-9). However, before substituting (B-6) and (B-9) into (B-12) we need to find a convenient way to express a_{it-1} and λ_{it-1}. By using (B-5) and (B-10) we have:

$$\lambda_{it-1} = r_{it-1}\mu_{it-1} - q_{it-1} = r_{it-1}\frac{\alpha_M}{s_{Mit-1}} - q_{it-1}. \tag{B-13}$$

At the same time plugging (B-13) into (B-12) and re-arranging yields:

$$a_{it-1} = \alpha_M LHS_{it-1} - \alpha_L\left(l_{it-1} - k_{it-1}\right) - \gamma k_{it-1} - \left(r_{it-1}\frac{\alpha_M}{s_{Mit-1}} - q_{it-1}\right). \tag{B-14}$$

Finally, by combining (B-6), (B-9), (B-13) and (B-14) into (B-12) we obtain:

$$LHS_{it} = \frac{\gamma}{\alpha_M}k_{it} + \frac{\alpha_L}{\alpha_M}(l_{it} - k_{it}) + \phi_a LHS_{it-1} - \phi_a \frac{\gamma}{\alpha_M}k_{it-1} - \phi_a \frac{\alpha_L}{\alpha_M}(l_{it-1} - k_{it-1})$$

$$+ (\phi_\lambda - \phi_a)\left(\frac{r_{it-1}}{s_{Mit-1}} - \frac{q_{it-1}}{\alpha_M}\right) + \frac{1}{\alpha_M}(G_{ar} + G_{\lambda r}) + \frac{1}{\alpha_M}(\nu_{ait} + \nu_{\lambda it}). \quad \text{(B-15)}$$

Note that the revenue equation (B-15) is, besides the idiosyncratic productivity and demand shocks ν_{ait} and $\nu_{\lambda it}$, now entirely written in terms of observables and useful parameters. There are various ways of estimating (B-15) and here we use perhaps the simplest one. More specifically, we rewrite (B-15) as the following linear regression:

$$LHS_{it} = b_1 z_{1it} + b_2 z_{2it} + b_3 z_{3it} + b_4 z_{4it} + b_5 z_{5it} + b_6 z_{6it} + b_7 z_{7it} + u_r + u_{it}, \quad \text{(B-16)}$$

where z_{1it}=k_{it}, z_{2it}=$(l_{it} - k_{it})$, z_{3it}=LHS_{it-1}, z_{4it}=k_{it-1}, z_{5it}=$(l_{it-1} - k_{it-1})$, z_{6it}=$\frac{r_{it-1}}{s_{Mit-1}}$, z_{7it}=q_{it-1}, u_r=$\frac{1}{\alpha_M}(G_{ar} + G_{\lambda r})$, u_{it}=$\frac{1}{\alpha_M}(\nu_{ait} + \nu_{\lambda it})$ as well as b_1=$\frac{\gamma}{\alpha_M}$, b_2=$\frac{\alpha_L}{\alpha_M}$, b_3=ϕ_a, b_4=$-\phi_a \frac{\gamma}{\alpha_M}$, b_5=$-\phi_a \frac{\alpha_L}{\alpha_M}$, b_6=$(\phi_\lambda - \phi_a)$ and b_7=$-(\phi_\lambda - \phi_a)\frac{1}{\alpha_M}$.

Given our assumptions, the error term u_{it} in (B-16) is uncorrelated with current capital and labour as well as with lagged inputs use, quantity and revenue. Therefore, z_{1it} to z_{7it} are uncorrelated to u_{it}. Concerning u_r, it is also uncorrelated with u_{it} and so (B-16) can be estimated by OLS. Indeed, given we do not need to separately identify the impact of geographical factors affecting productivity and demand but simply control for them, we simply replace u_r with the log of the 2009 population and the log of the land area of region r and use them as controls in (B-16). Operationally, we augment (B-16) with a full battery of 8-digit product dummies, as well as year dummies, and finally set $\widehat{\frac{\gamma}{\alpha_M}}$=$\hat{b}_1$, $\widehat{\frac{\alpha_L}{\alpha_M}}$=$\hat{b}_2$ and $\hat{\phi}_a$=$\hat{b}_3$ without exploiting parameters' constraints in the estimation.

We now turn to estimating γ from the quantity equation in a second step. Combining (B-1) and (B-5) we have:

$$q_{it} = \mu_{it} s_{Mit}(m_{it} - k_{it}) + \alpha_L(l_{it} - k_{it}) + \gamma k_{it} + a_{it}. \quad \text{(B-17)}$$

Further using $\alpha_M = \frac{\gamma}{\hat{b}_1}$ as well as $\alpha_L = \frac{\gamma \hat{b}_2}{\hat{b}_1}$ and we get:

$$q_{it} = \frac{\gamma}{\hat{b}_1}(m_{it} - k_{it}) + \frac{\gamma \hat{b}_2}{\hat{b}_1}(l_{it} - k_{it}) + \gamma k_{it} + a_{it}, \quad \text{(B-18)}$$

where we replace b_1 and b_2 with their estimates $\hat{b}_1$ and $\hat{b}_2$ coming from (B-16). Finally, using (B-6) to substitute for a_{it} and using (B-14) we obtain:

$$\begin{aligned} q_{it} &= \frac{\gamma}{\hat{b}_1}(m_{it}-k_{it})+\frac{\gamma\hat{b}_2}{\hat{b}_1}(l_{it}-k_{it})+\gamma k_{it}+\gamma\frac{\hat{\phi}_a}{\hat{b}_1}LHS_{it-1}-\frac{\gamma\hat{b}_2\hat{\phi}_a}{\hat{b}_1}(l_{it-1}-k_{it-1}) \\ &- \gamma\hat{\phi}_a k_{it-1}-\hat{\phi}_a\left(r_{it-1}\frac{\gamma}{\hat{b}_1 s_{Mit-1}}-q_{it-1}\right)+G_{ar}+\nu_{ait}. \end{aligned} \tag{B-19}$$

Note that the only unobservable in (B-19) is the idiosyncratic productivity shock ν_{ait} while the only important parameter left to identify is γ. Indeed, the impact of geographical factors affecting productivity is simply a control in our framework and we use the log of the 2009 population and the log of the land area of region r to replace G_{ar}. We can more compactly write (B-19) as the following linear regression:

$$\overline{LHS}_{it}=b_8 z_{8it}+G_{ar}+\nu_{ait} \tag{B-20}$$

where:

$$\begin{aligned} \overline{LHS}_{it} &= q_{it}-\hat{\phi}_a q_{it-1} \\ z_{8it} &= \frac{1}{\hat{b}_1}(m_{it}-k_{it})+\frac{\hat{b}_2}{\hat{b}_1}(l_{it}-k_{it})+k_{it}+\frac{\hat{\phi}_a}{\hat{b}_1}LHS_{it-1} \\ &- \frac{\hat{b}_2\hat{\phi}_a}{\hat{b}_1}(l_{it-1}-k_{it-1})-\hat{\phi}_a k_{it-1}-\frac{\hat{\phi}_a r_{it-1}}{\hat{b}_1 s_{Mit-1}} \end{aligned}$$

as well as b_8=γ. Concerning z_{8it}, it is endogenous but we can use and have used several moment conditions for identification: $E\{\nu_{ait}k_{it}\}=E\{\nu_{ait}l_{it}\}=E\{\nu_{ait}l_{it-1}\}=E\{\nu_{ait}m_{it-1}\}=E\{\nu_{ait}k_{it-1}\}=E\{\nu_{ait}q_{it-1}\}=E\{\nu_{ait}r_{it-1}\}=0$. As for G_{ar}, it is under our assumptions uncorrelated with ν_{ait}. Operationally, we perform estimations of (B-20) separately for each two-digit industry (NACE Sections) and consider a full battery of 8-digit product dummies, as well as year dummies. IV estimation of (B-20) provides an estimate of γ that, together with $\widehat{\frac{\gamma}{\alpha_M}}$ and $\widehat{\frac{\alpha_L}{\alpha_M}}$ coming from the first stage revenue equation, uniquely delivers production function parameters ($\hat{\alpha}_L$, $\hat{\alpha}_M$ and $\hat{\gamma}$) as well as productivity $\hat{a}_{it}$=$q_{it}-\hat{\alpha}_L l_{it}-\hat{\alpha}_M m_{it}-(\hat{\gamma}-\hat{\alpha}_L-\hat{\alpha}_M)k_{it}$. (3.6) and (3.9) can then be used to recover the firm-specific markup and marginal cost while (3.2) delivers demand heterogeneity $\tilde{\lambda}_{it}$.

B-2 Measurement error in output and unanticipated shocks

As customary in productivity analyses, an issue to account for before proceeding to any estimations of the production function is the presence of measurement error in output and/or unanticipated productivity shocks. In the former case, instead of q_{it}, the econometrician might be observing q'_{it}=$q_{it} + e_{it}$ where e_{it} is standard measurement error. Another interpretation of the same equation is that e_{it} represents productivity shocks unanticipated by the firm. (B-1) thus becomes:

$$q'_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + a_{it} + e_{it}.$$

The approach suggested by the literature (Ackerberg et al., 2015; De Loecker et al., 2016) to deal with measurement error in output and/or unanticipated shocks e_{it} is based on the proxy variable framework and a semi-parametric implementation. We follow this approach and, building on the same logic of equation (19) in DGKP, we estimate:

$$q'_{it} = poly(l_{it}, m_{it}, p_{it}, k_{it}) + e_{it}, \tag{B-21}$$

where q'_{it} is (log) quantity as reported in the data and $poly(.)$ is a third-order polynomial in l_{it}, m_{it}, p_{it} and k_{it}.[VI] We run (B-21) separately for each two-digit industry while including the log of the 2009 population and the log of the land area of region r, as well as a full set of 8-digit product dummies and year dummies, to (B-21). We then use the OLS prediction of q'_{it}, that we label $\hat{q'}^{OLS}_{it}$, as quantity in the both the DGKP and FMMM procedures.

We also use the same approach for revenue and consider:

$$r'_{it} = poly(l_{it}, m_{it}, p_{it}, k_{it}) + \bar{e}_{it}, \tag{B-22}$$

where $\bar{e}_{it}$ now contains measurement error in both quantity and prices, as well as unobserved productivity shocks, and use the OLS prediction of r'_{it}, that we label $\hat{r'}^{OLS}_{it}$, as revenue in the both the DGKP and FMMM procedures. Again, we run (B-22) separately

[VI]The logic behind using (B-21) to purge quantity from measurement error and unanticipated shocks is quite simple. From the quantity equation (B-1) q_{it} is a function of l_{it}, m_{it}, k_{it} and a_{it}. Using prices p_{it} as a proxy for a_{it}, while assuming invertibility, one can then write a_{it} as a function of l_{it}, m_{it}, p_{it} and k_{it}. Overall, q_{it} is thus a function of l_{it}, m_{it}, p_{it} and k_{it} than can be semi-parametrically approximated by a polynomial function. Crucially, measurement error and/or unanticipated shocks to do influence a firm's choices and so their are not part of the polynomial approximation but rather the residual of equation (B-21).

for each two-digit industry while including the log of the 2009 population and the log of the land area of region r, as well as a full set of 8-digit product dummies and year dummies, to (B-22). Also note that, as suggested in DGKP, by purging revenue from measurement error and using $\hat{r'}_{it}^{OLS}$ instead of r'_{it}, we obtain a more reliable measure of the share of materials in revenue (s_{Mit}) that is needed to compute markups.

B-3 Multi-product firms

Produced quantities and generated revenues may be observable for the different products of each firm in databases like ours. However, information on inputs used for a specific product is typically not available. We report here an extension of the MULAMA model from FMMM to solve the problem of assigning inputs to outputs for multi-product firms.

As usual we denote a firm by i and time by t. A firm i produces in t one or more products indexed by p and the number of products produced by the firm is denoted by I_{it}. In our data p is an 8-digit prodcom product code but in other data, like the bar-code data used in Hottman et al. (2016), can be much more detailed. We assume product appeal is firm-time specific (λ_{it}) while we allow markups (μ_{ipt}) and productivity (a_{ipt}) to be firm-product-time specific. The production function for product p produced by firm i is given by:

$$Q_{ipt} = C_p C_t A_{ipt} L_{ipt}^{\alpha_{Lg}} M_{ipt}^{\alpha_{Mg}} K_{ipt}^{\gamma_g - \alpha_{Mg} - \alpha_{Lg}}, \tag{B-23}$$

where C_p and C_t are innocuous product and time constants (that will be controlled for by suitable dummies) we disregard in what follows and g identifies a product group/industry. Production function coefficients are the same for products within a product group because a certain level of data aggregation is needed to deliver enough observations to estimate parameters. (B-23) means we allow for technology (α_{Lg}, α_{Mg}, γ_g) to differ across the different products p produced by a multi-product firm. At the same time productivity is allowed to vary across products within a firm and information coming from single-product firms need to be used to infer the technology of multi-product firms, i.e., we rule out physical synergies in production but allow for some of the economies (diseconomies) of scope discussed in DGKP. Furthermore, we assume firm i to maximize profits and choose (for each product p) the amount of labour L_{ipt} and materials M_{ipt} in order to minimize short-term costs while taking capital K_{ipt}, as well as productivity a_{ipt} and product appeal

λ_{it} as given. We make use of (3.4) and in particular:

$$r_{ipt} \simeq \frac{1}{\mu_{ipt}}(q_{ipt} + \lambda_{it}). \tag{B-24}$$

Profit maximization implies:

$$P_{ipt} = \mu_{ipt}\frac{\partial C_{ipt}}{\partial Q_{ipt}}, \tag{B-25}$$

so that we can, starting from data on prices and markups, recover marginal costs. Also note that the marginal cost is equal to[VII]

$$\frac{\partial C_{ipt}}{\partial Q_{ipt}} = A_{ipt}^{-\frac{1}{\alpha_{Lg}+\alpha_{Mg}}} Q_{ipt}^{\frac{1-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}} K_{ipt}^{\frac{\gamma_g-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}} \tag{B-26}$$

Firms minimize costs and so markups are such that:

$$\mu_{ipt} = \frac{\alpha_{Mg}}{s_{Mipt}} \tag{B-27}$$

where s_{Mipt} is the expenditure share of materials for product p at time t in firm revenue for product p at time t.

As far as single-product firms are concerned, the DGKP procedure or the FMMM procedure described in Appendix A can be used to recover quantity TFP, markups, marginal costs and demand heterogeneity. Turning to multi-product firms we impose, as in DGKP, that the same technology parameters coming from single-product producers extend to the products of the former. Yet, in order to quantify multi-product firms productivity, markups, marginal costs and demand heterogeneity we still need to solve the issue of how to assign inputs to outputs and we do so by building on the above assumptions and the parameters estimated for single-product firms. As far as materials are concerned, we need to assign the observable total firm material expenditure M_{it} across the I_{it} products produced by firm i at time t, i.e., we need to assign values to M_{ipt} such that $\sum_{p=1}^{I_{it}} M_{ipt} = M_{it}$. We can use this condition along with (B-27) and (B-24) to operate this assignment. Substituting (B-27) into (B-24) and adding $\sum_{p=1}^{I_{it}} M_{ipt} = M_{it}$ provides a system of $I_{it}+1$ equations in $I_{it}+1$ unknowns; the I_{it} inputs expenditures M_{ipt} plus λ_{it}. Indeed, at this stage we have data on r_{ipt}, q_{ipt}, α_{Mg} and M_{it}. Operationally, one can actually proceed in two stages. Combining the above equations one has $\sum_{p=1}^{I_{it}} \frac{\alpha_{Mg} r_{ipt} R_{ipt}}{q_{ipt}+\lambda_{it}} = M_{it}$. This

[VII] We omit the innocuous product-time constant $\left(\frac{W_{Lpt}}{\alpha_{Lg}}\right)^{\frac{\alpha_{Lg}}{\alpha_{Lg}+\alpha_{Mg}}} \left(\frac{W_{Mpt}}{\alpha_{Mg}}\right)^{\frac{\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}}$

equation is solved for each firm and delivers λ_{it}. With this at hand one can then obtain materials expenditure from $M_{ipt} = \frac{\alpha_{Mg} r_{ipt} R_{ipt}}{q_{ipt}+\lambda_{it}}$. By recovering inputs expenditures M_{ipt} we subsequently compute materials expenditure shares in revenues s_{Mipt} and so use (B-27) to recover a firm-product-time specific markup μ_{ipt} as well as the marginal cost from (B-25). Since labour is a variable input a condition analogous to (B-27) holds for this input and so we use the computed markups μ_{ipt} and information on α_{Lg} to derive labour expenditure: $L_{ipt} = \frac{\alpha_{Lg} R_{ipt}}{\mu_{ipt}}$. Operationally, this is not guaranteed to satisfy the constraint $\sum_{p=1}^{I_{it}} L_{ipt} = L_{it}$ for each firm and so the L_{ipt} are re-scaled for each firm.

The above procedure allows so far to obtain markups, marginal costs and product appeal/demand heterogeneity, as well as information on labour and materials use, for each of the products of a multi-product firm. However, in order to recover productivity a_{ipt} we still need values for capital K_{ipt}. To do this one can proceed as follows. Combining the marginal cost, profit maximization and quantity equations one gets:

$$K_{ipt} = \left(\frac{P_{ipt}}{\mu_{ipt} Q_{ipt}^{a+b} L_{ipt}^{-a\alpha_{Lg}} M_{ipt}^{-a\alpha_{Mg}}} \right)^{\left(\frac{1}{c-a\alpha_{Kg}}\right)} \tag{B-28}$$

where $a = -\frac{1}{\alpha_{Lg}+\alpha_{Mg}}$, $b = \frac{1-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}$, $c = \frac{\gamma_g-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}$ and $\alpha_{Kg} = \gamma_g - \alpha_{Mg} - \alpha_{Lg}$ is the capital coefficient. We further refine those values by running an estimation where the computed K_{ipt} from (B-28) is regressed on R_{ipt}, M_{ipt}, L_{ipt} as well as total firm expenditure on materials and labour plus the capital stock and a full battery of year and product dummies. The predicted values of such regression are then re-scaled for each firm to meet the constraint $\sum_{p=1}^{I_{it}} K_{ipt} = K_{it}$.

B-4 Additional Tables

Table B.1: Revenue productivity, density and product composition effects (number of firms weighted, SP sample)

Dep. var.	log VA per worker		OLS TFP-R		Wooldridge TFP-R		TFP-R	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log density	0.0432	0.0328	0.0136	0.0108	0.0210	0.0175	0.0210	0.0163
	(0.0073)***	(0.0077)***	(0.0021)***	(0.0019)***	(0.0052)***	(0.0037)***	(0.0050)***	(0.0033)***
R^2	0.02	0.17	0.56	0.69	0.87	0.91	0.78	0.87
N	55,432	55,432	55,432	55,432	55,432	55,432	55,432	55,432
2-digit dummies	Yes	No	Yes	No	Yes	No	Yes	No
8-digit dummies	No	Yes	No	Yes	No	Yes	No	Yes

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by ZE. Regressions are weighted and include year dummies as well as either 2-digit or 8-digit product dummies. Estimations are carried on the sample of SP firms. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.2: Revenue productivity, density and product composition effects (revenue weighted, SP sample)

Dep. var.	log VA per worker		OLS TFP-R		Wooldridge TFP-R		TFP-R	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log density	0.0899	0.0621	0.0189	0.0120	0.0525	0.0298	0.0515	0.0300
	(0.0147)***	(0.0134)***	(0.0027)***	(0.0022)***	(0.0116)***	(0.0078)***	(0.0107)***	(0.0074)***
R^2	0.04	0.33	0.57	0.80	0.90	0.96	0.86	0.95
N	55,432	55,432	55,432	55,432	55,432	55,432	55,432	55,432
2-digit dummies	Yes	No	Yes	No	Yes	No	Yes	No
8-digit dummies	No	Yes	No	Yes	No	Yes	No	Yes

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by ZE. Regressions are weighted and include year dummies as well as either 2-digit or 8-digit product dummies. Estimations are carried on the sample of SP firms. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.3: OLS regressions of standard revenue productivity measures on ZE population density (revenue weighted, 2-digit dummies, various samples)

Dep. var.	log VA per worker				OLS TFP-R				Wooldridge TFP-R			
	Fare sample	Prodcom sample	SP+MP sample	SP sample	Fare sample	Prodcom sample	SP+MP sample	SP sample	Fare sample	Prodcom sample	SP+MP sample	SP sample
log density	0.0662	0.0759	0.0762	0.0899	0.0162	0.0170	0.0171	0.0189	0.0632	0.0554	0.0755	0.0525
	(0.0119)***	(0.0134)***	(0.0137)***	(0.0147)***	(0.0024)***	(0.0026)***	(0.0027)***	(0.0027)***	(0.0168)***	(0.0137)***	(0.0150)***	(0.0116)***
2-digit dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.12	0.12	0.13	0.04	0.81	0.83	0.77	0.57	0.80	0.93	0.88	0.90
N	628,940	201,261	189,017	55,432	628,940	201,261	189,017	55,432	628,940	201,261	189,017	55,432

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the ZE level. Regressions include time and industry (2-digit) dummies. The Fare sample includes firms with complete balance sheet data in NACE 2 industries 10-32 that remain after an initial cleaning of the data. The Prodcom sample includes the subset of such firms that are in the Prodcom dataset. In both samples, an observation is a firm-year combination. Each firm-year observation is weighted by R_{it}/R_r where R_{it} is firm i revenue at time t and R_r is the sum of R_{it} across the firm-year observations corresponding to the ZE r. SP and MP refer to single-product and multi-product firms in the Prodcom sample that have been subject to further data cleaning. We consider two samples: 1) the sample of SP and MP; 2) the sample of SP. In both samples an observation is a firm-product-year combination. For SP a firm-product-year combination corresponds to a unique firm-year combination. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.4: OLS regressions of standard revenue productivity measures on ZE population density (revenue weighted, 6-digit dummies, various samples)

Dep. var.	log VA per worker				OLS TFP-R				Wooldridge TFP-R			
	Fare sample	Prodcom sample	SP+MP sample	SP sample	Fare sample	Prodcom sample	SP+MP sample	SP sample	Fare sample	Prodcom sample	SP+MP sample	SP sample
log density	0.0565	0.0611	0.0596	0.0783	0.0137	0.0142	0.0139	0.0159	0.0572	0.0382	0.0545	0.0385
	(0.0119)***	(0.0134)***	(0.0114)***	(0.0147)***	(0.0024)***	(0.0026)***	(0.0021)***	(0.0027)***	(0.0168)***	(0.0137)***	(0.0094)***	(0.0116)***
6-digit dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.26	0.26	0.27	0.16	0.84	0.85	0.81	0.66	0.89	0.94	0.91	0.92
N	628,940	201,261	189,017	55,432	628,940	201,261	189,017	55,432	628,940	201,261	189,017	55,432

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the ZE level. Regressions include time and industry (6-digit) dummies. The Fare sample includes firms with complete balance sheet data in NACE 2 industries 10-32 that remain after an initial cleaning of the data. The Prodcom sample includes the subset of such firms that are in the Prodcom dataset. In both samples, an observation is a firm-year combination. Each firm-year observation is weighted by R_{it}/R_r where R_{it} is firm i revenue at time t and R_r is the sum of R_{it} across the firm-year observations corresponding to the ZE r. SP and MP refer to single-product and multi-product firms in the Prodcom sample that have been subject to further data cleaning. We consider two samples: 1) the sample of SP and MP; 2) the sample of SP. In both samples an observation is a firm-product-year combination. For SP a firm-product-year combination corresponds to a unique firm-year combination. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.5: OLS regression of log marginal costs on TFP and log quantity (SP firms)

Dep. Var.	log marg. cost	log marg. cost
TFP	-1.1936	-1.2092
	(0.0027)***	(0.0061)***
log quantity	0.2157	0.2294
	(0.0020)***	(0.0039)***
Weighting	un-weighted	revenue
R^2	0.9974	0.9985
N	55,432	55,432

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered by firm. Regressions include year dummies as well as 8-digit product dummies. Estimations are carried on the sample of SP firms. The first column reports results of an un-weighted OLS regression while column two provides results of a weighted OLS regression where each firm-product-year observation is weighted by R_{ipt} where R_{ipt} is firm i revenue corresponding to product p at time t.

Table B.6: FMMM procedure: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.1032	0.1353	0.0321	0.0346	-0.0025
	(0.0546)*	(0.0489)***	(0.0188)*	(0.0188)*	(0.0031)
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.7: FMMM procedure: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (number of firms weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.0386	0.0549	0.0163	0.0225	-0.0062
	(0.0270)	(0.0243)**	(0.0081)**	(0.0081)***	(0.0028)**
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.8: FMMM procedure: 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\tilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0158	-0.0163	0.0321	0.0040	-0.0722	0.0840	0.0042
	(0.0091)*	(0.0222)	(0.0188)*	(0.0444)	(0.0639)	(0.0336)**	(0.0036)
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.9: FMMM procedure: 2SLS regressions of firm TFP-R and Mulama measures on log density (number of firms weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\widetilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0211	0.0048	0.0163	0.0276	-0.0387	0.0323	0.0074
	(0.0034)***	(0.0091)	(0.0081)**	(0.0207)	(0.0312)	(0.0139)**	(0.0030)**
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.10: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.1196	0.1975	0.0779	0.0864	-0.0085
	(0.0629)*	(0.0554)***	(0.0305)**	(0.0311)***	(0.0040)**
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.11: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (number of firms weighted, SP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.0518	0.0770	0.0252	0.0329	-0.0077
	(0.0340)	(0.0288)***	(0.0151)*	(0.0160)**	(0.0028)***
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.12: 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\widetilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0428	-0.0351	0.0779	0.0135	-0.1026	0.1319	0.0117
	(0.0122)***	(0.0322)	(0.0305)**	(0.0460)	(0.0713)	(0.0425)***	(0.0044)***
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.13: 2SLS regressions of firm TFP-R and Mulama measures on log density (number of firms weighted, SP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\widetilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0168	-0.0084	0.0252	0.0148	-0.0403	0.0423	0.0081
	(0.0042)***	(0.0162)	(0.0151)*	(0.0200)	(0.0281)	(0.0134)***	(0.0025)***
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.14: DGKP procedure with total wage bill to measure the labour input: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.1487	0.1899	0.0413	0.0473	-0.0060
	(0.0602)**	(0.0556)***	(0.0228)*	(0.0230)**	(0.0034)*
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.15: DGKP procedure with total wage bill to measure the labour input: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (number of firms weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.0524	0.0675	0.0151	0.0237	-0.0086
	(0.0271)*	(0.0245)***	(0.0079)*	(0.0081)***	(0.0029)***
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.16: DGKP procedure with total wage bill to measure the labour input: 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0064	-0.0349	0.0413	-0.0250	-0.1013	0.1328	0.0085
	(0.0059)	(0.0254)	(0.0228)*	(0.0361)	(0.0492)**	(0.0349)***	(0.0032)***
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.17: DGKP procedure with total wage bill to measure the labour input: 2SLS regressions of firm TFP-R and Mulama measures on log density (number of firms weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0035	-0.0116	0.0151	0.0076	-0.0465	0.0424	0.0082
	(0.0025)	(0.0087)	(0.0079)*	(0.0176)	(0.0226)**	(0.0122)***	(0.0025)***
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.18: DGKP procedure with wage bill weights: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.1120	0.1496	0.0376	0.0397	-0.0021
	(0.0556)**	(0.0507)***	(0.0223)*	(0.0229)*	(0.0039)
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by W_{ipt}/W_r where W_{ipt} is firm i wage bill corresponding to product p at time t and W_r is the sum of W_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.19: DGKP procedure with wage bill weights: 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\tilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0340	-0.0036	0.0376	0.0200	-0.0680	0.0820	0.0047
	(0.0138)**	(0.0260)	(0.0223)*	(0.0403)	(0.0616)	(0.0390)**	(0.0036)
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by W_{ipt}/W_r where W_{ipt} is firm i wage bill corresponding to product p at time t and W_r is the sum of W_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.20: DGKP procedure excluding firms located in Île de France: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (revenue weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.1263	0.1794	0.0530	0.0605	-0.0074
	(0.0821)	(0.0767)**	(0.0275)*	(0.0277)**	(0.0046)
N	257	257	257	257	257
LM stat under-identif.	42.9	42.9	42.9	42.9	42.9
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	202.2	202.2	202.2	202.2	202.2

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.21: DGKP procedure excluding firms located in Île de France: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (number of firms weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.0650	0.0830	0.0180	0.0282	-0.0101
	(0.0317)**	(0.0273)***	(0.0100)*	(0.0100)***	(0.0033)***
N	257	257	257	257	257
LM stat under-identif.	42.9	42.9	42.9	42.9	42.9
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	202.2	202.2	202.2	202.2	202.2

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.22: DGKP procedure excluding firms located in Île de France: 2SLS regressions of firm TFP-R and Mulama measures on log density (revenue weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0357	-0.0173	0.0530	0.0361	-0.1148	0.1144	0.0099
	(0.0199)*	(0.0342)	(0.0275)*	(0.0596)	(0.0875)	(0.0529)**	(0.0051)*
N	257	257	257	257	257	257	257
LM stat under-identif.	42.9	42.9	42.9	42.9	42.9	42.9	42.9
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	202.2	202.2	202.2	202.2	202.2	202.2	202.2

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.23: DGKP procedure excluding firms located in Île de France: 2SLS regressions of firm TFP-R and Mulama measures on log density (number of firms weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\tilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\tilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0170	-0.0010	0.0180	0.0331	-0.0514	0.0353	0.0094
	(0.0051)***	(0.0118)	(0.0100)*	(0.0189)*	(0.0273)*	(0.0107)***	(0.0028)***
N	257	257	257	257	257	257	257
LM stat under-identif.	42.9	42.9	42.9	42.9	42.9	42.9	42.9
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	202.2	202.2	202.2	202.2	202.2	202.2	202.2

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Table B.24: DGKP procedure with Translog production function: 2SLS regressions of firm log quantity, log revenue, log price, log marginal cost and log markup on log density (number of firms weighted, SP+MP sample)

Dep. var.	log quantity	log revenue	log price	log marg. cost	log markup
log density	0.0458	0.0624	0.0166	0.0225	-0.0059
	(0.0264)*	(0.0239)***	(0.0080)**	(0.0081)***	(0.0026)**
N	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by R_{ipt}/R_r where R_{ipt} is firm i revenue corresponding to product p at time t and R_r is the sum of R_{ipt} across the firm-product-year observations corresponding to the ZE r.

Table B.25: DGKP procedure with Translog production function: 2SLS regressions of firm TFP-R and Mulama measures on log density (number of firms weighted, SP+MP sample)

Dep. var.	TFP-R	TFP	log price	Adj. TFP $\widetilde{a}$	rev. shifter $\widetilde{\lambda}$	Adj. scale $\widetilde{\tilde{q}}$	rev. slope $1/\mu$
log density	0.0076	-0.0090	0.0166	-0.0075	0.0073	0.0078	0.0043
	(0.0026)***	(0.0082)	(0.0080)**	(0.0125)	(0.0202)	(0.0095)	(0.0020)**
N	273	273	273	273	273	273	273
LM stat under-identif.	29.4	29.4	29.4	29.4	29.4	29.4	29.4
Under-identif. p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald F stat weak identif.	225.9	225.9	225.9	225.9	225.9	225.9	225.9

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Instruments for log density are 1831, 1861 and 1891 log density. Robust standard errors reported. The 'LM stat' is a LM test statistic for under-identification and 'Under-identif. p-value' is the corresponding p-value. 'Wald F stat' is the first-stage F test statistic corresponding to the excluded instruments and is a test statistic for weak identification. See Stock J.H. and Yogo (2005). Starting from the sample of SP and MP firms, firm-product-year variables are aggregated at the ZE level after demeaning by 8 digit product and year. Each firm-product-year observation is weighted by $1/N_r$ where N_r is the total number of firm-product-year observations corresponding to the ZE r.

Appendix C

Appendix to Chapter 4: The UK's Great Demand Recession

C-1 Data appendix

C-1.1 The ARDx

The ARDx covers around two-thirds of UK economic activity, comprising most SIC 2007 sections, except parts of sections A (agriculture) and K (finance), and all of O (public administration and defence), T (activities of households) and U (extraterritorial organisations). We split the data in order to run the full model on manufacturing and the restricted model on services as described in section 4.2.6 above. Tables C.1 and C.2 show the ARDx coverage for these sectors for firms with 10 or more employees, weighting the survey data using sampling weights based on the stratification of country, 4-digit SIC 2007 headings and employment band. Monetary values in the data are deflated using appropriate series published by the ONS (see below for details) and are reported in 2010 £s.

The recession of 2008-9 is clearly visible for both manufacturing and services, with output and gross value-added dropping substantially for both. The picture for inputs is mixed, however, with manufacturing employment shrinking substantially over the entire 2003-13 period, while real capital stock and real intermediates consumption also declined. The services sectors saw a substantial post-recession rebound in employment, capital stock, and intermediates consumption. We also note a pronounced fall in the number of firms sampled from the manufacturing industries from 2007 to 2008 when the survey changed from the ABI to the ABS, leading us to prefer our within-firm analysis that is likely to be

Table C.1: Manufacturing aggregates by year (using deflated prices)

	Observations	Firms	Output £bn	GVA £bn	Intermediates £bn	Capital £bn	Employment m
2003	9,277	37,048	390.47	110.33	280.14	111.45	2.75
2004	8,910	35,337	396.73	115.79	280.94	108.66	2.60
2005	8,406	34,279	389.92	117.58	272.35	106.38	2.47
2006	7,718	37,732	384.96	123.77	261.19	93.40	2.36
2007	8,360	36,271	421.76	136.63	285.13	91.70	2.31
2008	5,223	34,771	427.56	142.44	285.11	95.14	2.35
2009	4,813	30,349	372.08	118.64	253.44	72.74	2.25
2010	4,901	28,994	393.36	131.94	261.42	72.14	2.09
2011	4,554	28,485	401.68	137.72	263.96	71.35	2.05
2012	4,694	29,378	397.11	135.78	261.33	70.78	2.06
2013	4,543	29,034	394.84	136.57	258.27	70.87	2.07

GVA, output and intermediates all calculated at basic prices from ARDx dataset for firms with 10 or more employees. Manufacturing firms SIC 2007 industries 10-33. Capital stock estimated using ARDx and perpetual inventory method code supplied by ONS to SecureLab users. Employment at time of ARDx sample selection. First column shows un-weighted cell count, all other columns weighted using ARDx sample weights with appropriate auxiliary variable.

less heavily biased by selection effects (see main text).

C-1.2 Revenue, value added and labour measures

As a measure of firms' revenues from production we calculate what the ONS labels 'output at basic prices' by adjusting reported firm turnover for value-added tax, goods bought and sold for resale (i.e. where no production has taken place), changes in stocks and work-in-progress, and changes in stocks of materials, storage and fuels, see Ayoubkhani (2014). We similarly calculate gross value added, measuring intermediates as consumption at purchaser prices, less goods and services bought for resale and changes in stocks of materials, storage and fuels.

We use the firm wage bill as our preferred measure of labour input, instead of the number of employees, as it controls for ability differences across workers over time. In our setting in particular, if firms responded to the recession by substituting workers of different skill levels – with a corresponding difference in wages – we would if using the headcount measure mistakenly pick this up in our estimates of revenue TFP and TFP.

C-1.3 Capital stock

The ARD/ABS does not report a capital stock variable but it does have investment data allowing capital stock to be estimated using the perpetual inventory method. The ONS made available supplementary data files and its Stata code that allow researchers to con-

Table C.2: Services aggregates by year (using deflated prices)

Year	Observations	Firms	Output £bn	GVA £bn	Intermediates £bn	Capital £bn	Employment m
2003	24,059	162,782	888.43	437.05	451.38	151.25	11.42
2004	23,353	162,321	944.75	470.77	473.98	166.74	11.84
2005	22,810	166,425	963.46	475.33	488.13	179.60	12.03
2006	19,643	247,799	983.87	487.82	496.05	159.89	11.67
2007	21,766	270,784	1,089.66	536.77	552.89	179.76	12.38
2008	19,369	282,954	1,167.54	560.97	606.58	197.72	13.65
2009	17,627	179,346	1,074.91	514.30	560.62	170.20	13.93
2010	16,381	177,204	1,088.87	521.79	567.08	166.36	13.40
2011	17,532	181,681	1,096.46	522.17	574.30	168.53	13.60
2012	18,100	194,493	1,117.38	529.87	587.50	171.48	14.26
2013	17,949	201,314	1,179.53	557.94	621.59	177.70	14.58

GVA, output and intermediates all calculated at basic prices from ARDx dataset for firms with 10 or more employees.Services firms SIC 2007 sectors F-U, excluding K. Capital stock estimated using ARDx and perpetual inventory method code supplied by ONS to SecureLab users. Employment at time of ARDx sample selection. First column shows un-weighted cell count, all other columns weighted using ARDx sample weights with appropriate auxiliary variable.

struct firm capital stocks over any time period and with suitable parameters. We construct the variable for the years 2003-2013 using data for the period 1998-2013.

The perpetual inventory method requires a value for the capital stock on the first observation of a firm in the data. The ONS solution is to compare the amount of investment in the survey data with the amount of investment known to occur from National Accounts data, in order to establish the proportion of observed investment to total investment at a (letter code) industry level. This proportion is assumed to also apply to capital stocks. We then obtain a measure of industry level capital stocks (again using the National Accounts estimates of capital services) in the observed data. All that remains is to share this among firms within each industry, which we do based on the IDBR turnover variable.

The procedure also needs to deal with the unbalanced nature of the panel giving rise to many missing observations for investment for all but the largest firms. These missing values are imputed using the average of investment per employee for observed years. We keep observations when capex is negative, even if total capex for a firm over time is negative, dealing with negative capital stocks instead at the end of the process. This procedure first attempts to 'correct' negative capital stocks by rebasing the initial estimate and then recalculating subsequent observations, iterating five times or until the process leads to a positive capital stock. Those that remain negative are removed from our final dataset.

Standard ONS industry and national accounts data contains 10 types of capital, while the ABS/ABI has investment data for only three types: buildings, vehicles and plant and machinery, the latter being a residual 'other' category. Our preferred capital stock is to use plant and machinery, following Harris and Drinkwater (2000); Harris and Moffat (2017). In doing so we do not require detailed user costs in order to aggregate different types of capital stock. We use an annual depreciation rate of 10% for plant and machinery.

Firms are also asked about capital hires in the ABS/ABI surveys but, since in our model capital must be pre-decided in each period, we do not add this to the capital stock as it is presumably more flexible than the fixed stock. Instead, it will be included in firms' intermediates expenditures.

C-1.4 Product concordance

The core Prodcom list is published annually by Eurostat, however until 2005 it contained some aggregate codes and some national-level codes. These national level codes are dropped from the analysis. The Prodcom list typically has minor changes from year to year as new codes are added, obsolete products removed, and other codes either merged or split. However, the 2007/8 revision of NACE, and corresponding move from UK Standard Industry Classification (SIC) 2003 to SIC 2007, involved wholesale changes to Prodcom when 4,396 Prodcom codes were made obsolete and replaced with 3,864 new codes. Since we need to track a firm's production and sales of a product over time, we need a concordance procedure to give us a consistent set of products. For this we borrow the procedure and Stata code due to Van Beveren et al. (2012). The concordance works by grouping together new codes that map from obsolete codes into a synthetic code, and then keeps track of these family trees. Each 'tree' becomes a new product, and we are left with 3,795 consistent products. We drop firm-product-year observations that we are not able to match using the concordance procedure.

C-1.5 Deflators

We deflate all monetary variables to constant 2010 prices. For output price deflators we use a series provided by the ONS, 'Experimental Industry Level Deflators'[I].

These are a mixture of 2-, 3- and 4-digit deflators produced by aggregating industry

[I]Downloaded from https://www.ons.gov.uk/economy/inflationandpriceindices/adhocs/006718industryleveldeflatorsexperimentaluk1997to2015 product deflators based on their use of products in line with the National Accounts supple-use framework. Finance (SIC section K) is excluded, and we drop the section from the services analysis. For most service sectors, the deepest level of detail available is at the SIC division (2-digit) level.

Detailed input price deflators are available for manufacturing but not services. For manufacturing we turn to the ONS Producer Price Index (PPI) time series[II] which provides Gross Sector Input (GSI) deflators at the one to four digit SIC 2007 level, depending on GSI sub-section. Each GSI sub-section can span several SIC divisions (the 2 digit level), while data for some levels of aggregation are not available. We group the data in such a way as to produce deflators at the SIC division level. Most[III] SIC divisions map 1:1 to the GSI deflators we have available[IV].

For services inputs we deflate using the same output price series described above. Our estimates of the plant and machinery capital stock are deflated by the ONS deflator series[V] for other machinery & equipment by SIC section.

We use the output price deflators to convert wage costs into real terms. Calculations of value added for manufacturing are made using double deflation, deflating output and inputs separately before calculating value added.

C-1.6 Merging the ARDx and Prodcom

The merging process, and the impact it has on the mean values of key variables, is shown in Table C.3. Firms in Prodcom and the ARDx (column 2) are on average larger by every metric than those only in the ARDx (column 1) while single product firms are smaller (column 3). Firms in the final single product estimation sample (column 4) have slightly higher output, intermediates and capital stock than all the single product firms, and slightly lower value added, wage bills and employment.

[II]Downloaded from `https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/producerpriceindex`

[III]The exceptions are Food (SIC division 10), Beverages (11) and Tobacco (12) for which we use the combined GSI sub-section 'Food, Beverages and Tobacco'.

[IV]We map SIC division 22 ('Rubber and Plastic Products') to GSI sub-section 'Other manufactured goods n.e.c.'; division 23 ('Other non-metallic mineral products') to 'Cement, Lime & Plaster'; division 27 ('Electrical equipment') to 'Computer, Elect & Optical Products'; division 30 ('Other transport equipment') to 'Motor Vehicles'; and division 33 ('Repair and Installation') to 'Other manufactured goods n.e.c.'.

[V]Downloaded from: `https://www.ons.gov.uk/economy/economicoutputandproductivity/output/datasets/capitalservicesestimates`

Table C.3: Summary statistics and data constraints

	All variables available	...plus Prodcom	... plus single product	... plus data constraints
Output £m	45.86	50.65	35.80	37.08
Value-added £m	15.35	17.52	12.67	12.11
Intermediates £m	30.51	33.13	23.13	24.97
Capital £m	10.07	11.33	8.42	9.28
Wage Bill £m	8.49	9.72	7.35	6.89
Employment	239.40	279.48	217.79	217.07
Average N per year	5,731	4,268	2,266	1,140

Means of variables by sample. 'All variables available' is the number of manufacturing sector firms in the ARDx that have a) at least 10 employees and b) have the following variables available: employment, total turnover ex. VAT, purchases of goods and materials, capital stock, total wages and salaries. '...plus Prodcom' adds the requirement that the firm-year observation is also in the Prodcom dataset. '...plus single product' adds the requirement that at least 90% of a firm's output at basic prices is accounted for by sales of a single product . '...plus data constaints' adds the requirement that Prodcom measures a non-zero quantity of production, and that firm revenues reported by Prodcom are within 30% of the output calculated from the ARDx.

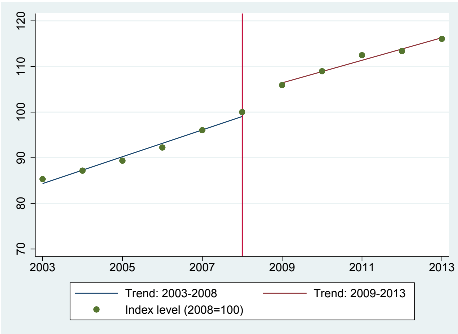
C-2 Additional results

Table C.4: Revenue production function estimates for the whole sample of manufacturing firms in the ARDx.

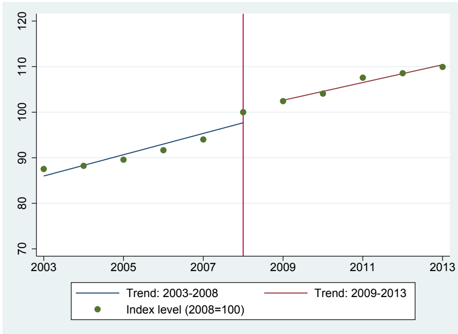
	log output		log value-added	
	(1) OLS	(2) WLD	(3) OLS	(4) WLD
log wage bill	0.421	0.358	0.946	0.914
	(0.003)***	(0.006)***	(0.002)***	(0.006)***
log intermediates	0.549	0.646		
	(0.002)***	(0.016)***		
log capital stock	0.016	0.013	0.030	0.064
	(0.001)***	(0.007)*	(0.001)***	(0.008)***
R^2	0.99		0.99	
Obs	63,043	16,050	63,043	16,050

All regressions include 2-digit industry and year dummies. Standard errors clustered by firm. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

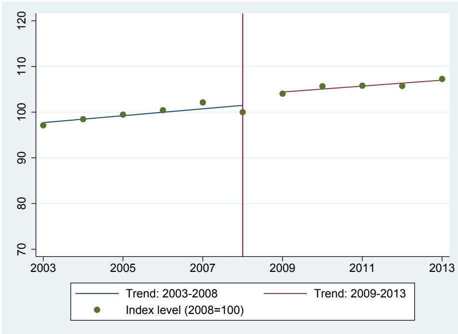
Figure C.1: Real, actual and official price index changes: manufacturing, revenue-weighted, within sample



(a) Actual prices



(b) Official price index



(c) Real prices

Indices of prices (2008=100) calculated using revenue-weighted changes from $t-1$ to t for the manufacturing firms within sample.

Table C.5: Real, actual and official price index changes: manufacturing, un-weighted, within sample

	Δ real prices	Δactual prices	Δ official price index	Obs
2004	0.009	0.018	0.009	641
2005	0.009	0.027	0.018	564
2006	0.002	0.025	0.023	617
2007	0.014	0.042	0.028	594
2008	-0.001	0.069	0.069	401
2009	0.024	0.047	0.024	443
2010	0.015	0.036	0.021	496
2011	-0.001	0.039	0.040	450
2012	0.003	0.012	0.009	478
2013	-0.001	0.012	0.012	477
2003-2008	0.007	0.033	0.026	2,817
2008-2013	0.008	0.029	0.021	2,344

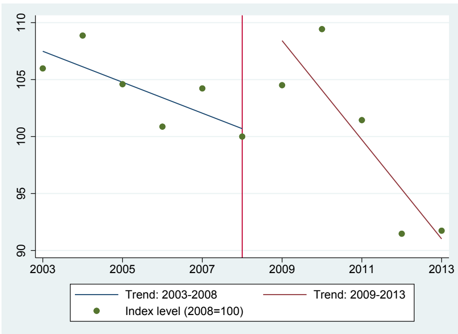
The Table shows mean un-weighted within-firm changes from $t-1$ to t for the within sample. The final two rows show the mean of changes over the two periods using all the annual observations.

Table C.6: Real, actual and official price index changes: manufacturing, multi-product firms, within changes, revenue-weighted

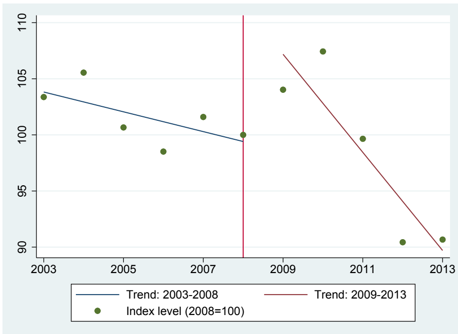
	Δ real prices	Δactual prices	Δ official price index	Obs
2004	0.002	0.011	0.008	2,189
2005	-0.007	0.008	0.015	1,891
2006	-0.005	0.017	0.023	2,376
2007	0.007	0.034	0.027	2,368
2008	-0.012	0.075	0.087	1,657
2009	0.010	0.040	0.030	2,079
2010	0.005	0.023	0.017	2,267
2011	0.008	0.056	0.048	2,277
2012	-0.012	0.002	0.014	2,243
2013	0.007	0.024	0.017	2,210
2003-2008	-0.003	0.029	0.032	10,481
2008-2013	0.003	0.029	0.026	11,076

The Table shows mean revenue-weighted within-firm-product changes from $t-1$ to t for the multi-product firms sample, where the weights used are average revenues in the two periods. The final two rows show the mean of changes over the two periods using all the annual observations.

Figure C.2: CES demand measure for different values of σ: manufacturing, within sample, un-weighted



(a) CES demand measure $\sigma = 5$, un-weighted



(b) CES demand measure $\sigma = 10$, un-weighted

Indices of changes in CES demand measures (2008=100). Indices are computed for the manufacturing firms within sample using un-weighted within changes and two alternative values for the elasticity of substitution: $\sigma = 5$ and $\sigma = 10$.

Table C.7: Chow tests on changes in revenue TFP and its components: manufacturing, within sample, revenue weighted.

	F(k,N-2k)	p	N
Δ TFP-R (WLD)	54.89	0.000	5,161
Δ TFP-R (DGKP)	28.54	0.000	5,161
Δa	19.16	0.000	5,161
$\Delta\lambda$	27.43	0.000	5,161
$\Delta\omega$	35.93	0.000	5,161
$\Delta\mu$	28.64	0.000	5,161
Δ scale	54.43	0.000	5,161

The Table shows the F-stat and corresponding p-value from Chow tests of the null hypothesis of no structural break in each listed variable in 2008, where $k = 2$.

Table C.8: Chow tests on changes in revenue TFP and its components: services, within sample, revenue weighted.

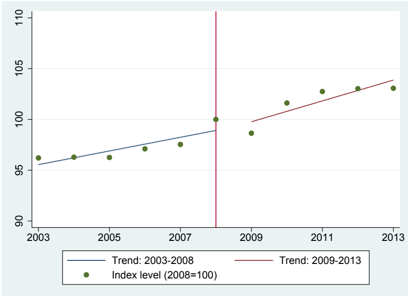
	F(k,N-k)	p	N
Δ TFP-R (WLD)	36.12	0.000	66,146
$\Delta\omega$	17.71	0.000	66,146
$\Delta\mu$	9.39	0.000	66,146
Δ scale	156.43	0.000	66,146
Δ adjusted ω	34.71	0.000	66,146
Δ adjusted scale	26.20	0.000	66,146

The Table shows the F-stat and corresponding p-value from Chow tests of the null hypothesis of no structural break in each listed variable in 2008, where $k = 2$.

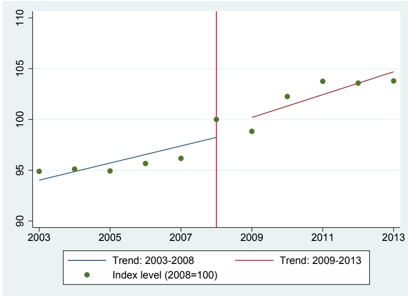
Table C.9: Manufacturing. Changes of revenue TFP and its components over the period 2003-2013. Multi-product firms, within changes, revenue-weighted.

	Δ TFP-R (WLD)	Δ TFP-R (DGKP)	Δa	$\Delta\lambda$	$\Delta\omega$	$\Delta\mu$	Δ scale	Obs
2004	0.001	0.002	0.000	-0.022	-0.022	-0.002	0.019	2,189
2005	-0.000	-0.002	0.006	0.004	0.010	0.004	-0.027	1,891
2006	0.009	0.008	0.013	0.061	0.074	0.005	-0.017	2,376
2007	0.004	0.005	-0.002	0.088	0.087	0.009	0.021	2,368
2008	0.026	0.041	0.053	-0.030	0.023	-0.000	-0.074	1,657
2009	-0.014	-0.012	-0.023	-0.170	-0.192	-0.017	-0.057	2,079
2010	0.031	0.036	0.031	0.272	0.302	0.023	0.008	2,267
2011	0.012	0.016	0.007	0.075	0.082	0.007	0.013	2,277
2012	0.003	-0.002	0.010	0.075	0.085	0.009	-0.025	2,243
2013	0.000	0.002	-0.005	0.018	0.014	0.000	0.009	2,210
2003-2008	0.008	0.011	0.014	0.023	0.036	0.003	-0.014	10,481
2008-2013	0.007	0.008	0.005	0.058	0.062	0.005	-0.009	11,076

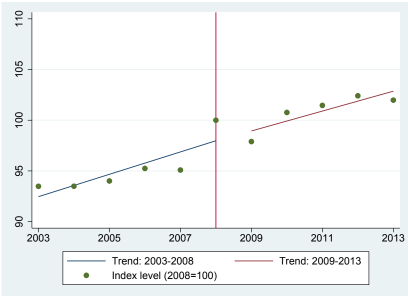
Figure C.3: Manufacturing. Evolution of revenue TFP and its components over the period 2003-2013. Multi-product firms, within changes, revenue-weighted.



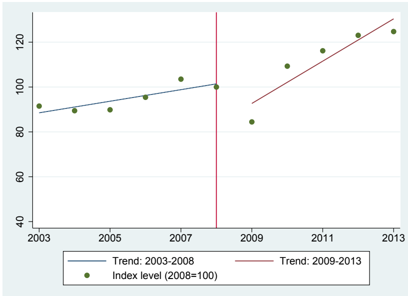
(a) WLD rev TFP



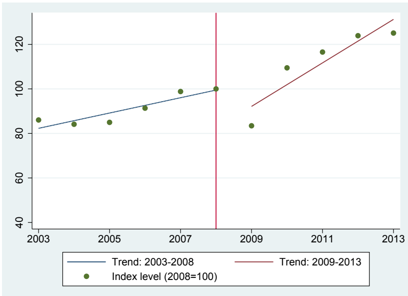
(b) DGKP rev TFP



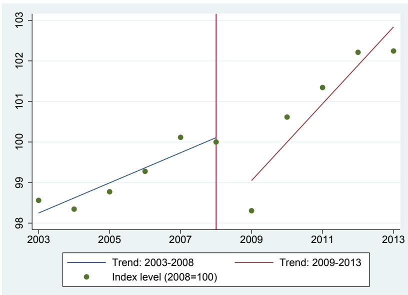
(c) a



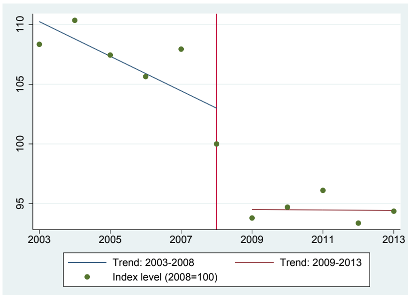
(d) λ



(e) ω



(f) μ



(g) scale

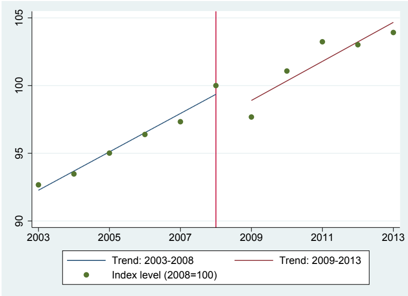
Table C.10: Manufacturing. Changes of revenue TFP and its components over the period 2003-2013. Within sample, un-weighted

	Δ TFP-R (WLD)	Δ TFP-R (DGKP)	Δa	$\Delta\lambda$	$\Delta\omega$	$\Delta\mu$	Δ scale	Obs
2004	0.008	0.018	0.010	0.162	0.171	0.016	0.036	641
2005	0.016	0.002	-0.007	0.078	0.071	0.008	0.001	564
2006	0.014	0.015	0.013	0.023	0.036	0.002	0.008	617
2007	0.010	0.014	0.001	0.150	0.150	0.016	0.015	594
2008	0.028	0.026	0.027	0.321	0.348	0.032	-0.054	401
2009	-0.025	-0.026	-0.050	-0.260	-0.310	-0.027	-0.084	443
2010	0.037	0.030	0.016	0.179	0.194	0.019	0.026	496
2011	0.022	0.022	0.023	0.195	0.218	0.019	0.015	450
2012	-0.001	-0.005	-0.009	0.050	0.042	0.007	0.001	478
2013	0.009	0.011	0.012	0.132	0.144	0.013	0.019	477
2003-2008	0.014	0.015	0.008	0.135	0.142	0.014	0.005	2,817
2008-2013	0.009	0.007	-0.001	0.063	0.062	0.007	-0.003	2,344

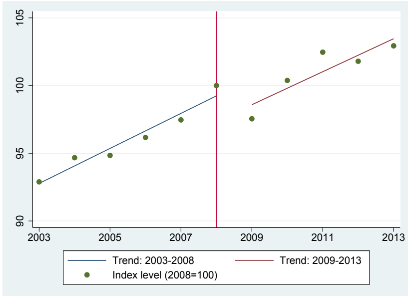
Table C.11: Services. Changes of revenue TFP and its components over the period 2003-2013. Within sample, un-weighted

	Δ TFP-R (WLD)	$\Delta\omega$	$\Delta\mu$	Δ scale	Δ adjusted ω	Δ adjusted scale	Obs
2004	0.002	0.096	0.008	0.042	0.044	-0.043	8,387
2005	-0.008	0.038	0.005	0.022	0.019	-0.026	7,813
2006	0.004	-0.024	-0.004	0.022	0.003	0.001	6,438
2007	0.006	0.129	0.008	0.033	0.058	-0.052	6,266
2008	-0.017	-0.219	-0.021	-0.005	-0.105	0.088	5,674
2009	-0.005	0.052	0.011	-0.048	-0.005	-0.000	6,493
2010	0.012	0.073	0.005	0.018	0.049	-0.037	6,079
2011	-0.002	-0.150	-0.015	0.012	-0.058	0.056	5,966
2012	0.011	0.053	0.002	0.012	0.044	-0.034	6,404
2013	0.002	-0.013	-0.003	0.025	0.016	-0.014	6,626
2003-2007	0.001	0.061	0.004	0.030	0.031	-0.030	28,904
2007-2013	0.000	-0.030	-0.003	0.002	-0.008	0.008	37,242

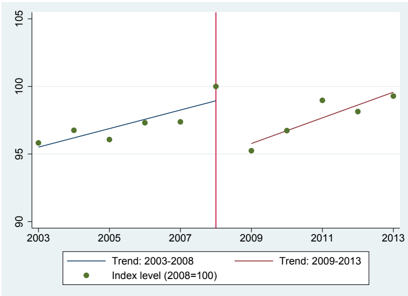
Figure C.4: Manufacturing. Evolution of revenue TFP and its components over the period 2003-2013. Within sample, un-weighted



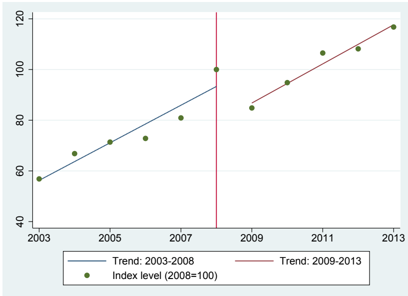
(a) WLD rev TFP



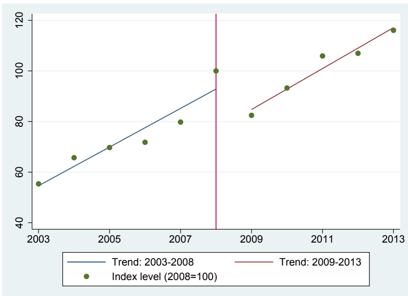
(b) DGKP rev TFP



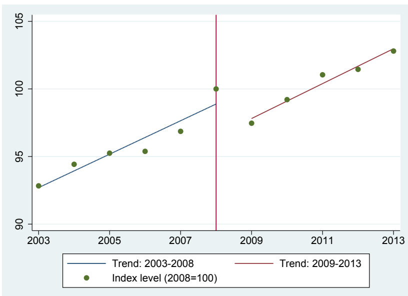
(c) a



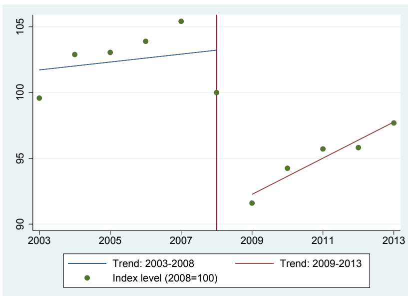
(d) λ



(e) ω

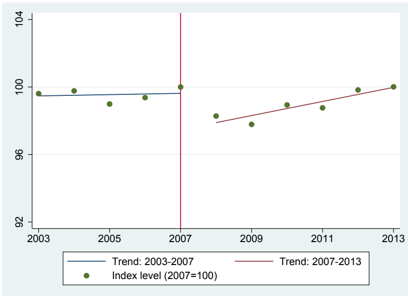


(f) μ

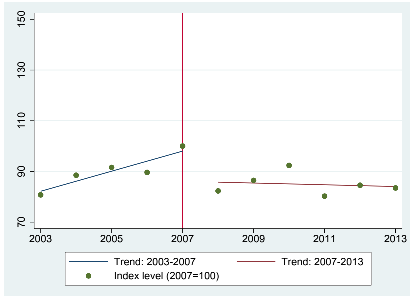


(g) scale

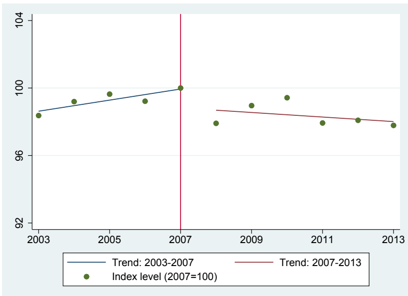
Figure C.5: Services. Evolution of revenue TFP and its components over the period 2003-2013. Within sample, un-weighted



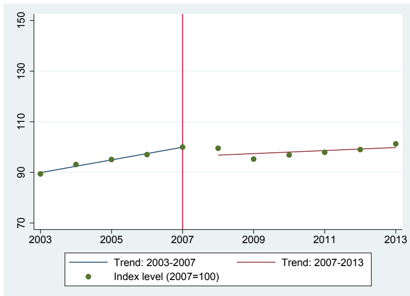
(a) WLD rev TFP



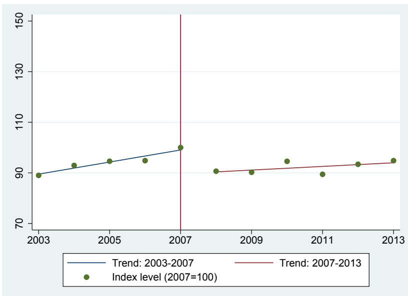
(b) ω



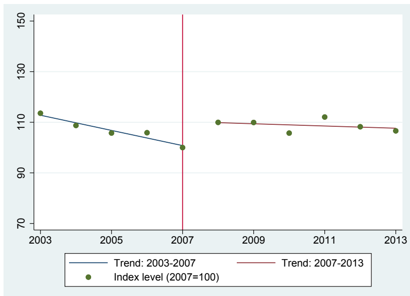
(c) μ



(d) scale



(e) adjusted ω



(f) adjusted scale

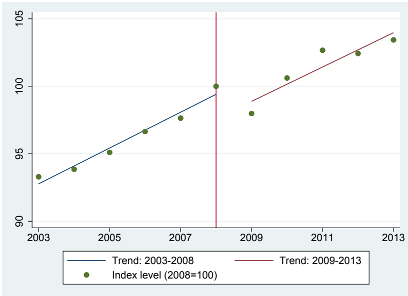
Table C.12: Manufacturing. Changes of revenue TFP and its components over the period 2003-2013. Within sample, employment-weighted

	Δ TFP-R (WLD)	Δ TFP-R (DGKP)	Δa	$\Delta\lambda$	$\Delta\omega$	$\Delta\mu$	Δ scale	Obs
2004	0.006	0.019	0.015	0.214	0.229	0.019	0.030	641
2005	0.013	-0.002	-0.017	0.065	0.049	0.006	0.007	564
2006	0.017	0.019	0.019	0.111	0.129	0.010	0.009	617
2007	0.011	0.017	-0.007	0.155	0.148	0.011	0.024	594
2008	0.025	0.022	0.030	0.352	0.382	0.033	-0.052	401
2009	-0.022	-0.019	-0.056	-0.306	-0.362	-0.031	-0.080	443
2010	0.028	0.022	0.009	-0.013	-0.004	-0.000	0.035	496
2011	0.022	0.022	0.017	0.295	0.312	0.026	0.011	450
2012	-0.003	-0.008	-0.000	-0.027	-0.027	-0.001	0.009	478
2013	0.011	0.014	0.001	0.177	0.178	0.014	0.021	477
2003-2008	0.014	0.015	0.008	0.176	0.184	0.016	0.006	2,817
2008-2013	0.007	0.006	-0.006	0.024	0.018	0.002	-0.001	2,344

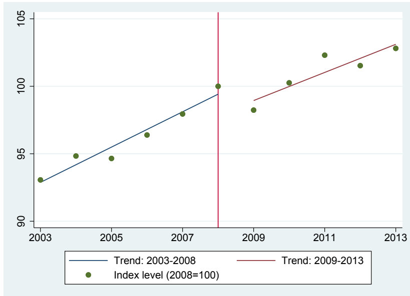
Table C.13: Services. Changes of revenue TFP and its components over the period 2003-2013. Within sample, employment-weighted

	Δ TFP-R (WLD)	$\Delta\omega$	$\Delta\mu$	Δ scale	Δ adjusted ω	Δ adjusted scale	Obs
2004	0.005	0.051	0.002	0.042	0.044	-0.039	8,387
2005	-0.008	-0.013	-0.002	0.047	-0.030	0.022	7,813
2006	0.002	-0.036	-0.004	0.024	0.019	-0.017	6,438
2007	0.006	0.174	0.013	0.031	0.061	-0.055	6,266
2008	-0.021	-0.238	-0.019	0.013	-0.118	0.098	5,674
2009	-0.006	-0.033	-0.001	-0.029	-0.090	0.084	6,493
2010	0.009	0.131	0.010	0.021	0.081	-0.071	6,079
2011	-0.002	-0.298	-0.026	0.012	-0.068	0.066	5,966
2012	0.011	0.062	0.001	0.029	0.058	-0.047	6,404
2013	0.001	-0.064	-0.006	0.040	0.025	-0.024	6,626
2003-2007	0.001	0.045	0.002	0.036	0.024	-0.023	28,904
2007-2013	-0.001	-0.070	-0.007	0.015	-0.016	0.015	37,242

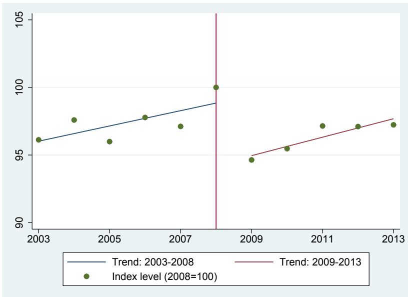
Figure C.6: Manufacturing. Evolution of revenue TFP and its components over the period 2003-2013. Within sample, employment-weighted



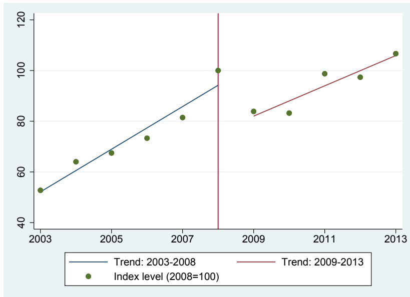
(a) WLD rev TFP



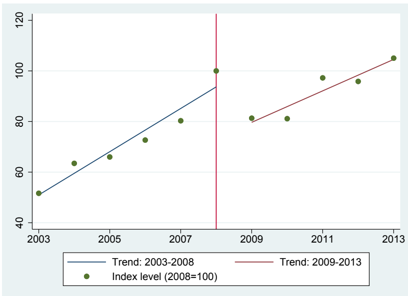
(b) DGKP rev TFP



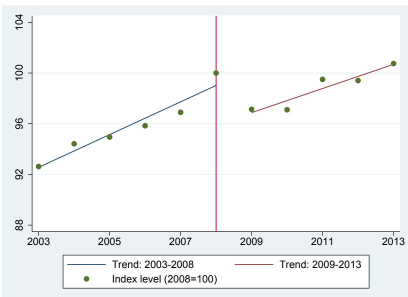
(c) a



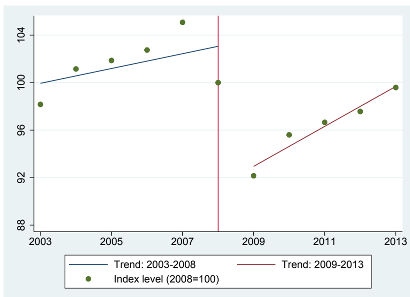
(d) λ



(e) ω

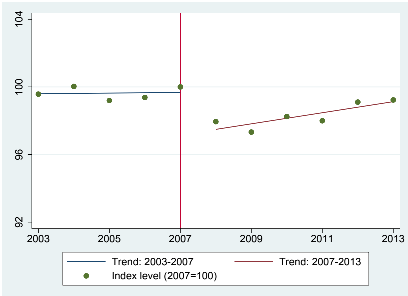


(f) μ

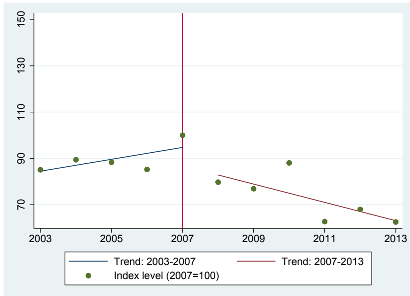


(g) scale

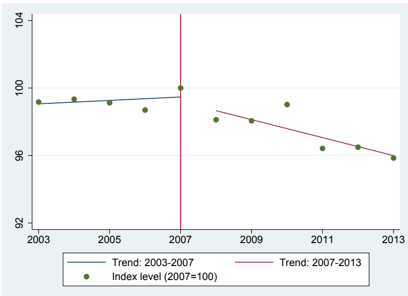
Figure C.7: Services. Evolution of revenue TFP and its components over the period 2003-2013. Within sample, employment-weighted



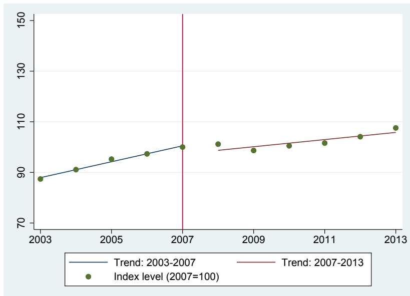
(a) WLD rev TFP



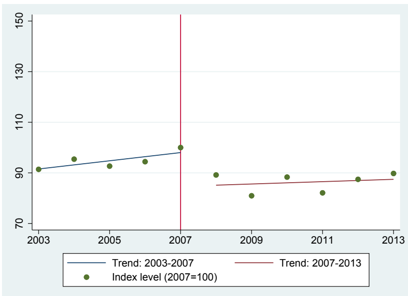
(b) ω



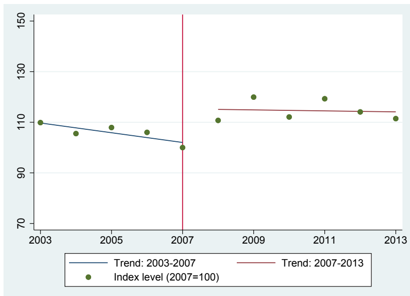
(c) μ



(d) scale



(e) adjusted ω



(f) adjusted scale

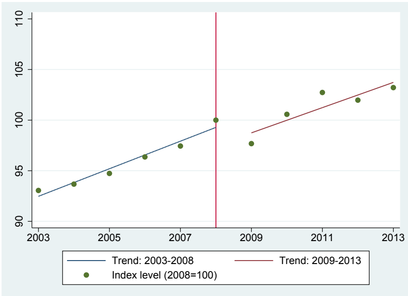
Table C.14: Manufacturing. Changes of revenue TFP and its components over the period 2003-2013. FMMM estimation procedure, within sample, revenue-weighted

	Δ TFP-R (WLD)	Δ TFP-R (FMMM)	Δa	$\Delta\lambda$	$\Delta\omega$	$\Delta\mu$	Δ scale	Obs
2004	0.007	0.024	0.011	0.123	0.133	0.012	0.031	576
2005	0.011	-0.011	-0.012	0.026	0.014	0.003	0.022	498
2006	0.018	0.046	0.039	0.060	0.099	0.006	-0.016	548
2007	0.012	0.025	0.015	-0.004	0.010	0.001	0.026	519
2008	0.028	0.032	0.064	0.149	0.213	0.016	-0.045	355
2009	-0.025	-0.000	-0.046	-0.096	-0.142	-0.015	-0.093	364
2010	0.031	0.031	0.003	0.005	0.008	0.001	0.032	435
2011	0.023	0.038	0.050	0.168	0.217	0.017	0.003	387
2012	-0.008	-0.015	-0.021	0.005	-0.016	0.000	0.018	405
2013	0.013	0.012	-0.012	0.094	0.082	0.007	0.025	417
2003-2008	0.015	0.024	0.023	0.070	0.094	0.007	0.004	2,496
2008-2013	0.007	0.013	-0.005	0.035	0.030	0.002	-0.003	2,008

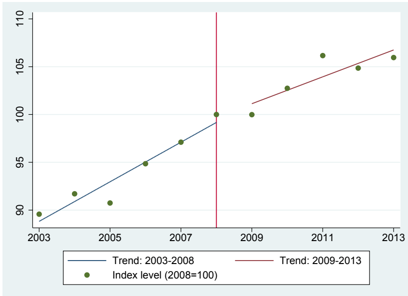
Table C.15: Services. Changes of revenue TFP and its components over the period 2003-2013. FMMM estimation procedure, within sample, revenue-weighted

	Δ TFP-R (FMMM)	$\Delta\omega$	$\Delta\mu$	Δ scale	Δ adjusted ω	Δ adjusted scale	Obs
2004	-0.004	0.083	0.007	0.052	0.050	-0.054	8,387
2005	-0.012	-0.034	-0.002	0.049	-0.017	0.006	7,813
2006	0.008	-0.017	-0.002	0.035	0.026	-0.017	6,438
2007	-0.001	0.166	0.014	0.051	0.113	-0.114	6,266
2008	-0.014	-0.159	-0.012	0.010	-0.124	0.110	5,674
2009	-0.027	-0.018	0.001	-0.016	0.011	-0.038	6,493
2010	0.011	0.097	0.008	0.014	0.079	-0.068	6,079
2011	0.018	-0.132	-0.013	-0.002	-0.108	0.126	5,966
2012	0.030	0.101	0.006	0.009	0.088	-0.058	6,404
2013	0.004	-0.022	-0.002	0.024	-0.012	0.016	6,626
2003-2007	-0.002	0.051	0.004	0.047	0.044	-0.046	28,904
2007-2013	0.004	-0.019	-0.002	0.007	-0.008	0.012	37,242

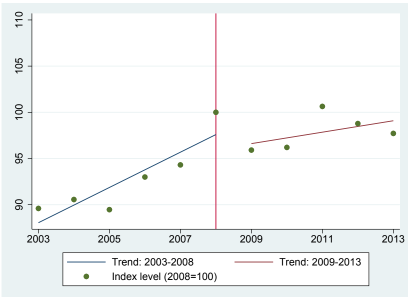
Figure C.8: Manufacturing. Evolution of revenue TFP and its components over the period 2003-2013. FMMM estimation procedure, within sample, revenue-weighted



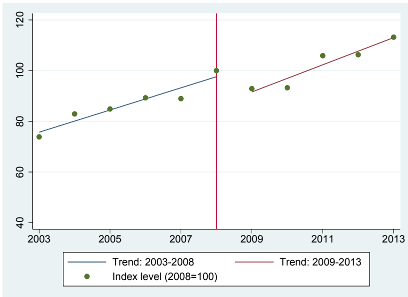
(a) WLD rev TFP



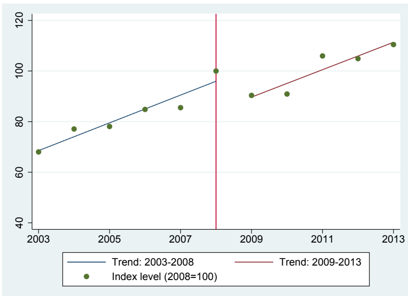
(b) FMMM rev TFP



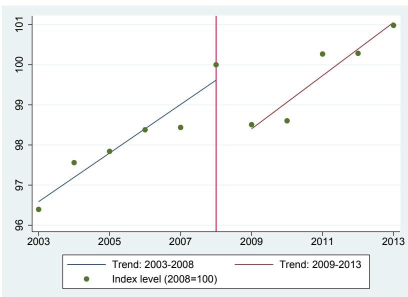
(c) a



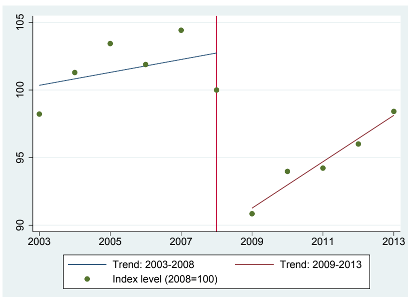
(d) λ



(e) ω

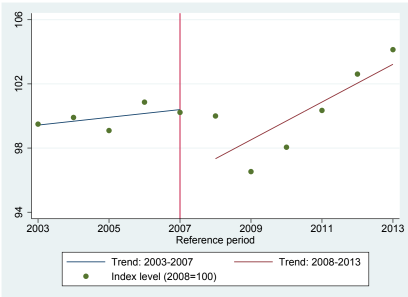


(f) μ

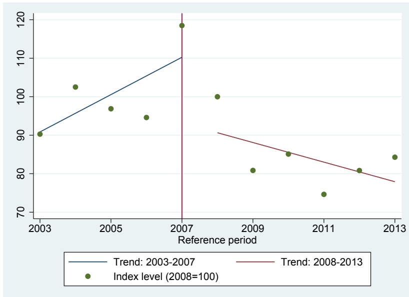


(g) scale

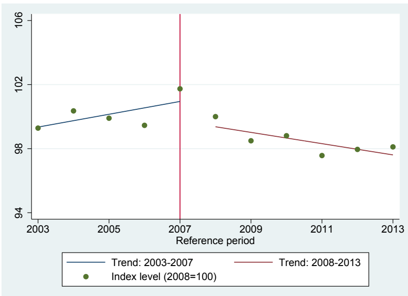
Figure C.9: Services. Evolution of revenue TFP and its components over the period 2003-2013. FMMM estimation procedure, within sample, revenue-weighted



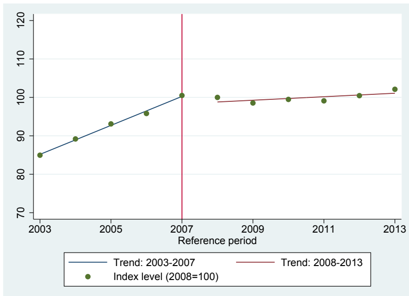
(a) FMMM rev TFP



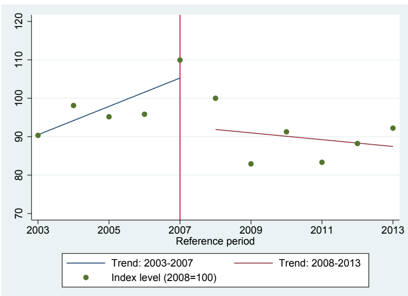
(b) ω



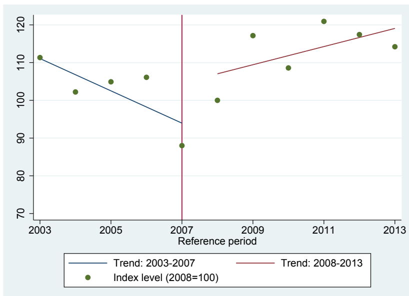
(c) μ



(d) scale



(e) adjusted ω

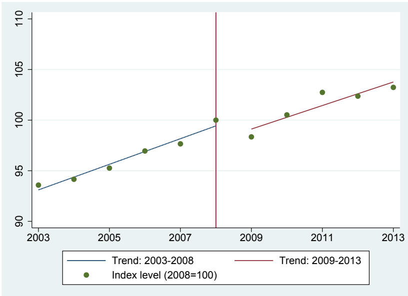


(f) adjusted scale

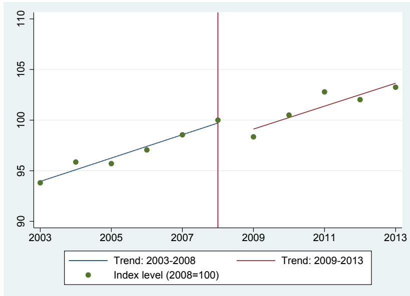
Table C.16: Manufacturing. Changes of revenue TFP and its components over the period 2003-2013. Translog production function, within sample, revenue-weighted

	Δ TFP-R (WLD)	Δ TFP-R (DGKP)	Δa	$\Delta\lambda$	$\Delta\omega$	$\Delta\mu$	Δ scale	Obs
2004	0.006	0.021	0.012	0.227	0.240	0.020	0.040	659
2005	0.012	-0.002	-0.015	-0.005	-0.020	-0.001	0.018	572
2006	0.018	0.015	0.005	0.134	0.140	0.011	0.024	624
2007	0.008	0.016	0.000	0.222	0.222	0.017	0.033	605
2008	0.025	0.015	0.047	0.144	0.191	0.016	-0.036	396
2009	-0.018	-0.018	-0.051	-0.210	-0.261	-0.022	-0.059	434
2010	0.023	0.023	0.013	0.318	0.332	0.028	0.039	504
2011	0.024	0.024	0.022	0.277	0.299	0.024	0.021	457
2012	-0.004	-0.008	-0.014	-0.127	-0.141	-0.012	0.015	485
2013	0.009	0.013	-0.004	0.197	0.194	0.015	0.024	484
2003-2008	0.013	0.013	0.009	0.149	0.158	0.013	0.017	2,856
2008-2013	0.007	0.007	-0.006	0.093	0.087	0.007	0.009	2,364

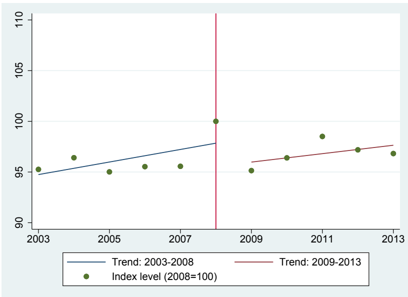
Figure C.10: Manufacturing. Changes of revenue TFP and its components over the period 2003-2013. Translog production function, within sample, revenue-weighted



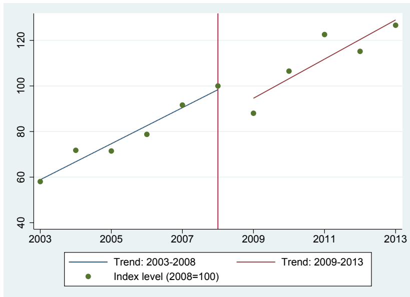
(a) WLD rev TFP



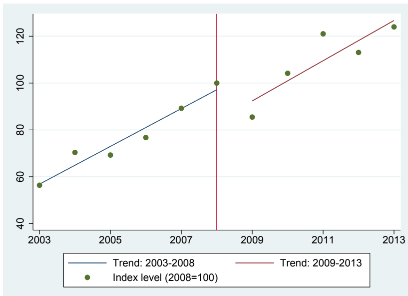
(b) DGKP rev TFP



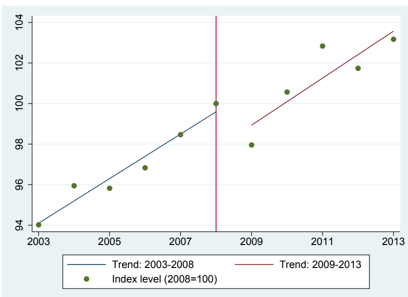
(c) a



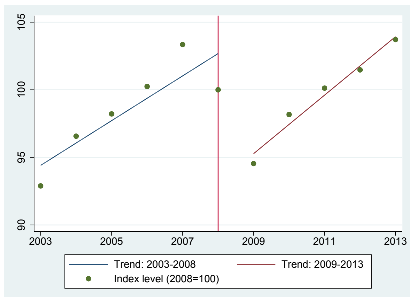
(d) λ



(e) ω



(f) μ



(g) scale