Productivity Dispersion, Import Competition, and Specialization in Multi-product Plants

Scott Orr *

Sauder School of Business

University of British Columbia

January 11, 2019

Abstract

Import competition can increase plant level efficiency, as competitive pressure forces plants to reallocate more of their resources towards their highest performing products. There has been little empirical work directly documenting these gains, since within plant differences in quality and productivity, and the allocation of inputs across production lines, are rarely observed. This paper develops a flexible methodology that uses profit maximization conditions to solve for unobserved input allocations, TFP, and quality across product lines in multi-product plants. I apply this methodology to a panel of plants manufacturing machinery in India from 2000-2007. I emphasize three results. First, there is substantial variation in both TFP and quality across product lines within a given plant. Second, quality and quantity based TFP (TFPQ) are negatively correlated. Third, increases in Chinese import competition led plants to reallocate their inputs towards higher quality products and away from their high TFPQ products, thereby generating within plant quality improvements, rather than productivity gains.

JEL codes: D24, F12, F14.

*I wish to thank my committee, Daniel Trefler, Victor Aguirregabiria, and Aloysius Siow, for their advice on this project, as well as Stephen Ayerst, Bernardo Blum, Kunal Dasgupta, Jan De Loecker, Kevin Devereux, Daniel Ershov, Garth Frazer, Leandro Freylejer, Nicolas Gendron-Carrier, Daniel Goetz, Ashique Habib, Ignatius Horstmann, Amit Khandelwal, Nicholas Li, Kevin Lim, Gaelan MacKenzie, Mathieu Marcoux, Pamela Medina, Juan Morales, Peter Morrow, Jean-William P.Laliberté, Swapnika Rachapalli, Eduardo Souza-Rodrigues, Sharon Traiberman, Eric Verhoogen, Frederic Warzynski, David Weinstein, Erhao Xie, Genet Zinabou, and a number of participants at the 2017 CEA and EARIE conferences, the 2018 Danish International Economics Workshop, as well as seminar participants at the Bank of Canada, Columbia, McMaster, UBC Sauder, University of Chicago Booth, and the U of T SWEAT, International Economics, and Industrial Organization seminars, who provided very helpful feedback. Any remaining errors are my own. Financial support from the Social Sciences and Humanities Research Council (SSHRC) is gratefully acknowledged. Please send any comments to scott.orr@sauder.ubc.ca

1 Introduction

Quantitative estimates of the gains from trade are of first-order importance for developing trade policy. Empirical work examining the effect of international trade on aggregate productivity has been very successful on this front, finding that trade liberalization can lead to sizeable efficiency gains that are straightforward to quantify.[1] These efficiency gains are generally found to come from two key sources: *across firm reallocations*, where high productivity firms gain market share at the expense of low productivity firms, and *within firm* efficiency gains, where existing firms increase their productivity.[2]

Recent studies, including Eckel and Neary (2010), Bernard et al. (2011), Mayer et al. (2014) and Mayer et al. (2016), have emphasized that trade induced efficiency gains may also be generated by *within firm reallocations*. Specifically, import competition may force firms to reallocate a larger proportion of their inputs towards their highest performing products, improving firm level efficiency. However, these within firm reallocations can only generate significant efficiency gains if there are sizeable differences in productivity and quality across production lines within a firm. Unfortunately, there has been very little empirical work measuring the magnitude of within firm heterogeneity.[3] This is largely because data on multi-product firms will typically only record input use at the firm level, rather than the firm-product level, thereby making standard approaches to productivity estimation infeasible.

In this paper, I develop a flexible methodology that identifies unobserved input allocations, TFP, and quality across product lines in multi-product firms. This approach exploits the fact that firm-product level prices incorporate information on within firm input use. More formally, I show that profit maximization restrictions, combined with estimates of a firm-product level demand function, provide enough information for prices and quantities to reveal input use across production lines. One can then use these restrictions to estimate a firm-product level production function for multi-product firms, generating estimates of firm-product level TFP. Since the general approach requires estimation of a product-level demand function, I use differences in the demand residual across firm-products to estimate product specific *quality*, as in Khandelwal (2010) and Amiti and Khandelwal (2013).[4] Altogether, the approach generates estimates of firm-product input use, and two dimensions of within firm heterogeneity, TFP and quality, allowing the researcher to answer questions related to within-firm reallocations and efficiency.

I then apply my methodology to a panel of plants manufacturing machinery in India, using data

[1]See Melitz and Trefler (2012) and De Loecker and Goldberg (2014) for a review of this literature.

[2]For the theoretical underpinnings of these across firm reallocations, see Melitz (2003). See Pavcnik (2002), Bernard and Jensen (2004) and Trefler (2004), for empirical work documenting the existence of both across and within-firm productivity gains.

[3]Some notable exceptions include Hottman et al. (2016), who examine within firm marginal cost and quality dispersion in US retail, and Garcia-Marin and Voigtländer (2013) and Garcia-Marin and Voigtländer (2017), who estimate within-plant TFP dispersion in Chile using a novel dataset containing information on cost shares by product line within a plant.

[4]Since products with large demand residuals generate a higher level of purchases than can fully be explained by price, by revealed preference these products must include product characteristics that consumers find more appealing. I will refer to differences in these unobserved characteristics within a given product-code as quality variation. Note, however, that these product characteristics can also be immaterial characteristics, such as the quantity of advertising, that nevertheless increase quantity sold conditional on price.

from the 2000-2007 Annual Survey of Industries (ASI). I find that around one third of the variation in product level quality and TFP is due to *within-plant* variation, implying that within-firm reallocations can indeed have sizeable efficiency effects. I further find that tariff cuts, as well as increased Chinese import competition, led to within plant reallocations towards a plant's highest *quality* products, but away from their products with the highest *quantity* based TFP, or TFPQ. This is in large part because I find that TFPQ and quality are negatively correlated. Put differently, quality is costly to produce, in the sense that it often requires more inputs per unit of output. While these reallocations directly decrease plant level TFPQ, increased quality specialization tends to increase revenue based TFP, or TFPR. On net, I find that these reallocations were approximately *TFPR-neutral*. Hence, while I find evidence for within-plant reallocations in response to increased import competition, these reallocations generally do not increase plant-level productivity, but rather imply large gross changes in the *composition* of high versus low quality goods within the plant.

This paper generates some of the first estimates of the impact of within-plant reallocations on plant level efficiency, in large part because there is no standard approach for estimating productivity in datasets containing multiple outputs produced by the same firm or plant.[5] To highlight the issues generated by multi-product firms in productivity analysis, consider a firm producing two outputs using only labour. Suppose that a researcher observes the level of output across the two production lines, as well as the total quantity of labour purchased by the firm. Further suppose that the researcher is confident that both production lines are characterized by constant returns to scale. Without further information, differences in output across production lines can either be due to differences in labour input, or differences in unit labour requirements. As a result, productivity and differences in input use are not separately identified.

I show that if product level prices are observed, then profit maximization restrictions can be used to separately identify input use from productivity dispersion. For example, if the markets for each output are characterized by perfect competition, the price of each good will equal its marginal cost. Given constant returns to scale, this implies that the price of each good will be equal to the wage over the unit labour requirement for that product. As a result, the ratio of output prices within the plant will be proportional to ratio of unit labour requirements, which means that price variation *reveals* productivity dispersion. One can then combine this information with the structure of the production technologies, as well as an aggregate resource constraint, to generate four equations that uniquely pin down the four unknowns; the two unobserved unit labour requirements, and the two unobserved labour allocations.[6]

While the above example imposes a number of strong assumptions (single input technologies, perfect competition, constant returns to scale), in this paper I show that one can use profit maximization to separately identify input allocations from within-firm productivity heterogeneity for a much wider class

[5] I survey some of the other approaches used in the literature in more detail in Section 2.

[6] More formally, since labour is the only input and production is characterized by constant returns to scale, we can write $Y^j = A^j L^j$, where $j = 1, 2$ indexes products within the plant, A^j is the unit labour requirement for output j, and L^j is the quantity of labour used to produce j. This provides two equations. As long as labour is completely attributable to each production line, this provides a third restriction, $L^1 + L^2 = L$, where L is the total quantity of observed labour. Profit maximization and perfect competition imply that $P^j = \frac{w}{A^j}$ for $j = 1, 2$, which implies that $\frac{P^1}{P^2} = \frac{A^2}{A^1}$. This is the fourth condition necessary to uniquely determine (A^1, A^2, L^1, L^2) using information on (Y^1, Y^2, P^1, P^2, L).

of pricing models and production technologies. These include models of oligopolistic competition and collusion, as well as more general multiple input production technologies with increasing or decreasing returns to scale. These are new identification results that should prove useful to other researchers who wish to estimate the productivity of multi-product firms in datasets that only contain information on aggregate inputs.[7]

The identification results developed in this paper shows that one cannot ignore the demand side of an industry when engaging in productivity analysis, if one wishes to separately identify within-firm input use from within-firm productivity dispersion. In particular, I show that allowing firms to have pricing power usually requires that a product-level *demand function* be estimated, before one can identify within-firm input allocations.[8] This is because separately identifying input allocations from productivity will require information on demand elasticities, since one must be able to distinguish between variation in prices due to productivity, versus variation in prices due to differences in market power.

In my empirical application, I estimate a product-level discrete/continuous choice demand system based on Björnerstedt and Verboven (2013), under the assumption of Bertrand-Nash price competition. This approach allows for within-plant *cannibalization effects*, where individual plants account for the fact that decreasing the price of one of their product lines will tend to decrease demand for their other products. To estimate the parameters governing this demand system, I make use of a novel set of instrumental variables, which harness exogenous variation in input prices due to demand and supply shocks in *other* output markets that use similar inputs. This new identification strategy takes advantage of the detailed input price variation observed in my data, while avoiding the endogeneity problems generated by the potential correlation between individual input prices and output quality.[9]

I further show that as long as the production function satisfies some relatively standard restrictions consistent with much of the empirical work on productivity analysis, then one can uniquely determine input allocations from price, quantity, and demand elasticity data, without prior knowledge of the production function parameters. This is a useful identification result, since one can estimate within plant input allocations *before* production function parameters are estimated. As a result, one can use these estimates of plant-product level input use to directly estimate a plant-product level production function, using standard tools from the production function estimation literature.[10] This property allows me to use all of the plants in my sample to estimate the production function, rather than using a selected sample of single product plants as in De Loecker et al. (2016).[11]

[7]Examples of this type of data include the U.S. Census of Manufactures used in Foster et al. (2008), the Belgian PRODCOM data used in Dhyne et al. (2017), the Canadian manufacturing data used in Baldwin and Gu (2009), the Indian Prowess data used in Goldberg et al. (2010b) and De Loecker et al. (2016), as well as the Indian ASI data used in this paper.

[8]I also show that the restrictions on the demand system necessary for identification of within-firm input allocations are simply a variant of those required by Berry and Haile (2014) for the nonparametric identification of demand systems in differentiated product markets. As noted in their paper, these restrictions are satisfied by almost all the demand systems used in applied work.

[9]See Kugler and Verhoogen (2012) for evidence on the correlation between plant level input prices and output quality.

[10]See Ackerberg et al. (2007) for a review of this literature.

[11]Note that De Loecker et al. (2016) account for sample selection using a variant of the control function approach described in Olley and Pakes (1996), which partly restricts the nature of selection into multi-product production. Since my approach uses information on multi-product plants directly, I do not need these restrictions.

Having estimated the demand and production function parameters for the Indian machinery industry, I then examine whether import shocks lead to within-plant reallocations towards higher performing products. Using a series of simple OLS and IV regressions, where I leverage changes in Indian tariffs in the mid 2000s, as well as plausibly exogenous increases in Chinese exports, I find that increased import competition is associated with reallocations *towards* high quality goods, but *away* from high TFPQ products. This potentially counter-intuitive result is largely due to the fact that high TFPQ products are found to be lower quality, a finding that has also been uncovered in recent papers by Jaumandreu and Yin (2016) and Forlani et al. (2016), who estimate production and demand functions for single product firms. Intuitively, this means that quality is costly to produce (i.e. it requires more inputs), and therefore reallocations towards quality also generate increased costs. Decomposing the increased specialization in quality into intensive and extensive margin effects, I find that increased quality specialization is primarily driven by plants dropping low quality, high TFPQ varieties in response to increased import competition.

These findings add to the empirical literature focused on trade's impact on multi-product producers, including Baldwin and Gu (2009), Liu (2010), Goldberg et al. (2010a), Goldberg et al. (2010b), Bernard et al. (2011), Mayer et al. (2014), and Eckel et al. (2015), and Medina (2017). In particular, I show that within-firm reallocations may partly explain the positive impact of import competition on quality upgrading, a feature of the data recently explored in Amiti and Khandelwal (2013) and Medina (2017). Moreover, these results highlight that the within-firm productivity gains from trade, emphasized by theoretical models developed in Eckel and Neary (2010), Bernard et al. (2011), Mayer et al. (2014) and Mayer et al. (2016), primarily operate through a quality improvement margin in this industry, where low quality goods are more likely to be dropped in the face of increased import competition.

This paper also contributes to a growing literature on separately identifying the different sources of firm heterogeneity, including Foster et al. (2008), Forlani et al. (2016), Jaumandreu and Yin (2016), who focus on separately identifying demand versus supply-side heterogeneity, De Loecker and Warzynski (2012) and De Loecker et al. (2016) who focus on separately identifying TFP and markups, Goldberg et al. (2010a) and Goldberg et al. (2010b), who focus on identifying within-firm improvements driven by product adding and dropping, Valmari (2016) and Garcia-Marin and Voigtländer (2017) who focus on identifying within-firm TFP heterogeneity, and Hottman et al. (2016) who provides a framework for separately identifying markups, quality, and marginal costs across and within firms. The approach described in this paper provides a unified framework that identifies all of these margins of firm heterogeneity, including across and within-firm variation in TFP (both TFPQ and TFPR), quality, markups, and marginal costs, while previous approaches have only developed techniques which can uncover a subset of these different margins. One exception to this is recent work by Blum et al. (2018), who also provide a framework to examine the various margins of firm heterogeneity, although with a slightly different notion of demand heterogeneity than that used in this paper. In particular, their approach uses estimated markups to provide a local linear approximation of the slope and intercept of a firm's demand function around equilibrium price and quantity, while I estimate a global demand function, and recover the estimates of the demand shifters (quality) from there.

The paper is structured as follows. In Section 2, I outline the problems generated by multi-product firms for productivity analysis, and outline how my approach to solving these problems differs from other methods that have previously been proposed in the literature. In Section 3, I present a general model of industry demand and supply, and show how the structure of this model can be used to obtain within-firm input allocations as a function of observable demand side data. I then discuss my data in Section 4, and describe the estimation strategy I use to estimate the demand and production function in Section 5. Section 6 presents the empirical results, and I conclude in Section 7.

2 The General Productivity Estimation Problem

2.1 The "Standard" Problem: Single Product Firms

The standard approach to measuring firm-level productivity involves estimating a firm-level production function. For example, a typical study examining firm productivity will estimate a model of the form:

$$y_{it} = f(\vec{X}_{it}) + \omega_{it} \tag{1}$$

where y_{it} is log output of firm i at time t, $\vec{X}_{it}$ is an $1 \times P$ vector of P different input quantities, $f(.)$ is the log production function and ω_{it} is firm-level log TFP.[12] Note I will denote the natural log of any variable X, by its corresponding lower case letter x, e.g. $y_{it} \equiv \ln(Y_{it})$, where Y_{it} is the level of output produced by firm i.

Typically, it is assumed that the researcher has data on output and all of the relevant inputs at the firm level, while the exact form $f(.)$ takes in (1) is unknown. To obtain TFP estimates, the researcher must determine a way to estimate the shape of the production function.

The key challenge to estimating $f(.)$ is the endogeneity of input use, or, as described by Griliches and Mairesse (1995), the "transmission bias" problem. Since firms know their own productivity, input usage will tend to be correlated with ω_{it}, which means that one cannot simply estimate (1) using OLS. As a result, much of the literature on production function estimation is concerned with solutions to the transmission bias problem, with some popular approaches being the proxy-variable estimation methods described by Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg et al. (2015), as well as the dynamic panel approaches described by Arellano and Bond (1991), Blundell and Bond (1998), Blundell and Bond (2000) and Bond and Söderbom (2005).

[12] If y_{it} is measured in revenue units, ω_{it} would correspond to log TFPR, while if y_{it} is measured in quantity units, ω_{it} measures log TFPQ. Note that this distinction is not important for the general identification results discussed in this and the following section, as long as a well defined TFPR or TFPQ generating production function exists. As a result, I will simply refer to ω_{it} as TFP when discussing the general approach to identification.

2.2 Complications Generated by Multi-product Production

While the transmission bias problem inherent in estimating the production function $f(.)$ has been widely studied, with many estimation routines available to applied researchers who wish to estimate productivity, firm-level datasets containing multi-product firms have generated new problems that the researcher must deal with to obtain estimates of TFP, *even if $f(.)$ were known.* Specifically, most firm-level datasets only contain information on input use at the firm-level, even if they provide output information at the firm-product level.[13] This generates significant difficulties, since if one wishes to estimate firm-product varying TFP, one must estimate some variant of the following model

$$y_{it}^{j} = f^{j}(\vec{X}_{it}^{j}) + \omega_{it}^{j}, \tag{2}$$

where y_{it}^{j} is output of product j produced by firm i at time t, $\vec{X}_{it}^{j}$ is the $1 \times P$ vector of inputs used in the production of good j, $f^{j}(.)$ is the production function, which may vary with j, and ω_{it}^{j} is the overall productivity of good j. Unfortunately, while y_{it}^{j} is generally observed in most firm-level datasets, the researcher generally only has access to data on aggregate inputs, $\vec{X}_{it}$. As a result, *all* of the variables on the right-hand side of (2) are unobserved, making standard production function estimation infeasible.

There have been, roughly speaking, two general classes of solutions to the problems posed by multi-product firms. The first solution makes observing $\vec{X}_{it}^{j}$ unnecessary for productivity analysis, by redefining the object of interest. For example, one can instead focus on simply recovering estimates of firm-level TFP, either by creating a firm-level output index to deflate firm-level revenues, as in Eslava et al. (2004) and Smeets and Warzynski (2013), or by making use of an explicit within-firm aggregation model that allows one to estimate the transformation rates of outputs across uses within a firm, as in Balat et al. (2016). While these approaches are useful for answering questions related to across-firm variation in productivity, they by construction remain silent on within-firm productivity dispersion and allocative efficiency.

Another approach in this vein, recently pursued by Dhyne et al. (2017), examines firm as well as firm-product level productivity by estimating a *transformation* function, which is a generalization of the production function for multi-product firms which only requires information on aggregate input use. This approach allows for joint-production of outputs within a firm, which may allow for more careful analysis of the sources of economies of scope. On the other hand, the approach is not well suited for examining within-firm allocative gains, largely because the input allocation problem has already implicitly been solved in the formulation of the transformation function, and therefore cannot be used to consider changes in within-plant *specialization*.[14]

[13] Although there are some exceptions, including the data on vertically integrated Chinese steel producers used in Brandt et al. (2017) , and the Chilean data used Garcia-Marin and Voigtländer (2013, 2017).

[14] Note that results from Hall (1973) imply that transformation functions that are specified to be additively separable in inputs and ouputs, as in Dhyne et al. (2017), almost always describe *non-joint* production technologies, i.e. technologies that *cannot* be described by a series of product line specific production functions. It is extremely difficult to capture the notion of specialization in non-joint environments, since this framework rules out a standard input allocation problem. See also Appendix H, which considers the relationship between transformation functions and the approach to recovering input allocations described in this paper.

A second solution involves using information at the firm-level to construct estimates of the within-firm input allocations. While many authors simply use revenue shares to allocate inputs, including Foster et al. (2008), Atalay (2014) and Collard-Wexler and De Loecker (2015), a number of recent papers, including De Loecker et al. (2016) and Valmari (2016), have shown how one may use restrictions on the nature of the technology and market structure to recover equilibrium input allocations.[15] To see why restrictions are necessary, note that if both $\vec{X}_{it}^{j}$ and ω_{it}^{j} are unobserved in (2), then researchers face a significant identification problem, since increased output across production lines can be due to both increases in *productivity*, which are unobserved, or increases in the *quantity of inputs* going into a particular production line j, which are also unobserved. More formally, even if $f^{j}(.)$ is known *ex-ante*, then for each multi-product firm producing J_{it} products, there are $J_{it}(P+1)$ unobservables that the researcher must identify: J_{it} unknown ω_{it}^{j} terms, and $J_{it} \times P$ unknown input allocations. Unfortunately, (2) only provides, at most, J_{it} restrictions, leaving productivity and input allocations severely underidentified.

Hence, if the researcher wishes to separately identify input allocations from productivity, the researcher must impose some restrictions on the nature of production. De Loecker et al. (2016) (henceforth DGKP), have recently generated some important insights to this problem, by showing that one can separately estimate TFP and input allocations under the following key restrictions:[16]

- (DGKP 1): Aggregate inputs, $\vec{X}_{it}$, are completely attributable to each production line, i.e. $X_{it}^{j} = S_{it}^{jX} X_{it}$, where $S_{it}^{jX} \in [0,1]$ and $\sum_{j=1}^{J_{it}} S_{it}^{jX} = 1$, for any input X.
- (DGKP 2): Input shares do not vary by input, i.e. $S_{it}^{jX} = S_{it}^{j}$ for any input X.
- (DGKP 3): There is no within-firm TFP dispersion, i.e. $\omega_{it}^{j} = \omega_{it}$ for all j.
- (DGKP 4): $f^{j}(.)$ is known.

These restrictions imply that there are only $J_{it}+1$ unobservables that need to be identified by the researcher: the J_{it} input shares, S_{it}^{j}, and the single unknown TFP term, ω_{it}. More importantly, there are now enough equations to pin down the remaining unknowns, since (2) provides the J_{it} equations necessary to determine the input shares, while restrictions (DGKP 1) and (DGKP 2) generate the extra equation, $\sum_{j=1}^{J_{it}} S_{it}^{j} = 1$, necessary to pin down firm-level TFP. While this requires that $f^{j}(.)$ be known, DGKP accomplish by estimating $f^{j}(.)$ using a sample of single product firms.[17] They then use $\sum_{j=1}^{J_{it}} S_{it}^{j} = 1$ and (2) to solve for the unknown input share and firm-level TFP terms numerically.[18]

[15] De Loecker (2011) and Collard-Wexler and De Loecker (2015) also describe sufficient conditions for input shares to be given by equal weights, or revenue shares, respectively, although this is not their primary focus.

[16] My presentation of their assumptions differs slightly from their own, since their goal is to estimate within-firm input *expenditure shares*, rather than quantity shares, as expenditure shares (combined with an estimate of within-firm input prices) are the only information needed to estimate product-level markups. Note, however, that their algorithm for backing out the expenditure allocations can also be applied to back out the quantity allocations.

[17] The estimation procedure used in DGKP introduces a selection correction term that, under some further restrictions, will consistently estimate the production function parameters.

[18] Note that having $J+1$ equations to pin down $J+1$ unknowns is only a necessary condition for identification. For this to be sufficient, one would need to impose further restrictions on the production technology to guarantee that a

The key advantage of using these restrictions to identify input shares and TFP, is that they do not require any restrictions on demand or market structure. This allows researchers to investigate questions related to competition and market power, without imposing some of the (potentially strong) restrictions on demand and market structure, that are commonly used in empirical industrial organization studies.[19] For example, DGKP combine this insight with the fact that markups can be calculated using only production side data, as in De Loecker and Warzynski (2012), to show that trade liberalization in India resulted in an *increase* in markups, due to incomplete pass-through of the costs savings generated by input-market trade liberalization.[20]

Note, however, that the restriction (DGKP 3) simply rules out variation in within-firm productivity by assumption, which is precisely the sort of within-firm heterogeneity emphasized by Eckel and Neary (2010), Bernard et al. (2011), Mayer et al. (2014) and Mayer et al. (2016) that can lead to within-firm allocative efficiency gains. In this paper, I develop an alternative set of restrictions that can be used to separately identify TFP from input allocations, *even in the presence of within-firm TFP dispersion*. The approach is based on a fully specified demand and supply model of production, where I show that conditional cost-minimziation for the $J_{it} \times P$, inputs combined with J_{it} pricing first-order conditions, generate enough restrictions to separately identify the $J_{it} \times P$ input allocations from the J_{it} TFP terms.[21]

Valmari (2016) also takes advantage of pricing information to estimate within-firm input use, although only under the special case of Cobb-Douglas production technologies and monopolistic pricing. I provide a general approach to this problem, providing a series of sufficient conditions for observable prices and quantities to reveal the unobservable equilibrium input allocations chosen by a firm. The general approach allows for relatively general production technologies, as well as models of multi-product firms incorporating upward pressure on prices due to cross-product *cannibalization effects*, as well as some models of collusive pricing. In the subsequent section, I outline a general model of industry demand and supply that includes each of these particular models as special cases, and show how to use this class of models to identify within firm input allocations.

3 Model

This section describes a simple model of industry-level demand and supply, the structure of which can be used to determine within-firm input allocations as a function of commonly observed firm-level data (output quantities, prices, and aggregate inputs), and parameters to be estimated. First I describe the basic structure of the model, including the key assumptions necessary to generate a mapping between

unique solution to this system of equations exists. However, log-linear production functions, such as the Cobb-Douglas and translog production technologies, which are used in the vast majority of applied empirical work, will satisfy these restrictions.

[19] See Ackerberg et al. (2007) for a review of some of this literature.

[20] In particular, while trade liberalization led to price decreases, marginal costs fell *more* than prices, generating an increase in markups.

[21] Another way to state this is that $J_{it} \times P$ first-order conditions for inputs, combined with the J_{it} equations provided by (2), generate all of the necessary equations to pin down the $J_{it}(P+1)$ unknowns. However, since input use will generally depend on output, and hence price, it is generally easier to work with the $J_{it} \times P$ first order conditions for cost minimization conditional on output levels, and the J_t pricing first-order conditions conditional on the cost function.

input allocations and observables, that does not depend on production function parameters. I then derive this mapping, and discuss some simple examples and extensions.

3.1 Basic Environment

During each period t, a set of differentiated products, Ω_t, are sold on the market by N_t firms. Each product is produced by a particular firm i, with $\mathbb{Y}_{it} \subset \Omega_t$ denoting the *set* of products produced by firm $i = 1, 2, ..., N_t$. Each product (or variety) $j \in \mathbb{Y}_{it}$ is produced using the following production technology:

$$Y_{it}^j = \exp(\omega_{it}^j) F(\vec{X}_{it}^j), \tag{3}$$

where Y_{it}^j is total output of variety $j \in \mathbb{Y}_{it}$, $\vec{X}_{it}^j$ is the vector of inputs used in the production of j, and ω_{it}^j the log TFP, or productivity, of variety j.

To obtain an input allocation rule that *only* depends on demand-side data (prices, quantities, and demand elasticities), I require that the production function satisfy the following restrictions:

Assumption 1. *$F(.)$ is continuous and differentiable, equal to zero if any of its arguments are equal to zero, strictly increasing in all arguments, quasi-concave, and homogeneous of degree $\phi > 0$.*

Assumption 2. *The production technology differs across product lines within a firm only due to differences in ω_{it}^j, i.e. $F(.)$ does not depend on $j \in \mathbb{Y}_{it}$.*

While Assumption 1 mostly imposes standard regularity conditions on the production technology, Assumption 2 requires some discussion. While restricting the shape of the production technology to not differ across production lines within a firm may appear quite restrictive, note that in practice most production data sets only have enough observations to feasibly estimate production function parameters at the *industry* level, rather than the *industry-product level.* For example, Levinsohn and Petrin (2003) only allow production function parameters to differ at the 3-digit ISIC code, while De Loecker et al. (2016) only allow production function parameters to differ at the 2-digit NIC level. Hence, in practice, Assumption 2 simply requires that one focus on firms that only sell products within the same *industry.*

Moreover, this assumption, along with a little more structure, will imply restriction (DGKP 2), i.e. input shares do not depend on the identity of the input. This will allow me to determine an input allocation rule that *only* depends on demand side data. This shall prove useful for the purpose of estimation, since it will allow the researcher to estimate within-firm productivity dispersion in three sequential steps. *First*, estimate the demand side of the model. In the *second* step, recover the equilibrium input allocations. Finally, in the *third* step, use the estimated input allocations to estimate a firm-product level production function.

However, as I show in Section 3.3, Assumptions 1 and 2 can both be relaxed, but at the cost of generating an input allocation rule that depends on production function parameters. If production technologies were already known, this would generate no additional difficulties. However, for the purpose of estimation, this can generate significant problems, which I discuss in Section 3.3. In either case, the researcher should regard Assumption 2 simply as a sufficient condition for the input allocation rule to only depend on demand side data.

While the production technology can use an arbitrary number of inputs, I will generally assume that production may require at least one *dynamic input*, such as capital, which can only be obtained through a dynamic investment process, and requires at least one *static input*, such as materials, that can be purchased from the market each time period and has no dynamic implications. More formally:

Assumption 3. *$F(.)$ takes as an input at least one element from the set $\mathbb{M}$, where $\mathbb{M}$ is set of static inputs, which can be purchased for one-period use from the market according to some known price schedule $W_{it}^M = \boldsymbol{W}^M\left(\sum_{j\in\mathbb{Y}_{it}} M_{it}^j, A_{it}^M\right)$ for each $M \in \mathbb{M}$, where A_{it}^M is a vector of input price shifters.*

Note that Assumption 3 allows for, but does not require, input buyers having market power, as in Morlacco (2018), in that they recognize that buying more units of an input may affect the market price of that input. If producers have no market power, then $\mathbf{W}^M\left(\sum_{j\in\mathbb{Y}_{it}} M_{it}^j, A_{it}^M\right)$ is simply a constant for all levels of input purchases.

Similarly, I let $\mathbb{K}$ denote the set of dynamic inputs, which face some form of adjustment costs at the firm-level.[22] Each dynamic input $K \in \mathbb{K}$, evolves over time according to the law of motion $K_{i,t+1} = l^K\left(K_{it}, I_{it}^K, I_{i,t+1}^K\right)$, where K_{it} is the stock of dynamic input K in firm i at time t, and I_{it}^K is firm i's current investment in dynamic input K.[23] Investment, as well as upkeep of the current stock of dynamic input K, costs the firm $d^K(K_{i,t-1}, K_{i,t}, I_{it}^K)$, where the stock of the dynamic input K in the previous period is included in the dynamic cost function to allow for adjustment costs. This formulation of the cost function allows dynamic inputs to have adjustment costs at the *firm-level.* To obtain an allocation rule that only depends on demand side data, I require that there not be any adjustment costs *within* a firm. In particular, I assume that:

Assumption 4. *All inputs can be costlessly transferred across production lines within a firm.*

Note that Assumption 4 is immediately satisfied for static inputs. For dynamic inputs, Assumption 4 simply means that given a stock K_{it} of dynamic input $K \in \mathbb{K}$, each firm i can costless allocate this resource across their production lines, with K_{it}^j denoting the quantity of dynamic input $K \in \mathbb{K}$ going into production line $j \in \mathbb{Y}_{it}$.

Given these assumptions, it will be useful to distinguish between *input allocations*, which describe the allocation of inputs across uses within a firm *conditional* on the aggregate resources firm i commands, and *aggregate input vectors*, which correspond to total quantity of inputs used by the firm i in period t. Formally, I will denote input allocations by the input matrix $\mathbb{X}_{it}$, with typical element (j, X), X_{it}^j,

[22]Labour could either be a static variable, or a dynamic variable, depending on the structure of the labour market. For example, if hiring and costs are significant, then it may be more appropriate to model labour as a dynamic variable, as I do in my empirical application.

[23]Note that I *allow* investment at time $t+1$ to affect the total stock of dynamic input K at time $t+1$, but do not require that current investment affect the total stock of the dynamic input. In particular, the case of predetermined dynamic inputs simply corresponds to the law of motion where $l^K\left(K_{it}, I_{it}^K, I_{i,t+1}^K\right) = l^K\left(K_{it}, I_{it}^K, \widetilde{I}_{i,t+1}^K\right) \quad \forall I_{i,t+1}^K, \widetilde{I}_{i,t+1}^K$. Note that this restriction would be satisfied by the standard law of motion $K_{i,t+1} = I_{i,t} + \delta K_{it}$, that is often used in the literature. Note, however, that the identification results in this section do not require that the law of motion for capital take this linear form, nor do they require that K_{it} be predetermined.

denoting the total quantity of input $X \in (\mathbb{K}, \mathbb{M})$ allocated to production line $j \in \mathbb{Y}_{it}$. On the other hand, I will denote aggregate input vectors by $\vec{X}_{it} \equiv \left(\vec{K}_{it}, \vec{M}_{it}\right)$, where $\vec{K}_{it}$ is the vector dynamic input stocks owned by firm i at time t, and $\vec{M}_{it}$ is the vector of total static inputs purchased by firm i.

To ensure that input allocations are well defined, I also make the following assumption:

Assumption 5. *Aggregate inputs, $\vec{X}_{it}$, are completely attributable to each production line, i.e. $X^j_{it} = S^{jX}_{it} X_{it}$, where $S^{jX}_{it} \in [0,1]$ and $\sum_{j=1}^{J_{it}} S^{jX}_{it} = 1$, for any input X.*

Note that Assumption 5 rules out public inputs, such as a machine that can be used in more than one production processes simultaneously. This may rule out some forms of economies of scope. While this is a standard assumption in the literature on estimating production functions with multi-product firms, in Section 3.3 I relax this assumption. Under some restrictions on the form of public inputs, I show that the *share* of *effective inputs* across production lines can still be identified, although the quantity of public versus private inputs cannot. As a result, TFP estimates from multi-product firms may be "scaled-up" due to public inputs, which, under the restrictions discussed in Section 3.3, will be observationally equivalent to a common TFP shifter for multi-product firms. This echoes the discussion in De Loecker et al. (2016) that Assumption 5 is compatible with some economies of scope, as long as they are embodied in differences in TFP across firms.

I also make the following restriction on relationship between inputs and outputs:

Assumption 6. *The input sets $(\mathbb{K}, \mathbb{M})$, do not contain any products produced within the same firm, i.e. $\mathbb{Y}_{it} \notin (\mathbb{K}, \mathbb{M})$*

Assumption 6 rules out vertical integration of the production process. While this may be an interesting channel for thinking about productivity in some industries (e.g. Steel), the methods outlined in this paper do not easily generalize to this case.

The next set of assumptions describe the industry structure. I assume that each product sold on the market faces a downward sloping demand function, $Q^j_{it}(\vec{P}_t)$. Note that the demand function depends on the entire *vector* of prices charged on the market, $\vec{P}_t$, allowing for fairly general patterns of cross-product substitutability. Letting $\vec{Q}_{it}(\vec{P}_t)$ denote the vector of demand functions, I also require that the overall demand system satisfy the following restriction, described in more detail in Berry et al. (2013) and Berry and Haile (2014):

Assumption 7. *The demand system $\vec{Q}_{it}(\vec{P}_t)$ exhibits connected substitutes in prices.*

Assumption 7 is a fairly weak restriction that is satisfied by most of the demand systems used in applied work. Roughly speaking, this restriction requires that all goods be weak substitutes for each other (demand for each good j is non-decreasing in the price of all goods $k \neq j$), and some goods are strict substitutes for one another (demand for good j is strictly increasing in the price of some subset goods $m \neq j$).

Each firm then chooses their aggregate input vectors $\vec{X}_{it}$, i.e. total labour and capital, input allocations $\mathbb{X}_{it}$, i.e., the quantity of labour and capital in each production line, prices $\vec{P}_{it}$, and their investment levels, $\vec{I}_{it}$ to maximize the present discounted value of their profits, taking into account the constraint that sales cannot be greater than quantities produced, as well as the laws of motion for the dynamic inputs. The Bellman equation associated with this problem is given by:[24]

$$\begin{aligned}
V_t(\chi_{it}) = \underset{\vec{I}_{it},\mathbb{X}_{it},\vec{X}_{it},\vec{P}_{it}}{\text{Max}} & \sum_{j\in\mathbb{Y}_{it}} P^j_{it}Q^j_{it}\left(\vec{P}_t\right) - \sum_{M\in\mathbb{M}}\sum_{j\in\mathbb{Y}_{it}} \mathbf{W}^M\left(\sum_{j\in\mathbb{Y}_{it}} M^j_{it}, A^M_{it}\right) M^j_{it} \\
& - \sum_{K\in\mathbb{K}} d^K(K_{i,t-1}, K_{it}, I^K_{it}) + \beta\mathbb{E}\{V_{t+1}(\chi_{i,t+1})\,|\chi_{it}\} \\
\text{subject to:} & \qquad\qquad (4) \\
& \exp(\omega^j_{it})F(\vec{X}^j_{it}) \geq Y^j_{it} = Q^j_{it}\left(\vec{P}_t\right) \quad \forall j\in\mathbb{Y}_{it} \\
& \sum_{j\in\mathbb{Y}_{it}} X^j_{it} = X_{it} \quad \forall X\in\mathbb{K} \\
& K_{i,t+1} = l^K(K_{it}, I^K_{it}, I^K_{i,t+1}) \quad \forall K\in\mathbb{K},
\end{aligned}$$

where χ_{it} is the vector state variables, which include the set of products produced by the firm $\mathbb{Y}_{it}$, the set of products produced by *other* firms $\mathbb{Y}_{-i,t} \equiv \Omega_t \setminus \mathbb{Y}_{it}$, the vector of firm-level TFP terms, $\vec{\omega}_{it}$, the vector of firm-level input price shifters $\vec{A}_{it}$, the vector of prices charged by all other firm $P_{-i,t}$, and the vector of lagged stocks of dynamic inputs $\vec{K}_{i,t-1}$.[25] [26]

3.2 Using First-Order Conditions to Recover Input Allocations

I now show that Assumptions 1 though 7 imply a mapping from observable prices and quantities, to the unobserved input allocations chosen by a set of firms each independently solving (4).

First, note that since dynamic inputs are costlessly transferable across uses within the firm (Assumption 3), and $F(.)$ is strictly increasing in all of its inputs (Assumption 1), each firm will choose an input allocation that minimizes total static input costs *conditional on its stock of firm-level dynamic inputs*, $\vec{K}_{it}$, and some desired set of output levels $\vec{Y}_{it}$. More formally, any solution to (4) will involve an input allocation $\mathbb{X}_{it}$ that minimizes static input costs, subject to some desired output levels Y^j_{it} for each $j \in \mathbb{Y}_{it}$, and a given stock of dynamic inputs.[27] The Lagrangian for this conditional cost minimization

[24] I slightly abuse notation here, letting Y^j_{it} denote the *desired* output quantity a firm wishes to produce. In equilibrium this will always equal actual quantity produced through (3).

[25] In the special case of a *pre-determined* dynamic inputs, where $l^K\left(K_{it}, I^K_{it}, I^K_{i,t+1}\right) = l^K\left(K_{it}, I^K_{it}, \tilde{I}^K_{i,t+1}\right) \quad \forall I^K_{i,t+1}, \tilde{I}^K_{i,t+1}$, then K_{it} will also be a state variable.

[26] Note given $\left(\mathbb{X}_{it}, \vec{X}_{it}\right)$, prices are implicitly pinned down in the above formulation of the firm's problem by the constraints $\exp(\omega^j_{it})F(\vec{X}^j_{it}) \geq Q^j_{it}\left(\vec{P}_t\right)$, which hold with equality in equilibrium. Therefore, one could also write the problem as firms choosing prices and all static inputs *except for one static input* $M \in \mathbb{M}$, although this would be notationally cumbersome.

[27] Note that reallocating dynamic inputs across production lines, conditional on any level of firm-level dynamic inputs and desired output levels Y^j_{it}, will not affect a firm's current costs or future profits, except by allowing the firm to potentially purchase less static inputs by reorganizing the dynamic inputs more efficiently. Hence, one can "concentrate out" the optimal levels of static inputs in this problem, conditional on some level of dynamic inputs, by solving the conditional cost minimization sub-problem for any given value of ($\vec{K}_{it}$, Y^j_{it}) .

sub-problem is given by[28]

$$
\begin{aligned}
L = &- \sum_{M\in\mathbb{M}}\sum_{j\in\mathbb{Y}_{it}} \mathbf{W}^M\left(\sum_{j\in\mathbb{Y}_{it}} M_{it}^j, A_{it}^M\right) M_{it}^j + \sum_{j\in\mathbb{Y}_{it}} \lambda_{it}^j \left(\exp\left(\omega_{it}^j\right) F(\vec{X}_{it}^j) - Y_{it}^j\right) \\
&+ \sum_{K\in\mathbb{K}} \mu_{it}^K \left(K_{it} - \sum_{j\in\mathbb{Y}_{it}} K_{it}^j\right),
\end{aligned}
\tag{5}
$$

where λ_{it}^j is the Lagrangian multiplier for the production constraint, and μ_{it}^K is the Lagrangian multiplier for resource constraint for dynamic input $K \in \mathbb{K}$.

Letting $\nu_{it}^X = W_{it}^X + \frac{\partial \mathbf{W}^X\left(\sum_{j\in\mathbb{Y}_{it}} X_{it}^j, A_{it}^X\right)}{\partial\left(\sum_{j\in\mathbb{Y}_{it}} X_{it}^j\right)} \sum_{j\in\mathbb{Y}_{it}} X_{it}^j$ if $X \in \mathbb{M}$, and $\nu_{it}^X = \mu_{it}^X$ if $X \in \mathbb{K}$, the first-order necessary condition for any input X_{it}^j can be written as:

$$
-\nu_{it}^X + \lambda_{it}^j \exp\left(\omega_{it}^j\right) \frac{\partial F(\vec{X}_{it}^j)}{\partial X} = 0. \tag{6}
$$

To obtain a simple input allocation formula that only depends on demand side information, I make use of the following Lemma, which shows that property (DGKP 2) from De Loecker et al. (2016) follows from the previously stated assumptions:

Lemma 1. *If Assumptions 1 through 6 hold, then there exists a solution to the firm's conditional cost minimization problem satisfying* $X_{it}^j = S_{it}^j X_{it} \quad \forall X \in (\mathbb{K}, \mathbb{M})$*, where* $S_{it}^j \in [0,1]$ *and* $\sum_{j\in\mathbb{Y}_{it}} S_{it}^j = 1$

Proof. See Appendix A. □

While a formal proof of the above is provided in the Appendix for completeness sake, the intuition for this result is a relatively straightforward implication of homogeneous production technologies. Since homogeneous production technologies generate isoquants that have constant slopes along any ray from the origin, this means that a cost minimizing firm will choose constant input ratios across production lines within the firm, i.e. $\frac{X_{it}^j}{Z_{it}^j} = \frac{X_{it}^k}{Z_{it}^k}$ for each $X, Z \in (\mathbb{K}, \mathbb{M})$ and each $j, k \in \mathbb{Y}_{it}$. This immediately implies that the input shares within a particular production line j do not depend on the identity of the input, and therefore $\frac{X_{it}^j}{X_{it}} = S_{it}^j \ \forall X \in (\mathbb{K}, \mathbb{M})$.

As a result, Assumptions 1 through 6 guarantee that the vector of inputs allocated to each production line can be written as $\vec{X}_{it}^j = S_{it}^j \vec{X}_{it}$. Substituting this into (6), and then using the fact that all of the partial derivatives of a homogeneous of degree $\phi > 0$ function are homogeneous of degree $\phi - 1$, one obtains:

$$
\nu_{it}^X = \lambda_{it}^j \exp\left(\omega_{it}^j\right) \left(S_{it}^j\right)^{\phi-1} \frac{\partial F(\vec{X}_{it})}{\partial X}. \tag{7}
$$

To get rid of the unknown TFP terms in (7), divide this expression by $Y_{it}^j = \exp\left(\omega_{it}^j\right) F(S_{it}^j \vec{X}_{it}) = \exp\left(\omega_{it}^j\right) \left(S_{it}^j\right)^{\phi} F(\vec{X}_{it})$, which yields, after some minor manipulations:

[28] Formal statement of the conditional cost minimization problem is stated in Appendix A.

$$S^j_{it} = \frac{\partial F(\vec{X}_{it})}{\partial X} \frac{\lambda^j_{it} Y^j_{it}}{F(\vec{X}_{it})\nu^X_{it}}. \tag{8}$$

One can then sum (8) over all $j \in \mathbb{Y}_{it}$, yielding $\frac{\partial F(\vec{X}_{it})}{\partial X} \frac{\sum_{j\in\mathbb{Y}_{it}} \lambda^j_{it} Y^j_{it}}{F(\vec{X}_{it})\nu^X_{it}}$. Dividing (8) by this expression yields:

$$S^j_{it} = \frac{\lambda^j_{it} Y^j_{it}}{\sum_{k\in\mathbb{Y}_{it}} \lambda^k_{it} Y^k_{it}} = \frac{MC^j_{it} Y^j_{it}}{\sum_{k\in\mathbb{Y}_{it}} MC^k_{it} Y^k_{it}}, \tag{9}$$

where the second equality in (9) follows from the envelope theorem $\lambda^j_{it} = \frac{\partial C\left(\vec{K}_{it}, \vec{Y}_{it}, \vec{\omega}_{it}, \vec{A}_{it}\right)}{\partial Y^j_{it}} \equiv MC^j_{it}$, where $C\left(\vec{K}_{it}, \vec{Y}_{it}, \vec{\omega}_{it}, \vec{A}_{it}\right)$ is the cost function for static inputs, conditional on the level of dynamic inputs $\vec{K}_{it}$ and a desired level of output $\vec{Y}_{it}$.[29]

While (9) may not at first glance appear all that useful, as marginal costs are almost always unobservable in any firm-level data set, the key insight that I exploit is that many models of inter-firm competition imply a direct mapping from observable *demand-side* variables to unobservable product level conditional marginal costs. This is an old insight, originally formalized by Rosse (1970) for the case of a monopolist, who noted that one can obtain marginal costs without supply-side data using pricing first-order conditions, since a monopolist will choose prices such that marginal revenues equal marginal costs. Hence, one can use demand-side price and quantity data to estimate demand elasticities, which can then be used to back out a firm's equilibrium marginal costs from marginal revenue.

The general insight that observed demand elasticities can be used to recover marginal costs applies to a much wider class of imperfect competition models beyond that of simple monopolists - see Berry and Haile (2014). To see this in the context of the model described in this paper, note that once the input allocation problem has been solved for any potential level of dynamic inputs $\vec{K}_{it}$, and desired output levels $\vec{Y}_{it}$, one can determine the static conditional cost function and substitute this into (4), yielding the simplified firm's problem:

$$\begin{aligned} V_t(\chi_{it}) = \underset{\vec{P}_{it}, \vec{I}_{it}, \vec{K}_{it}}{\text{Max}} & \sum_{j\in\mathbb{Y}_{it}} P^j_{it} Q^j_{it}\left(\vec{P}_t\right) - C\left(\vec{K}_{it}, \vec{Q}_{it}\left(\vec{P}_t\right), \vec{\omega}_{it}, \vec{A}_{it}\right) \\ & - \sum_{K\in\mathbb{K}} d^K(K_{i,t-1}, K_{it}, I^K_{it}) + \beta\mathbb{E}\{V_{t+1}(\chi_{i,t+1})\,|\chi_{it}\} \end{aligned} \tag{10}$$

subject to:

$$K_{i,t+1} = l^K(K_{it}, I^K_{it}, I^K_{i,t+1}) \quad \forall K \in \mathbb{K},$$

where I have used the fact that quantity produced will equal quantity sold in equilibrium, i.e. $Y^j_{it} = Q^j_{it}\left(\vec{P}_t\right)$.

Taking the first-order condition for any P^j_{it} in (10) yields:

[29] More formally, $C\left(\vec{K}_{it}, \vec{Y}_{it}, \vec{\omega}_{it}, \vec{A}_{it}\right)$ is the objective function associated associated with solution to (CM) in Appendix A.

$$Q^j_{it} + \sum_{k \in \mathbb{Y}_{it}} \frac{\partial Q^k_{it}}{\partial P^j_{it}} \left(P^k_{it} - MC^k_{it} \right) = 0. \tag{11}$$

One can then stack the $J_t \equiv |\Omega_t|$ first order conditions defined by (11), which, in matrix notation, defines the following system of equations:

$$\vec{Q}_t + \Delta_t \left(\vec{P}_t - \vec{MC}_t \right) = 0, \tag{12}$$

where $\Delta_t = \mathbb{O}_t \circ \partial_t$, with ∂_t corresponding to an $J_t \times J_t$ matrix of demand derivatives, with typical element (j,k) equal to $\frac{\partial Q^j_{it}}{\partial P^k_{it}}$, and $\mathbb{O}_t$ being the *ownership matrix*, with element (j,k) equal to 1 if product j and k are both produced by the same firm, and equal to zero otherwise.

Note that one can use (12) to solve for the equilibrium marginal costs as a function of quantities produced, prices, and demand derivatives, by pre multiplying by Δ_t^{-1}, yielding:

$$\vec{MC}_t = g\left(\vec{Q}_t, \vec{P}_t, \partial_t, \mathbb{O}_t\right) = \Delta_t^{-1}\vec{Q}_t + \vec{P}_t. \tag{13}$$

For the marginal cost inversion described by (13) to be valid, Δ_t needs to be invertible. Berry and Haile (2014) have recently shown that Δ_t will be invertible under fairly general conditions. In particular, as long as Assumption 7 holds, a relatively weak restriction that is satisfied by many of the demand systems used in applied work, then Δ_t will be invertible.[30]

An important point worth emphasizing is that as long as Assumption 7 holds, Δ_t is invertible for *any* ownership matrix $\mathbb{O}_t$. This means that the this marginal cost inversion is also possible under imperfect competition with some forms of *collusion*. In particular, as noted by Nevo (1998), collusion between firms can be thought of as *joint-profit maximization*, i.e. legally separate firms choosing their prices *as if* they were one multi-product firm.[31] As long as the set of colluding firms fully internalizes their pricing decisions across products, then this simply corresponds to an alternative ownership matrix, $\mathbb{O}'_t$, than that observed in the data. For example, *full market collusion* simply corresponds to the case where $\mathbb{O}'_t$ is a $J_t \times J_t$ matrix of ones, implying that all prices are chosen to internalize all possible cannibalization effects.

Combining (13) and (9) yields the following Theorem:

Theorem 1. *As long as the cost minimizing input allocation is unique, then Assumptions 1 though 7 imply that if each firm chooses prices, input quantities, and investment to solve (4), the share any input* $X \in (\mathbb{K}, \mathbb{M})$ *going into production line* $j \in \mathbb{Y}_{it}$ *satisfies* $S^j_{it} = \frac{g^j_{it}\left(\vec{Q}_t, \vec{P}_t, \partial_t, \mathbb{O}_t\right) Y^j_{it}}{\sum_{j \in \mathbb{Y}_{it}} g^k_{it}\left(\vec{Q}_t, \vec{P}_t, \partial_t, \mathbb{O}_t\right) Y^k_{it}}$, *where each* $g^j_{it}(.)$ *is a known function of prices, quantities, demand derivatives, and the ownership matrix.*

Proof. Follows from the discussion in text and equations (9) and (13), which will always hold if there is a unique solution to the firm's input allocation problem, as per through Lemma 1. Note that uniqueness

[30] Roughly speaking, Assumption 7 is satisfied if all goods are weak substitutes for each other (demand for each good j is non-decreasing in the price of all goods $k \neq j$), and some goods are strict substitutes for one another (demand for good j strictly increases if the price of good k rises). See Berry et al. (2013) and Berry and Haile (2014) for more details.

[31] Note that this form of collusion requires that firms only cooperate in output markets, i.e. through pricing first-order conditions, *not* through input markets, i.e. firm's cannot engage in labour or capital sharing.

of the input allocation problem is guaranteed for standard homogenous production functions such as Cobb-Douglas and CES, as it is straightforward to verify that (6) implies that $X_{it}^j = S_{it}^j X_{it} \; \forall j \in \mathbb{Y}_{it}$. $\square$

3.2.1 Theorem 1 In Practice: Examples

While Theorem 1 is formulated in fairly general terms, one can obtain some useful insights by applying it to some special cases that are often used in the applied literature. For example, models of perfect competition with price-taking firms imply that all firms choose quantities such that $P_{it}^j = MC_{it}^j$. Substituting this into (9) yields: $S_{it}^j = \frac{P_{it}^j Y_{it}^j}{\sum_{k \in \mathbb{Y}_{it}} P_{it}^k Y_{it}^k}$, i.e the optimal input allocations are simply revenue shares.

While perfect competition is a highly unrealistic assumption for a model where firms produce differentiated varieties, note that revenue share input allocation rule will also hold whenever multiplicative markups $\mu_{it}^j \equiv \frac{P_{it}^j}{MC_{it}^j}$, are constant within a firm. Equation (9) then implies:

$$S_{it}^j = \frac{\frac{P_{it}^j}{\mu_{it}} Y_{it}^j}{\sum_{k \in \mathbb{Y}_{it}} \frac{P_{it}^k}{\mu_{it}} Y_{it}^k} = \frac{P_{it}^j Y_{it}^j}{\sum_{k \in \mathbb{Y}_{it}} P_{it}^k Y_{it}^k} \quad \text{if } \mu_{it}^j = \mu_{it} \;\; \forall j \in \mathbb{Y}_{it}. \tag{14}$$

The fact that constant within-firm markups allows one to use revenue shares to allocate inputs, was previously noted by Collard-Wexler and De Loecker (2015), although their derivation of this property only proceeded after making the restriction $\omega_{it}^j = \omega_{it} \; \forall j \in \mathbb{Y}_{it}$. Note that, as the derivation of (9) makes clear, a revenue share input allocation rule will be valid with constant within-firm markups *even if* there is within-firm TFP heterogeneity.

Interestingly, it turns out that equilibrium markups will be constant within a firm if variety-level demand is CES, i.e. $Q_{it}^j(\vec{P}_t) = \frac{\left(P_{it}^j\right)^{-\sigma}}{\sum_{l=1}^{N_t} \sum_{k \in \mathbb{Y}_{lt}} \left(P_{lt}^k\right)^{1-\sigma}}$, *even if firms are not atomistic*, a property of CES demand systems previously noted by Feenstra and Ma (2007). In particular, it is straightforward to show the pricing first-order conditions (11) for this demand system imply the following marginal cost inversion:[32]

$$MC_{it}^j = \frac{P_{it}^j}{\left(1 + \frac{1}{(\sigma-1)(1-RS_{it})}\right)} \quad \text{if } Q_{it}^j(\vec{P}_t) \text{ is CES,} \tag{15}$$

where $RS_{it} \equiv \frac{\sum_{j \in \mathbb{Y}_{it}} P_{it}^j Q_{it}^j}{\sum_{l=1}^{N_t} \sum_{k \in \mathbb{Y}_{lt}} P_{lt}^k Q_{lt}^k}$ is the revenue share of *all* products produced by firm i. Note that as long as the number of competitors in the market is finite, so $RS_{it} > 0$ for at least some i, there will be heterogeneous markups *across* firms, but since RS_{it} does not vary with $j \in \mathbb{Y}_{it}$, markups will constant within firm, and therefore according to (14) input allocations will follows a simple revenue share rule.[33] Hence, in practice, if a researcher believes that demand for a particular industry can be

[32]Note that this marginal cost inversion also holds if the product produced have heterogeneous qualities, i.e. $Q_{it}^j(\vec{P}_t) = \frac{\left(P_{it}^j\right)^{-\sigma} \eta_{it}^j}{\sum_{l=1}^{N_t} \sum_{k \in \mathbb{Y}_{lt}} \eta_{lt}^k \left(P_{lt}^k\right)^{1-\sigma}}$, where η_{it}^j is a variety specific demand shifter.

[33]Interestingly, note that this revenue sharing rule is valid for *any* ownership matrix $\mathbb{O}_t$. This means that revenue share allocations can be used with CES demand functions, *even if firms are colluding*, and the *sets of colluding firms are unknown*.

well approximated by CES, then no other information, besides revenue at the firm-product level, is necessary to estimate input allocations for multi-product firms.

Note, however, that CES demand for all products is a very strong restriction. If demand is non-CES, the allocation rule described in Theorem 1 will depend on price elasticities, which will have to be estimated. For example, if there is a logit demand system for each product, $Q^j_{it}(\vec{P}_t) = \frac{\exp\left(\eta^j_{it} - \alpha P^j_{it}\right)}{\sum_{l=1}^{N_t} \sum_{k \in \mathbb{Y}_{lt}} \exp\left(\eta^k_{lt} - \alpha P^k_{lt}\right)}$, where η^j_{it} can be interpreted as product-specific quality, then (11) implies the following marginal cost inversion:

$$MC^j_{it} = P^j_{it} - \frac{1}{\alpha \left(1 - QS_{it}\right)} \quad \text{if } Q^j_{it}(\vec{P}_t) \text{ is Logit,} \tag{16}$$

where $QS_{it} \equiv \frac{\sum_{j \in \mathbb{Y}_{it}} Q^j_{it}}{\sum_{l=1}^{N_t} \sum_{k \in \mathbb{Y}_{lt}} Q^k_{lt}}$ is the *quantity-share* of firm i.

Substituting (16) into (9), yields the following input allocation rule for simple logit demand systems:

$$S^j_{it} = \frac{\left(P^j_{it} - \frac{1}{\alpha(1-QS_{it})}\right) Y^j_{it}}{\sum_{k \in \mathbb{Y}_{it}} \left(P^k_{it} - \frac{1}{\alpha(1-QS_{it})}\right) Y^k_{it}} \quad \text{if } Q^j_{it}(\vec{P}_t) \text{ is Logit.} \tag{17}$$

Note that the allocation rule (17) depends on the demand parameter α, which governs the own and cross-product elasticities. Hence, to determine the input allocations from price and quantity data, the demand parameter α must first be estimated.

In practice, researchers may wish to specify richer demand systems, such as the nested logit model of McFadden (1978), or the random coefficients demand system described in Berry et al. (1995). These will require estimation of more demand side parameters to determine the exact functional form (13) takes. Whether the added computational burden associated with the estimation of richer demand systems is appropriate will vary with the particular application a researcher has in mind. Nevertheless, note that as long as the demand system satisfies Assumption 7, which is the case for nested logit as well as random coefficient models, an appropriate inversion mapping will exist.

3.3 Discussion and Extensions

The key significance of Theorem 1 is that under the maintained assumptions, demand side information (prices and quantities), can be used to infer the allocations of inputs across production lines, as long as demand derivatives and ownership structures are known. This immediately suggests a straightforward strategy for estimating TFP in a dataset with multiproduct firms when prices and quantities are observed. First, use the price and quantity data to estimate the shape of the demand function. This is a standard problem in industrial organization, where a number of standard techniques have been developed just for this purpose. Having estimated the demand function, the researcher can then obtain estimates of demand derivatives at the product level, and then apply Theorem 1 to obtain the allocations of inputs across production lines within a firm. This provides the researcher with estimates of input use at the firm product level, which can then be used for the purpose of production function estimation, using standard techniques from the literature.

Before describing this approach in detail, which I turn to in Section 5, it is useful to examine the role played by some of the stronger assumptions used in Theorem 1, and the cost of relaxing these restrictions.

3.3.1 General Function Forms

To examine the role the functional form restrictions in Assumption 1 and 2 play, suppose instead that $Y_{it}^j = \exp\left(\omega_{it}^j\right) F^j(\vec{X}_{it}^j)$, i.e. $F^j(.)$ *does* depend on j, and, in principal, is neither homogeneous or quasi-concave. One can then modify (6) to allow the production function to vary with j, and then divide this expression by $Y_{it}^j = \exp\left(\omega_{it}^j\right) F^j(\vec{X}_{it}^j)$ and rearrange, yielding:

$$X_{it}^j = \frac{\theta^j\left(\vec{X}_{it}^j\right)\lambda_{it}^j Y_{it}^j}{\nu_{it}^X}, \tag{18}$$

where $\theta^j\left(\vec{X}_{it}^j\right) \equiv \frac{\partial F^j\left(\vec{X}_{it}^j\right)}{\partial X}\frac{X_{it}^j}{F^j\left(\vec{X}_{it}^j\right)}$ is the *output elasticity* for input $X \in (\mathbb{K}, \mathbb{M})$. Summing (18) over all $j \in \mathbb{Y}_{it}$, and dividing (18) by this new expression, yields, after using the envelope theorem and marginal cost inversion results described in Section 3.2:

$$X_{it}^j = \frac{\theta^j\left(\vec{X}_{it}^j\right) g_{it}^j\left(\vec{Q}_t, \vec{P}_t, \partial_t, \mathbb{O}_t\right) Y_{it}^j}{\sum_{k\in\mathbb{Y}_{it}}\theta^k\left(\vec{X}_{it}^k\right) g_{it}^k\left(\vec{Q}_t, \vec{P}_t, \partial_t, \mathbb{O}_t\right) Y_{it}^k}\sum_{j\in\mathbb{Y}_{it}} X_{it}^j. \tag{19}$$

Note that equilibrium inputs shares, if they exist, will form a *fixed point* of the system of equations described by (19). If such a solution exists, note that this mapping *does not depend on the unobservable TFP terms*, ω_{it}^j, implying that that input shares are separately identified from TFP if F^j is known, *even if Assumptions 1 and 2 do not hold.*[34]

Note, however, that while Assumptions 1 and 2 are unnecessary if $F^j(.)$ is known, in general *the production function must also be estimated.* This severely restricts the practicality of (19) for applied work, since, if one wishes to use multi-product firms to estimate the production function, *one would have to already know the production technology to determine their input allocations.*[35] Moreover, in practice,

[34]Note that a fixed-point is guaranteed to exist in (19) for the case of heterogeneous Cobb-Douglas production functions, i.e. $F^j\left(\vec{X}_{it}^j\right) = \prod_{X\in(\mathbb{K},\mathbb{M})}\left(X_{it}^j\right)^{\beta_X^j}$. Since output elasticities are constant for each input with Cobb-Douglas, i.e. $\theta^j\left(\vec{X}_{it}^j\right) = \beta_X^j$, (19) simplifies to:

$$X_{it}^j = \frac{\beta_X^j g_{it}^j\left(\vec{Q}_t, \vec{P}_t, \partial_t, \mathbb{O}_t\right) Y_{it}^j}{\sum_{k\in\mathbb{Y}_{it}}\beta_X^k g_{it}^k\left(\vec{Q}_t, \vec{P}_t, \partial_t, \mathbb{O}_t\right) Y_{it}^k}\sum_{j\in\mathbb{Y}_{it}} X_{it}^j, \tag{20}$$

i.e. X_{it}^j will be known as long as the various production function parameters, demand elasticities, ownership structures, and $\sum_{j\in\mathbb{Y}_{it}} X_{it}^j$ are also known. Imposing Assumption 5 on top of this, then (20) implies that input shares $S_{it}^{jX} \equiv \frac{X_{it}^j}{X_{it}}$, will generally depend on the identity of the input, i.e. $S_{it}^{jX} \neq S_{it}^{jZ}$, which means that Assumption (2) will often be necessary for restriction (DGKP 2) to hold.

[35]Note that there exist approaches to estimation that can deal with this circularity problem. For example, one could simply estimate the production function parameters using single product firms, as in De Loecker et al. (2016), although

most of the applied literature estimates production function by *industry* rather than by *product*, largely due to the fact that most datasets do not have sufficient observations to calculate production functions at the product level *even if input allocations were known*. Hence, for the purposes of most applied work, I regard Assumptions 1 and 2 as relatively weak, in the sense that most work in this area implicitly makes such assumptions, and moreover, fairly reasonable from a pragmatic point of view, as they can allow the researcher to include multi-product producers in a straightforward and transparent manner.

3.3.2 Within-Firm Input Price Variation and Costless Input Transferability

While (19) shows that Assumptions 1 and 2 are not crucial for separately identifying TFP from input allocations, Assumptions 3, 4, 6 and 7 are of first-order importance. In particular, Assumptions 3 and 7 are useful because they imply that static first-order conditions in prices can be used to solve for the marginal costs, while the costless transferability of resources across production lines (Assumption 4) and no vertically integrated production (Assumption 6) guarantee that the *shadow cost* of each resource, ν_{it}^X, is constant across production lines, and hence, cancels out in (19). In general, violation of these assumptions would generate an input allocation rule that depends on further unobservables, such as the allocation of dynamic inputs across production lines in the *previous* period, making identification of input allocations a much more difficult task.[36]

Note, however, that the restriction that static inputs have the *same* price across production lines, W_{it}^M, which is implicit in the formulation of Assumption 3, can be easily be relaxed as long as firms only operate Cobb-Douglas technologies, and do not have market power in input markets, i.e. W_{it}^M does *not* depend on M_{it}. In particular, suppose that input prices *differ* by production line $j \in \mathbb{Y}_{it}$, with W_{it}^{jM} denoting the constant price of static input $M \in \mathbb{M}$ for production line j. This will change the first-order conditions for static inputs $M \in \mathbb{M}$ in (6) to $W_{it}^{jM} - \lambda_{it}^j \exp(\omega_{it}^j) \frac{\partial F(\vec{X}_{it}^j)}{\partial M} = 0$. It is straightforward to show that these FOCs imply:[37]

$$\frac{M_{it}^j}{M_{it}} = \frac{\frac{\lambda_{it}^j Y_{it}^j}{W_{it}^{jM}}}{\sum_{k \in \mathbb{Y}_{it}} \frac{\lambda_{it}^k Y_{it}^k}{W_{it}^{kM}}}. \tag{21}$$

this approach requires potentially strong restrictions on the process governing selection into producing multiple products, and will be less efficient as one will have to drop a large numbers of firms for estimation. Moreover, GMM based estimators only require a set of orthogonality conditions with some set of unobservables (E.g. TFP), and these unobservables need only be expressed as a function of data and parameters to be estimated. Hence, a GMM estimator based on orthogonality of some instruments to TFP, could be used after substituting (19) into $Y_{it}^j = \exp\left(\omega_{it}^j\right) F^j(\vec{X}_{it}^j)$, which would allow the researcher to express TFP as a function of parameters to be estimated and data, as Valmari (2016) does in the case of heterogeneous Cobb-Douglas technologies. Note, however, that this approach may suffer from identification problems, since the model would be inherently non-linear, and therefore verification of the *asymptotic identification conditions*, i.e. that the parameters to be estimated are *uniquely* pinned down by the assumed moment conditions, would be quite difficult to verify in practice.

[36] In principal, if one had a sufficient number of firms who switched from being single product to multi-product, one might be able to identify these adjustment costs as the lagged quantity of dynamic inputs would be observable when they were a single product firm. I leave this problem for future research

[37] One obtains this expression by dividing the new FOC by $Y_{it}^j = \exp(\omega_{it}^j) \prod_{X \in (\mathbb{K},\mathbb{M})} \left(X_{it}^j\right)^{\beta_X}$, yielding $M_{it}^j = \frac{\beta_M \lambda_{it}^j Y_{it}^j}{W_{it}^{jX}}$. One then sums this expression over all $j \in \mathbb{Y}_{it}$, dividing the previous expression by this new sum.

One could then account for these differences in input prices across production lines using the modified static input share formula described by (21). Note, however, that input price variation across production lines within a plant *is generally not observable*. As a result, simply using the input share formula (9) to allocate static inputs across production lines, rather than (21), will likely allocate too many inputs to high-cost production lines, and too few inputs to low productivity production lines, which could generate misleading estimates of within-firm productivity dispersion.

While few datasets contain information on input prices by product lines, many datasets, including the one used in this paper, have information on *plant-specific* input prices. As a result, one can attempt to deal with these potential biases by estimating within-firm input price dispersion using information on *output quality*, and the input prices charged to single product firms. In particular, if one is willing to make the assumption that across-product input price variation is primarily driven by output quality differences and geographic location, as in De Loecker et al. (2016), then one can simply *predict* input prices by estimating the following model on a subset of single-product firms:[38]

$$W_{it}^{M} = G_t(\text{Location}_{it}, \eta_{it}) + \epsilon_{it}, \tag{22}$$

where W_{it}^{M} is (observed) input prices paid by single product firm i, Location_{it} is the local market that single product firm i operates within (e.g. state or county), while η_{it} is an estimate of product specific input quality for single product plant i, and ϵ_{it} is a mean zero error term, capturing approximation error in $G_t(.)$, as well as other factors that determine input prices across products.

While output quality, η_{it}, is generally unobserved, De Loecker et al. (2016) deal with this by noting that many industrial organization models imply a direct mapping from output prices, market shares, and product codes, to unobserved product-specific quality. Thus, one can simply substitute $\eta_{it} = H_t(P_{it}, S_{it}, \text{Code}_{it})$, where S_{it} denotes the market share of firm i and Code_{it} denotes the product code of the output produced by firm i, into $G_t(.)$, making estimation of (22) feasible.[39] Letting $\widehat{G}_t\left(\text{Location}_{it}, P_{it}, S_{it}, Code_{it}\right)$ denote the predicted values from this regression, one can then predict input price dispersion within multi-product plants using:

$$\widehat{W}_{it}^{jM} = \widehat{G}_t\left(\text{Location}_{it}, P_{it}^{j}, S_{it}^{j}, Code_{it}^{j}\right), \tag{23}$$

where $\widehat{W}_{it}^{jM}$ is the predicted input price of product $j \in \mathbb{Y}_{it}$, and P_{it}^{j}, S_{it}^{j}, and $Code_{it}^{j}$ denote the prices, market shares, and product codes of product j, respectively. Direct substitution of (23) into (21) yields a modified static input inversion rule that accounts for within-firm input price variation. Note one can follow this approach for *any* static input where input prices are observed. Hence, if one is willing to assume labour is a static input, one can also estimate within-firm wage variation using the method described above, adjusting the labour input shares to account for the estimated differences in labour

[38] Since De Loecker et al. (2016) do not actually observe input prices for *any* firms, they do not directly estimate a pricing model as in (22), but rather include $G_t(\text{Location}_{it}, \eta_{it})$ as a control function in their production function estimation algorithm.

[39] An alternative approach, pursued in Section 6 of this paper, is to directly estimate η_{it} using the structure of the demand system. For details on the demand inversion used to estimate η_{it}, see Section 5.

quality across production lines.

Note that this solution to unobserved within-firm input price dispersion is far from perfect. In particular, (22) imposes the very strong restriction that input price variation is primarily driven by output quality and location, rather than differences in *input efficiency*, i.e. the ability of an input to produce quantities of output in shorter amounts of time or with less physical units, or firm-specific "wedges", such as unobserved subsidies or taxes, that have been emphasized in some of the misallocation literature (E.g. Restuccia and Rogerson (2008), and Hsieh and Klenow (2009)). Moreover, input price schedules need to be the same for multi-product and single-product firms, which may not be the case if multi-product firms use very different inputs, compared to single product firms. Nevertheless, even with these limitations in mind, since within-firm input price variation may potentially lead to biases in the estimation of within firm price dispersion, the approach described above provides a straightforward method to verify whether ignoring within-firm input price dispersion may be driving one's results.[40]

3.3.3 Public Inputs

Finally, it is worth considering whether Assumption 5 is appropriate, i.e. all inputs are attributable to each production line, so that $X_{it}^j = S_{it}^{jX} X_{it}$. While this assumption is made implicitly by basically all the literature that uses input allocation rules to deal with multi-product firms, note that it is restrictive in the sense that it rules out public inputs, which may generate economies of scope. Note, however, that a variant of the identification result discussed in the main text carries through, even with public inputs.

Suppose that the use of some input X_{it} within a firm can be divided into a *public* or *common* component, X_{it}^C and a *rivalrous* component, X_{it}^R. Rivalrous inputs X_{it}^R can only be allocated to a single production line, as in Assumption 5, so that $X_{it}^{jR} = S_{it}^{jXR} X_{it}^R$ with $\sum_{j\in\mathbb{Y}_{it}} S_{it}^{jXR} = 1$. Common inputs are allocated to *every* production line automatically. Hence, the quantity of *effective* inputs allocated to each production lines is given by $X_{it}^j = X_{it}^C + X_{it}^{jR}$.

Suppose further that the public component of X_{it} is a *constant fraction* of total inputs owned by the firm, i.e. $X_{it}^C = \kappa^X X_{it}$ and $X_{it}^R = \left(1-\kappa^X\right) X_{it}$, and the production technology is Cobb-Douglas, so that $F\left(\vec{X}_{it}^j\right) = \prod_{X\in(\mathbb{K},\mathbb{M})} \left(X_{it}^j\right)^{\beta_X}$. Firms will then allocate their private inputs X_{it}^R across production lines to minimize static production costs conditional on aggregate dynamic inputs. It is straightforward to show, following the derivation in the text, that this slight modification implies that *effective inputs* will satisfy the following:

$$X_{it}^j = \frac{g_{it}^j\left(.\right) Y_{it}^j}{\sum_{k\in\mathbb{Y}_{it}} g_{it}^k\left(.\right) Y_{it}^k} \sum_{j\in\mathbb{Y}_{it}} X_{it}^j = \frac{g_{it}^j(.) Y_{it}^j}{\sum_{k\in\mathbb{Y}_{it}} g_{it}^k(.) Y_{it}^k} \left(1 + \left(J_{it}-1\right)\kappa^X\right) X_{it}, \tag{24}$$

where $J_{it} \equiv |\mathbb{Y}_{it}|$ is the number of products produced by firm i.[41]

While the fractions of public inputs, κ^X, are unobservable, and hence the *level* of input useage will

[40] I consider this approach as a robustness check in Section 6.

[41] Implicitly, this derivation assumes that rivalrous inputs are chosen to be non-negative, and this constraint is not binding, i.e. the firm does not wish to move common inputs from one production line to another.

not be identified in this framework, note that under Cobb-Douglas, (24) implies that the unobservable component of these input allocations, $(J_{it}-1)\,\kappa^X$, are observationally equivalent to a TFP shifter received by multiproduct firms. To see this, substitute (24) into the production function, yielding:

$$Y_{it}^j = \exp(\omega_{it}^j + SC_{it}) \prod_{X\in(\mathbb{K},\mathbb{M})} \left(\widehat{X}_{it}^j\right)^{\beta_X}, \tag{25}$$

where $\widehat{X}_{it}^j$ is the level of input useage obtained from Theorem 1, and SC_{it} is the *economies of scope* shifter, given by $SC_{it} \equiv \sum_{X\in(\mathbb{K},\mathbb{M})} \beta_X \ln\left(1 + (J_{it}-1)\,\kappa^X\right)$.

Hence, when public inputs take this form, they are observationally equivalent to TFP shifters that depend on the number of products produced by a firm. As a result, one can deal with the complications introduced by public inputs by controlling for the number of products in the production function estimation routine.[42]

Note, however, that under these assumptions the production function residual, $\widehat{\omega}_{it}^j$ will be composed of both a "pure" TFP component and an economy of scope shifter, SC_{it}, i.e. $\widehat{\omega}_{it}^j = \omega_{it}^j + SC_{it}$. This means that multi-product firms may have higher measured TFP either due to *selection*, i.e. firms with high ω_{it}^j terms are more likely to produce many products, as emphasized by Bernard et al. (2010) and Mayer et al. (2014), or due to economies of scope, i.e. multi-product production scales up input effectiveness due to public inputs. Since I do not know of a clean source of exogenous variation that would randomly allocate plants across multi-product firm status, I do not attempt to separately identify these two different mechanisms which could potentially explain a "multi-product production premium." Instead, I simply note that both mechanisms can be present in my data, and can lead to higher measured TFP in multi-product firms.

4 Data

The primary data set used in this paper comes from the 2000-2007 Indian Annual Survey of Industries (ASI), provided by by the Indian Ministry of Statistics.[43] The sample frame for the survey is all manufacturing plants in India that employ more than 10 workers. Plants with more than 100 workers ("census" firms) are surveyed every year, while smaller plants are randomly sampled each year. The data contains consistent plant-level identifiers across years for both census and non-census plants, allowing me to construct plant-level panels for both types of firms.[44] As described in Martin et al. (2017), the panel

[42] More generally, one could allow the fraction of public inputs, κ^X, to depend on the *sets* of outputs produced by the firm, in which case one would have to include product set fixed effects.

[43] Years in the ASI are recorded from April 1 to March 31. While the Ministry of Statistics refers to years by the end year, I will refer to years by the start year since the majority of production time takes place in that year. I follow this convention when matching the data to other datasets (e.g. trade data).

[44] Since the unit of observation is a plant-item, rather than a firm-item, I will consider separate plants as separate firms in my empirical analysis, and generally use these terms interchangeably. This may be problematic for my empirical strategy, as multi-plant firms may choose prices in a coordinated way that would violate the Bertrand-Nash assumption at the *plant-level* assumed by my model. Roughly speaking, this will only matter if within-firm and across-plant cannibalization effects are large. Note, however, that many of the observations are likely to approximate the decision making unit, as firms have the option of filing a *joint return* to the census for all of their factories located within the same state. 13% of

data is fairly high quality, and covers a much larger subset of Indian producers than other comparable datasets for the country, such as Prowess.

I focus on a single industry in my empirical application, Machinery, Equipment, and Parts, the details of which are described in Appendix B. I focus on this industry for two reasons. *First*, I wish to focus on an industry where it is appropriate to think of inputs as being directly allocated to different production lines, as per Assumption 5. Unfortunately, this is not the case for two of India's largest industries, sugar and textiles, as many plants in these industries produce groups of products that are *by-products*. In particular, most refined sugar producers also produce molasses, which is generated by the refining process, while many cotton producers also produce cotton waste, a by-product of cotton production that is often resold for further production purposes. Unfortunately, this means that Theorem 1 is unlikely to apply for these firms.[45] Machinery manufacturing, on the other hand, does not generate many by-products that are also sold on the market by the same plant, making Theorem 1 more likely to apply.

Second, I wish to focus on industries where there are many multi-product firms, but the output sets produced by a firm are not driven by vertical integration, as per Assumption 6. If some of the outputs produced by a firm are also inputs in a vertically integrated production line, then Theorem 1 will not apply, as the shadow value of inputs will vary depending on where they are being allocated along the vertical production line.[46] Unfortunately, this is not the case for many other industries. After constructing an input-output table, the details of which I describe in Appendix B, I find that more than 50% of the observed revenue in industries such as steel, food, and synthetic textiles, are produced by multi-product firms that are *potentially vertically integrated*, in the sense that one of their outputs is likely to be an input for another of their production lines. While there are many firms in the Machinery, Equipment, and Parts industry that also produce product sets that may indicate vertical integration, this is less of a problem than in other industries, with just over 20% of firm-year observations belonging to firms that are potentially vertically integrated (See Table 2).[47]

They key variables used in this study are described in Tables 1 and 2. Each plant lists the revenues and quantities produced and sold for up to ten different products produced within the plant. Associated with each product entry is a 5-digit ASI Commodity Classification, or ASICC, code, with just over 1000

all the large census plant-year observations within the 2001-2008 ASI panel reported these joint returns. Hence, for plants to approximate firms, I implicitly assume that the firms that operate plants in multiple states decentralize the pricing decision to local managers, or that with-firm, across-plant cannibalization effects are small enough to be safely ignored. Note that across state cannibalization effects could be safely ignored if I defined the market at the state, rather than country level. However, since many plants may ship goods across state borders, I define the market at the country-year level, rather than the state-year level, to allow for price competition across states.

[45]Note that by-products may be thought of as an extreme case of public inputs (i.e. all inputs into cotton are automatically allocated to cotton waste), which my approach can be extended to handle, as discussed in Section 3.3. Note, however, that this approach requires that public inputs enter all production lines symmetrically, which is problematic if a multi-product firm also produces other outputs that are not by-products, e.g. jute.

[46]For example, labour allocated to a downstream production line will be less costly (in terms of shadow values), since if it was instead allocated to an upstream production line, it would also be indirectly producing downstream output. This means that input allocation rules will, in general, depend on the details of the vertically integrated production process, which is beyond the scope of this paper.

[47]I drop potentially vertically integrated plants whenever I require information on the estimated within-plant input allocations, since these are unlikely to be measured correctly.

unique item codes belonging to the Machinery, Parts, and Equipment industry.[48] Each ASICC code is associated with a particular unit of quantity, such as kilograms, tonnes, or units sold, which allows one to use the information on revenues and quantities to construct a within product code consistent unit price, which I take to be item-level prices.[49] While each item produced by a plant is assigned a product code, approximately 12 % of census plant-year observations report multiple entries for the same product code. These are, according to the ASI documentation, not to be regarded as duplicates, indicating that plants also report separate product lines within a 5-digit ASICC code as well. In my empirical analysis, I consider each *entry* as a separate variety of the same general product class, rather than aggregating to the 5-digit ASICC code level.

Table 1: Plant-Product-Year Summary Statistics: Machinery, Equipment, and Parts

Variable	Obs	Mean	Std. Dev.	Min	Max	Median
Log Revenue $\left(r_{it}^{j}\right)$	64109	15.96	2.59	1.39	25.14	15.97
Log Quantity Sold $\left(q_{it}^{j}\right)$	64109	7.69	4.19	-6.91	24.17	7.5
Log Prices $\left(p_{it}^{j}\right)$	64109	8.07	3.58	-3	21.27	8.05
Log Quantity Produced $\left(y_{it}^{j}\right)$	64109	7.83	4.2	-3.44	24.19	7.69
Multi-product	64109	.76	.43	0	1	1
Multi-product × Single Industry	64109	.62	.48	0	1	1
Multi-product × Single Industry× No VI	64109	.43	.5	0	1	0

Notes: Multi-Product, Single Industry, and No VI are all dummy variables. Single industry refers to products produced by plants that only produce products belonging to ASICC codes 74-78. No VI refers to plants that do not produce output sets that I classify as potentially vertically integrated, using information on input use by single product plants. See Appendix B for more details.

For the purpose of estimation, I take the inputs of the production function to be labour, L_{it}, capital, K_{it}, and materials, M_{it}, as is standard in the literature. I then measure labour input L_{it} by the number of man-days worked, and capital, K_{it}, as net-value of fixed assets deflated by the yearly capital deflator used in Allcott et al. (2016).[50] Since I observe information on the price and quantity of various inputs at the 5-digit ASICC code, I generate a Cobb-Douglas quantity index to measure material inputs, M_{it}, that is not subject to the input price bias discussed by De Loecker and Goldberg (2014) and De Loecker

[48]These product codes correspond to all product codes that belong to the 2-digit ASICC categories 74-78. See Appendix B for more details and examples of the 5-digit codes.

[49]Note that some product codes do not have quantity information, although these are relatively rare, with only 15% of the item-plant level observations containing no quantity information, and just over 9% of plant-item level revenues involving no quantity information.

[50]I only include *worker* hours in L_{it}, rather than managerial hours, in part because this measure is less likely to be a public input.

et al. (2016).[51] [52]

Table 2: Plant-Year Summary Statistics: Machinery, Equipment, and Parts

Variable	Obs	Mean	Std. Dev.	Min	Max	P50
Log labour hours (l_{it})	35475	9.19	1.54	3.26	15.38	8.95
Log capital stock (rupees) (k_{it})	35475	15.74	2.34	.69	23.82	15.54
Log materials (Cobb-Douglas aggregator) (m_{it})	35475	7.08	3.27	-3.91	22.99	6.63
Number of Varieties (J_{it})	35475	2.52	2.06	1	10	2
Multi-product	35475	.57	.5	0	1	1
Multi-product × No VI	35475	.46	.5	0	1	0

Notes: Summary statistics only reported for plants which only produce products belonging to ASICC codes 74-78. Labour hours refers to mandays worked by workers, i.e. excluding managerial workers. Net Closing Value is the net closing value for *all* fixed assets reported in Block C of the ASI, deflated by the yearly capital deflator used in Allcott et al. (2016). Cobb-Douglas materials are constructed according to the following formula: $M_{it} = \frac{\sum_k \text{cost}_{it}^k}{\prod_k \left(\frac{W_{it}^k}{\gamma_{it}^k}\right)^{\gamma_{it}^k}}$, where cost_{it}^k is total expenditure on input k by firm i, W_{it}^k is the unit-value (price) of input k, and γ_{it}^k is the fraction of observed materials input expenditure going into material k. Multi-product and No VI are dummy variables, where Multi-product equals 1 if a plant produces multiple products, while No VI equals 1 if the plant does not produce an output sets that I classify as potentially vertically integrated. See Appendix B for more details.

Before I move on to discussing estimation, there are two features of the ASI data that require some minor tweaks to the baseline model discussed in Section 3. First, note that the ASI contains information on both *quantity sold* and *quantity manufactured* in each year (See Table 1). While these variables are generally quite similar, with just over 75 % of observations involving quantities sold and produced that are within 5 % of each other, there are some cases where quantity sold and quantity manufactured differ by a significant margin. Since this is due to inventory management, I augment the model described in Section 3 to account for inventories by modifying the output market clearing condition to be $Q_{it}^j = Y_{it}^j - \Delta INV_{it}^j$, where ΔINV_{it}^j is the change in inventories of product j at time t, while I take Q_{it}^j to be quantity sold and Y_{it}^j to be quantity produced. In practice, accounting for inventories in this way changes the model very little, except for slight differences in notation. In particular, inventories now become a state variable for the firm's problem, and marginal costs should

[51]This index is created using information on the value and quantity of up to ten major inputs that are purchased from domestic firms, and up to five material inputs that are imported from firms abroad. I use this information to construct firm-varying materials input prices, W_{it}^k, by dividing the purchase values by observed purchase quantities. I then construct a Cobb-Douglas price index for materials, given by $\prod_k \left(\frac{W_{it}^k}{\gamma_{it}^k}\right)^{\gamma_{it}^k}$, where γ_{it}^k is the fraction of observed materials input expenditure going into material k. Aggregate materials, M_{it}, is then determined by dividing total observed materials expenditures by the plant level materials input price index. Note that this formulation of the Cobb-Douglas price index implies that different plants may use Cobb-Douglas materials aggregators that have different input elasticities, γ_{it}^k. However, cost minimization by the firm implies that the observed materials cost shares will be equal to these heterogeneous input elasticities.

[52]Note that using such a Cobb-Douglas index to measure input quantities may generate issues in the comparability of TFP across firms, as plants using different types of inputs may not be in comparable units. These differences in units may then show up in one's TFP estimates, which will generate measured differences in productivity that have nothing to do with "real" productivity differences. To deal with this concern, I also re-estimate all of my main TFP regressions with *materials expenditures* in place of the materials quantity aggregator, to verify that my results are not being driven by materials aggregator issues. These results can be found in Appendix E.

be interpreted as marginal costs *conditional* on some desired change in inventories.[53]

The second feature of that the data that I have to account for is the fact that the ASI records information on the amount of tax paid by each item sold. Since taxes generate an *observable* wedge between the price paid by consumers and the price received by producers, I modify the pricing model in Section 3 to account for differences in *ad-valorem* tax rates that must be paid by consumers across products.[54] This generates a minor modification to the pricing first-order conditions, which I discuss in detail in Appendix C2, where I work out the details of the marginal cost inversion for my empirical model.

Finally, since I wish to examine the effect of import competition on within-plant specialization, I supplement the ASI data with information on aggregate imports and tariff rates by 4-digit HS code from UN Comtrade, and the UNCAD TRAINS database, respectively. I then match this information to the 5-digit ASICC codes I observe in the ASI, using a series of public crosswalks, as well as information on the value of imported inputs reported in the ASI, to deal with many-to-one matches of product codes. See Appendix B3 for details.

5 Estimation

While Theorem 1 applies to a fairly general demand and production functions, to make estimation feasible for the given sample size, I specify the functional form of the demand and production function up to small number of unknown parameters, as is standard in most applied work. In the following section, I first describe the specification of the demand system, present my identification strategy for identifying the demand parameters, and then present the estimates. I then describe the specification and estimation of the production function, and present my estimated production function parameters.

5.1 Demand Estimation

I specify the aggregate demand system using a variant of the discrete choice models incorporating continuous quantity choice described in Björnerstedt and Verboven (2013). While more details are described in Appendix C, this demand system is essentially a generalization of the standard nested logit model described in McFadden et al. (1978) or Berry (1994), that allows consumers to make continuous quantity choices.[55] As in nested logit, the consumer choice problem can be represented as a sequential choice problem, with consumers first choosing a *product-nest*, which I will consider to be separate 5-digit ASICC codes, and then a *variety* within that product nest, corresponding to a particular product

[53]Note that the conditional cost-minimization problem remains conditional on $Y_{it}^{j} = Q_{it}^{j} + \Delta INV_{it}^{j}$ in this case, which means the conditional cost function for (10) becomes $C\left(\vec{K}_{it}, \vec{Q}_{it}\left(\vec{P}_t\right) + \Delta INV_{it}^{j}, \vec{\omega}_{it}, \vec{A}_{it}\right)$, which, in practice, changes very little except for the *interpretation* of marginal costs. Note that I will treat inventory changes as exogenous, since they are not the focus of this paper. However, one could certainly augment the model to allow inventory changes to be a firm-level choice variable.

[54]I measure the *ad-valorem* tax rate as the ratio of total taxes recorded, over gross revenues minus total taxes. Note that most of the tax rates are quite small, with a median tax rate of around 4%.

[55]Note that this demand model is also equivalent to a two-stage nested CES demand function, as originally shown by Verboven (1996).

produced by an individual plant. Let Λ_t^g denote the *set* of varieties belonging to product-nest g, *including imported varieties*. Demand for product $j \in \Lambda_t^g \in \Omega_t^h$, where Ω_t^h is the set of varieties belonging to 3-digit ASICC code h, is then given by:

$$Q_{it}^{jg}\left(\vec{P}_t\right) = \frac{E_t^h}{P_{it}^{jg}} \frac{\exp\left(\frac{\delta_{it}^{jg}}{\sigma}\right)\left(\sum_{k \in \Lambda_t^g} \exp\left(\frac{\delta_{mt}^{kg}}{\sigma}\right)\right)^{\sigma-1}}{\sum_{\Lambda_t^l \in \Omega_t^h}\left(\sum_{k \in \Lambda_t^l} \exp\left(\frac{\delta_{mt}^{kl}}{\sigma}\right)\right)^{\sigma}}, \tag{26}$$

where $\alpha > 0$ and $\sigma \in (0, 1]$ are demand parameters, $\delta_{it}^{jg} \equiv \eta_{it}^{jg} - \alpha p_{it}^{jg}$ is the *mean-utility* of product j, and E_t^h is total expenditure by consumers in the market h at time t.[56] In this formulation of the demand system, I consider different 3-digit ASICC codes to be different markets, the sense that they correspond to different *choice-sets*, Ω_t^h.[57]

The general approach to demand estimation is based on the mean utility inversion approach outlined in Berry (1994), where a simple linear estimating equation is generated by the *inverse mapping* between market shares, and the mean-utility levels, δ_{it}^{jg}. To obtain this mapping, I need to take a stance on the *outside option*, which in this application corresponds to products within the market that are *not* within the ASI's sample frame, or product codes that are within consumers choice sets but are not fully accounted for within the given 3-digit ASICC code. Let $0 \in \Omega_t^h$ index this outside option for any market h, and normalize $\delta_t^0 \equiv 0$, since discrete choice models are only identified up to such a normalization. It is then straightforward to show that mean-utilities are revealed by the following function of *revenue-shares*:[58]

$$\delta_{it}^{jg} = \frac{rs_{it}^{jg}}{rs_t^{0h}} - (1-\sigma)\, rs_{it}^{j|g}, \tag{27}$$

where $rs_{it}^{jg} \equiv \ln\left(\frac{R_{it}^{jg}}{\sum_{\Lambda_t^g \in \Omega_t^h}\sum_{l \in \Lambda_t^g} R_{mt}^{kl}}\right)$ is the log of product j's revenue share, rs_t^{0h} is the natural log of the revenue share of the outside option in market h, and $rs_{it}^{j|g} \equiv \ln\left(\frac{R_{it}^{jg}}{\sum_{k \in \Lambda_t^g} R_{mt}^{kg}}\right)$ is the natural log of the revenue share of variety $j \in \Lambda_t^g$ *within* 5-digit ASICC code g. Note that I consider imported goods to

[56] Recall that lower case letters correspond to natural logs, so $p_{it}^j \equiv \ln\left(P_{it}^j\right)$.

[57] See Appendix B2 for a list of 3-digit ASICC codes. While one could certainly define choice sets in alternative ways, this particular formulation was chosen so cannibalization effects within multi-product firms are appropriately estimated. In particular, note that if all products within Machine, Parts, and Equipment are considered part of the same choice set, most firms would have approximately zero market shares, which will generate zero cross-product-code elasticities within this demand system (See Appendix C). While I regard Machine, Equipment, and Parts as fairly similar objects from the production side, on the consumption side (specifically, from the point of view of downstream firms), not all items within these sets are likely to substitute towards one another. Since the 3-digit ASICC codes appear to grouped in terms of downstream uses, such as Agricultural and Forestry Equipment (761) vs Food, Beverage and Tabacco machinery (762), these provide a relatively straightforward way to group products that should substitute towards one another.

[58] The derivation of the mean-utility inversion exactly follows the derivation for the nested logit presented in Berry (1994), with revenue shares replacing quantity shares. In particular, note that (26) can be rewritten as $RS_{it}^{jg} = \frac{\exp\left(\frac{\delta_{it}^{jg}}{\sigma}\right)\left(\sum_{k \in \Lambda_t^g} \exp\left(\frac{\delta_{mt}^{kg}}{\sigma}\right)\right)^{\sigma-1}}{\sum_{\Lambda_t^g \in \Omega_t^h}\left(\sum_{k \in \Lambda_t^l} \exp\left(\frac{\delta_{mt}^{kl}}{\sigma}\right)\right)^{\sigma}}$, where $RS_{it}^{jg} \equiv \frac{P_{it}^{jg} Q_{it}^{jg}}{E_t^h}$. Note that this is identical to the expression for the nested logit demand model in Berry (1994), except with revenue shares replacing quantity shares.

be part of each 5-digit ASICC nest, and hence include total imported revenue within Λ_t^g.[59]

Equation (27) defines the inverse mapping between market shares and mean utilities which makes linear IV estimation feasible. To see this, substitute into $\delta_{it}^{jg} \equiv \eta_{it}^{jg} - \alpha p_{it}^{jg}$ into (27), yielding:

$$\frac{rs_{it}^{jg}}{rs_t^{0h}} = (1-\sigma)\, rs_{it}^{j|g} - \alpha p_{it}^{jg} + \eta_{it}^{jg}. \tag{28}$$

Since this estimating equation comes from a discrete choice model, the two demand parameters that I wish to estimate in (28), (α, σ), can be interpreted as density parameters governing the distribution of idiosyncratic taste shocks of machinery purchasers in this market, as noted by Berry (1994). In particular, as I show in Appendix C, where I derive (26) from a discrete choice microfoundation, α governs the *variance* of taste shocks across all product codes and all varieties in the market, with larger values of α implying smaller variance of the idiosyncratic taste shocks. If there is less idiosyncratic taste variation, this means that horizontal product differentiation will be much less important in the aggregate, and therefore price elasticities will generally become larger for all varieties.

The parameter σ, on the other hand, governs the correlation of taste shocks *within* a product code. In particular, low values of σ imply that individual consumers draw a vector of idiosyncratic taste shocks that are strongly correlated within a product code. As a result, each consumer tends to, on average, prefer consuming varieties within one product code over all the others, leading to larger cross-price elasticities *within* product code, compared to across product codes. On the other hand, as σ approaches 1, this generates zero correlation in the taste shocks within product codes. Note that examining how a product's within-product code revenue share, $rs_{it}^{j|g}$ leads to variation in it's overall market share rs_{it}^{jg}, helps identify this consumer taste correlation parameter, with the special case of no correlation corresponding to the case where $rs_{it}^{j|g}$ drops out of (28).

Equation (28) can be estimated using linear instrumental variables methods, with unobserved quality, η_{it}^{jg}, functioning as the structural residual.[60] Note that instrumental variables are needed because plants choose their prices with knowledge of their quality levels, η_{it}^{jg}. As a result, plants that produce higher quality products will tend to charge higher prices, meaning that the OLS estimates of α are likely to be downward biased. Furthermore, the within group revenue shares, $rs_{it}^{j|g}$, are also endogenous, implying that the estimates of $1-\sigma$ obtained by OLS would be inconsistent.

To find a suitable instrument for prices in (28), one generally needs a cost-shifter, such as input

[59]See Appendix B3 for details on the mapping between 5-digit ASICC codes and 4 digit HS codes used to determine total imports by ASICC code.

[60]Note that the right hand side of (28) is only observable if potential market size, E_t^h, is known, in which case the outside option can be constructed by subtracting observed ASI revenue in each period from potential market size. Since India experienced significant growth over the sample period, it is important to allow E_t^h to grow over time. To accomplish this, I set potential market size in 2001 for each three-digit ASICC code, E_{2001}^h, to two times observed revenue in the ASI, for the year 2001, where total industry revenue is obtained by summing up all observed revenues in the ASI for the given three-digit ASICC codes, multiplied by the the firm-level multipliers for plants that are randomly sampled, plus imports. Potential market size in subsequent years is then calculated using $E_t^h = GR_t E_{t-1}^h$, where GR_t is the growth factor in total observed revenue for all plants I observed in the Machinery, Equipment, and Parts industry plus imports, from year $t-1$ to t. This formulation of potential market size means that 3-digit ASICC codes that experienced slower than average revenue growth will have larger shares of consumers choosing the outside option. Note that since this formula led to occasionally negative outside option shares for a small number of 3-digit ASICC codes, I instead scaled the initial levels of *these* codes by three times observed revenue in 2001, to guarantee that the outside option share is always positive.

prices, that are excluded from the demand function. While plant-varying input prices are observable in the ASI, the key difficulty in the differentiated product context considered in this paper is that the structural residual, η_{it}^{jg}, embodies quality differences across products. As a result, a standard instrument that may be valid in a homogeneous goods context, i.e. plant-varying input prices, are unlikely to be valid, since differences in input prices across firms may reflect differences in the product-level *quality choices*, as in, for example, Verhoogen (2008), Kugler and Verhoogen (2012) and De Loecker et al. (2016). Specifically, if plants who wish to produce higher quality outputs need to purchase higher quality inputs, input quality differences across firms are likely to be reflected in these prices, thus violating the required exclusion restriction for input prices to function as valid instruments.

To address this concern, I generate an input-price instrument that only harnesses variation in input prices in *other* output markets, thereby harnessing cost-variation that should be uncorrelated with demand or quality shocks in the Machinery, Equipment, and Parts industry. To construct these instruments, let $\mathbb{I}^g$ denote the set of 5-digit input product codes that I observe being used by *single product producers* of product code g, and let $\mathbb{F}_t^{kg}$ denote the set of *firms* observed in the ASI at time t who purchase an input with product code $k \in \mathbb{I}^g$, *who do not sell any outputs in the Machinery, Equipment, and Parts Industry.* Letting W_{it}^k denote the current price of input-code k paid by firm i, I define the following input price instrument which varies at the product-code and year level:

$$Z_{it}^{jg} = Z_t^g = \sum_{k \in \mathbb{I}^g} \gamma^{kg} \times \ln\left(\frac{\sum_{m \in \mathbb{F}_t^{kg}} W_{mt}^k}{|\mathbb{F}_t^{kg}|}\right), \tag{29}$$

where γ^{kg} is an input weight, given by the overall cost share of input $k \in \mathbb{I}^g$ in the production of product-code g by single product firms.[61]

Roughly speaking, Z_t^g simply uses average input prices in other output markets, and will function as a valid instrument as long as time-variation in the price of inputs paid by firms operating in other output markets are primarily driven by demand and supply shocks that are specific to those industries, rather than demand shocks or quality changes in the Machinery, Equipment and Parts industry. Put differently, I require that this input price variation be driven by demand and supply shocks in other industries, that, on average, are orthogonal to machinery demand shocks. Roughly speaking, this requires that average input prices not be driven by machinery industry demand, which will obviously not be satisfied in input markets where the machinery industry is the primary downstream consumer. To deal with this concern, I exclude any input codes k from $k \in \mathbb{I}^g$ if more than 30% of the revenue I observe going into purchases of k comes from Machinery, Equipment, and Parts producers.[62]

[61] These weights are chosen to put greater weight on the price changes of inputs that are more important for the production of product code g. Note that in practice I trim the 95th and 5th percentiles of these prices by product code to limit the influence of extreme outliers, and demean the value of this instrument by product code so as to only take advantage of across time variation in input prices for each product code, rather than variation in the levels. See also footnote 62.

[62] Since I am excluding some inputs from the construction of the input price instrument, this means that the input weights, γ^{kg}, do not necessarily sum to one for each output code g, and therefore *levels* of the input price instrument are not necessarily informative. To deal with this, I demean the value of each instrument by product code so as to only take advantage of the across time variation in the price shocks. While one could instead rescale the instrument so that the weights summed to one for the non-excluded input codes, this strategy would artificially put greater weight on cost shocks

Note that one limitation inherent in Z_t^g is that it does not vary across plants that produce the same product codes. While one could define an analogous instrument at the firm-level, taking advantage of the different input sets used by different plants, I do not pursue this strategy since the choice of different inputs is almost certainly correlated with product quality. Instead, I construct a second instrument that takes advantage of the variation in the *output sets* produced by different multi-product plants. In particular, this instrument is based on the average value of Z_t^g taken by *other* products produced within the same firm:

$$Z_{it}^{-jg} = \frac{\left(\sum_{k \in \mathbb{Y}_{it}} Z_{it}^{kl}\right) - Z_{it}^{jg}}{J_{it} - 1}, \tag{30}$$

where $J_{it} = |\mathbb{Y}_{it}|$. Note that this instrument will be correlated with price either through within-plant cannibalization effects, or through cost-shocks that affect common inputs that are used in each production line within a plant.

The results of this estimation strategy can be found in Table 3, below.[63] As expected, OLS estimation of (28) generates a price coefficient of the wrong sign, since prices tend to be positively correlated with quality. The instrumental variables strategy, on the other hand, appears to fix this bias, generating point estimates that imply an average own-price elasticity of approximately 2.4. Note that the coefficient on $rs_{it}^{j|g}$ is quite close to one, which, given the discrete choice formulation underlying this demand system, means that consumer level taste shocks are strongly correlated *within* a product code, and as a result, cross-product substitution will tend to primarily occur *within* product codes, rather than across product codes. As a result, cannibalization effects will tend to be strongest in plants producing multiple varieties of the same product code, rather than plants producing multiple product codes.

To alleviate potential weak instrument concerns, I also report the first stage estimates in Table 4. Note that the input-price instruments are strongly correlated with the endogenous variables, with the first stage F-statistics taking values of around 24 and 12. Moreover, note that the first-stage coefficients generally take the sign one would intuitively expect, with increased input prices being associated with output price increases.[64]

---

for inputs with small cost shares, γ^{kg}, in product codes where an important input was excluded from $k \in \mathbb{I}^g$, making this instrument less likely to scale the cost shocks correctly.

[63] Note that the regressions include a set of dummies for various plant-level observables, including state, urban versus rural, census status, as well as organization and ownership type, as an alternative to including plant fixed effects. This allows me to use the price variation of the randomly sampled plants in my demand estimates, which I would have to drop if I included plant fixed effects.

[64] The fact that the input price instrument is *positively* correlated with the within-product code revenue shares is perhaps surprising, since one would expect input price increases to decrease market shares. Note, however, that since this instrument does not vary *within* a product code, this is likely due decreased output (or exit) of *competitors*, which will mechanically increase $rs_{it}^{j|g}$ within a product code. On the other hand, Z_{it}^{-jg} being negatively associated with $rs_{it}^{j|g}$ makes sense, in so far as cost shocks that affect common inputs within a plant lead to price increases, and therefore loses in market shares.

Table 3: Demand Estimates

	(1) OLS	(2) IV
p_{it}^{jg}	0.00682***	-0.233**
	(0.00210)	(0.111)
$rs_{it}^{j\|g}$	0.949***	0.843***
	(0.00354)	(0.214)
Observations	60,066	60,066

Standard errors clustered by plant and product code

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

Notes: Dummies for product code, year, state, census status, age, urban, organization and ownership type, and number of products sold also included in all regressions.

Table 4: First Stage Estimates

	(1) p_{it}^{jg}	(2) $rs_{it}^{j\|g}$
Z_t^g	0.328***	0.157***
	(0.109)	(0.0448)
Z_{it}^{-jg}	0.353	-0.401***
	(0.267)	(0.145)
Observations	60,066	60,066
AP F-Stat	23.74	11.74

Standard errors clustered by plant and product code

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

Notes: Dummies for product code, year, state, census status, age, urban, organization and ownership type, and number of products sold also included in all regressions.

5.2 Production Function Estimation

As is standard in much of the productivity literature, I assume that all firms in the industry use a Cobb-Douglas production technology, $Y_{it}^{j} = \exp\left(\omega_{it}^{jg}\right)\left(L_{it}^{jg}\right)^{\beta_L}\left(K_{it}^{jg}\right)^{\beta_K}\left(M_{it}^{jg}\right)^{\beta_M}$.[65] To obtain an estimate of the within-firm input allocations, I apply Theorem 1, using the estimated demand parameters $(\widehat{\alpha}, \widehat{\sigma})$ to determine the appropriate mapping between observables and estimated input allocations, $(\widehat{L}_{it}^{jg}, \widehat{K}_{it}^{jg}, \widehat{M}_{it}^{jg})$.[66] Having solved for the input allocations, production function estimation can proceed using standard production function estimation tools, as outlined below. To generate a simple estimating equation to guide this discussion, take logs of the production function after substituting in the estimated input allocations, yielding:

$$y_{it}^{jg} = \beta_L \widehat{l}_{it}^{jg} + \beta_K \widehat{k}_{it}^{jg} + \beta_M \widehat{m}_{it}^{jg} + \omega_{it}^{jg}. \tag{31}$$

As discussed earlier, a key identification problem inherent in estimating models like (31) is the *transmission bias* problem that has been highlighted by much of literature on production function estimation.[67] Since the firm chooses input quantities with knowledge of their own productivity, each of the inputs is likely to be correlated with the structural residual, ω_{it}^{jg}, in (31). As a result, one cannot simply estimate (31) by OLS.

A popular approach to circumventing the transmission bias problem uses the proxy-variable techniques developed in Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015), which have fruitfully been applied to answer a number of different questions related to productivity.[68] The key insight exploited in these papers is that input demands (capital investment in Olley and Pakes (1996), materials in Levinsohn and Petrin (2003)), directly depend on productivity when firms are profit maximizing. As a result, we can use information on input demands to proxy for TFP(Q), thereby controlling for the endogeneity problem created by unobservable productivity shocks.[69]

While a proxy-variable approach could be used to estimate (31) in principle, there are a number of complications introduced by multi-product firms that make these estimators intractable in this context. First, as discussed by Ackerberg et al. (2015), the Olley and Pakes (1996) and Levinsohn and Petrin (2003) approach will only identify the labour input elasticity if there is optimization error in labour input demand, or under variable timing assumptions where labour and materials are chosen with different information sets. These extra assumptions would invalidate the assumption underlying Thereom 1, and thus are not particularly useful in this context.

On the other hand, while the proxy-variable estimator proposed by Ackerberg et al. (2015) could still be used, note that the semi-parametric approach they propose to estimate TFP(Q) is unlikely to be

[65] Note that since I measure output in *quantity* units, ω_{it}^{jg} will measure TFPQ, rather than TFPR.

[66] See Appendix C2 for more details on the appropriate mapping given the assumed demand system (26). Note that since I have only estimated the production function for the Machinery, Equipment, and Parts industry here, I drop multi-industry plants in the subsequent regressions, as I do not have the necessary information to determine their input allocations.

[67] See Griliches and Mairesse (1995) and Ackerberg et al. (2007).

[68] See, for example, Pavcnik (2002), De Loecker (2007), and Topalova and Khandelwal (2011).

[69] Note that TFP here should be regarded as TFPQ, since I y_{it}^{j} is in quantity units.

tractable for multi-product firms. In particular, a key assumption maintained in Ackerberg et al. (2015) is that there exists a one-to-one mapping between current TFPQ levels and firm-level state variables. This mapping, called the control function, is estimated non parametrically in a first stage regression of outputs on inputs.[70] Note, however, that the state-space for multi-product firms is the entire vector of product-level TFPQ terms, ($\vec{\omega}_{it}$). While in principle the general approach can be augmented to account for multiple unobservables, see Ackerberg et al. (2007), this approach will quickly run into dimensionality problem in my context, since the vector of unobserved quality terms, $\vec{\eta}_{it}$, should also be state variables, as well as the particular *set* of product codes produced by the firm.[71] Even if one could guarantee that a well-defined control function exists, given the large state space, it is unlikely that one could approximate it accurately in a first-stage regression.

To avoid the issues associated with adapting the semi-parametric control function to a multi-product firm context, I instead consider a dynamic-panel style estimator, as in Anderson and Hsiao (1982), Arellano and Bond (1991), Blundell and Bond (1998), Blundell and Bond (2000) and Bond and Söderbom (2005), for the purposes of production function estimation. This approach does not require the use of a control function, and thus will not run in to the dimensionality problems discussed above. Moreover, since this does not require the existence of a control function, it places less restrictions on the econometrician's information set, since there no longer needs to be a one-to-one mapping between observables and productivity. For example, there can be firm-varying cost shocks to to all inputs that are unobserved by the econometrician, which proxy-variable methods generally cannot allow.

To obtain the estimating equation for the production function, I assume that productivity follows an AR(1) process, i.e.:

$$\omega_{it}^{jg} = \rho_0^g + \rho\omega_{i,t-1}^{jg} + \xi_{it}^{jg}, \tag{32}$$

where ρ_0^g is the mean TFP of varieties belonging to product-code g.

Substituting this law of motion for productivity into (31), yields $y_{it}^{jg} = \rho_0^g + \beta_L \widehat{l}_{it}^{jg} + \beta_K \widehat{k}_{it}^{jg} + \beta_M \widehat{m}_{it}^{jg} + \rho\omega_{i,t-1}^{jg} + \xi_{it}^{j}$. Once can then quasi or "ρ-difference" this equation, by subtracting the lagged value of (31) multiplied by ρ from this expression, yielding :

$$y_{it}^{jg} = \rho_0^g + \rho y_{i,t-1}^{jg} + \beta_L\left(\widehat{l}_{it}^{jg} - \rho\widehat{l}_{i,t-1}^{j}\right) + \beta_K\left(\widehat{k}_{it}^{jg} - \rho\widehat{k}_{i,t-1}^{jg}\right) + \beta_M\left(\widehat{m}_{it}^{jg} - \rho\widehat{m}_{i,t-1}^{jg}\right) + \xi_{it}^{jg}, \tag{33}$$

where the innovation to productivity, ξ_{it}^{j}, becomes the structural residual. Since the above model is nonlinear in parameters, due to the interactions between ρ and the production function parameters, (33) is estimated using a nonlinear GMM estimation procedure.

Two related points are worth emphasizing concerning estimating equation (33). First, the structure of this estimating equation is very similar to that used in Ackerberg et al. (2015), with y_{it}^{j} essentially

[70]Technically, the control function will not actually be identified until the second-stage is estimated, since one cannot separate production function terms from control function terms in their first stage regression. Nevertheless, the key idea is that *some* unknown function of inputs and firm-level state variables must be accurately estimated in the first stage.

[71]The set of product codes being a state variable leads to most significant dimensionality problem. Roughly speaking, this means one must estimate a different control function for *every* combination of product codes observed in the data.

taking the place of their first stage predicted values.[72] This relationship between dynamic panel estimators and their estimator is also noted by Ackerberg et al. (2015), with the key tradeoff between the two approaches being that the dynamic panel approach relies more heavily on the linearity of the productivity process, while proxy-variable approaches generally allows productivity to follow a non-linear Markov process.[73] Second, and perhaps more importantly, the key difference between the specification of the production function (31) that I use to generate (33), compared to the approach in Ackerberg et al. (2015), is that there are no productivity shocks that are *unobserved* by the firm. This is a more restrictive assumption on a firm's information set, relative to Ackerberg et al. (2015), which the reader should bear in mind. Note that while one can modify the above estimating equation to incorporate these extra productivity shocks, as Ackerberg et al. (2015) do when comparing the dynamic panel method to their approach, this would be problematic in a multi-product firm context because shocks that are unobserved by the firm would break the link between observed outputs, and the input allocations firm chose with knowledge of their TFPQ levels, which was used to derived the optimal allocation rule (9).[74] Hence, while acknowledging that this approach does require that I make some stronger assumptions on the firm's information set, relative to proxy-variable approaches, note that these extra restriction are only likely to be of quantitative importance if the unobserved component of productivity is relatively large compared to the component known by firms.

I now turn to the nonlinear instrumental variables strategy I use to estimate (33). As is standard in much of the production function literature, I assume that materials is a *static* input, while capital is a *dynamic and predetermined* input. In particular, I assume one-period time to build in capital, as in Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015), so that firm-level capital follows the simple law of motion $K_{i,t+1} = l^K(K_{it}, I_{it}^K)$. This law of motion implies that K_{it} cannot react to any news that arrives during period t. Since ξ_{it}^{jg}, the "innovation" to productivity, functions as the structural residual in (33), *both* k_{it} *and* $k_{i,t-1}$ can be used as instruments to identify β_K, since they will both be uncorrelated with ξ_{it}^{jg}, as long as one accounts for the product-code specific intercepts, ρ_0^g, with a series of product-code fixed effects.[75]

Similar to Ackerberg et al. (2015), I also assume that labour is also a dynamic input, in the sense

[72] Or $\widehat{\Phi}_{it}^{jg}$, adapting their notation to my context.

[73] Strictly speaking, one can allow for a non-linear Markov process using a dynamic panel approach, as long as one is willing to abstract from measurement error in output. Note, however, that the linear AR(1) structure that I use allows me to account for differences in average TFPQ across product codes in a more parsimonious manner; i.e. by demeaning within product codes, or by including a series of product code fixed effects.

[74] It is easiest to think of this problem in the case where these unobserved productivity shocks occur *after* production. In this case, firm's choose expected output, rather than realized output. As a result, the link between their chosen level of output, and observable output which I use in (9), would be broken. While a "first-stage" estimating equation might allow the researcher to identify the difference between expected productivity and realized productivity, these first-stage estimating equations would run into the dimensionality problems discussed earlier, and multi-product firms would not be able to be included in the estimation procedure, as the existence of productivity shocks would make the input allocation rule no longer hold exactly.

[75] Note, however, that since I assume that capital is perfectly transferable across uses within the firm during each period, the allocation of capital input across production lines, k_{it}^{jg}, will generally be correlated with the current productivity shocks, since they are chosen in the current period. Hence k_{it} can be used as an instrument, while k_{it}^{jg} cannot. Moreover, as I mention in the main text, it is preferable to use firm-level inputs as instruments, rather than the estimated input allocations, since this will deal with attenuation bias due to the measurement error generated by the fact that the input allocations are *estimated*.

that it faces *dynamic adjustment costs*, such as hiring or firing costs due due labour market regulation and union activity, which are fairly substantial in many Indian states, see Besley and Burgess (2004). This allows me to use lagged labour as an instrument to identify β_L, since lagged labour inputs $l_{i,t-1}$, will tend to be correlated with the quasi-differences in labour, $\widehat{l}_{it}^{jg} - \rho \widehat{l}_{i,t-1}^{jg}$.

While it is fairly standard to also use the lagged value of static inputs in both proxy-variable and dynamic panel approaches to production function estimation, note that only using this type of variation to identify β_M is problematic. In particular, it is not entirely obvious why the lagged value of a static input should be correlated with differences in static input usage, *except through* productivity or demand changes, as has recently been pointed out by Gandhi et al. (2016). While the fact that lagged inputs may be weak instruments has been widely acknowledged in the dynamic-panel literature for a while, see Blundell and Bond (2000), note that this problem is most likely to be pronounced with static inputs.

To deal with this concern, I also use the current and lagged values of the input price instruments described in Section 5.1, Z_t^g and Z_{t-1}^g for any $j \in \Lambda_t^g$, to identify β_M. As long as these instruments are valid for demand, meaning that variation in the input-price instrument is driven by demand and supply shocks in *other* output markets, which affect the input prices of goods used to produce product code g, then they will function as valid instruments for estimating (33) as well.

Altogether, I use the following set of moments to identify the parameters of the production function:

$$\mathbb{E}\left[\xi_{it}^{jg}\begin{pmatrix} y_{i,t-1}^{jg} \\ l_{i,t-1} \\ k_{it} \\ k_{i,t-1} \\ m_{i,t-1} \\ Z_t^g \\ Z_{t-1}^g \\ \vec{C}_{it}^{jg} \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vec{0} \end{pmatrix} \quad \text{for } j \in \Lambda_t^g, \tag{34}$$

where $\vec{C}_{it}^{jg}$ is a vector of 5-digit product code fixed effects, as well as number of product fixed effects, which are included to account for the product-code specific intercepts, ρ_0^g, as well as the economy of scope effects discussed in Section 3.3. Note that I prefer to use *firm-level* input useage as instruments, x_{it}, rather than the estimated input useage by product line, $\widehat{x}_{it}^{jg}$ for $x \in (l, k, m)$, since using the estimated shares as instruments should generate finite sample attenuation bias, due to the fact that these shares are *estimated* from a first-stage regression, and thus, measured with error.

Estimation is done using the standard nonlinear IV estimator, and standard errors are obtained using the block bootstrap, where I re-sample with replacement over plant identifiers, and re-estimate the demand and production function parameters using the described two-step estimation procedure.[76] The OLS and GMM results are presented below, in Table 5.

[76]For computational simplicity, I first de-mean the data within each 5-digit product code, so I do not need to actually estimate the entire vector of product-fixed effects. Since (33) is linear conditional of $(\beta_L, \beta_K, \beta_M)$, I concentrate out the remaining parameters when minimizing the criterion function, so I only need to solve for $(\beta_L, \beta_K, \beta_M)$ non-linearly.

Table 5: Production Function Estimates

	OLS	GMM
β_L	0.527***	0.374**
	(0.0607)	(0.151)
β_K	0.245***	0.0819
	(0.0390)	(0.0739)
β_M	0.263***	0.780***
	(0.0227)	(0.128)
$\beta_L + \beta_K + \beta_M$	1.04***	1.24***
	(0.0327)	(0.0643)
Observations	3545	3545

Block bootstrapped standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The estimated production function parameters are fairly reasonable, with materials having the largest input elasticity, as is commonly the case with gross-output production function. Interestingly, these estimates allow me to reject constant returns to scale at the 5 % level. While constant returns to scale often cannot be rejected in many studies based on *revenue* production functions, actual returns to scale will tend to be biased downwards when one uses a revenue production function, as prices have to *fall* when a firm increases their output.[77] Hence, the machinery industry appears to be characterized by significant scale economies, which one would indeed expect in large scale manufacturing.

6 Results

Having estimated the demand and production parameters for the Machinery, Equipment, and Parts industry, I use this information to construct plant-product specific estimates of log TFPQ, defined as the estimated residual from (31), as well as plant-product specific quality, defined as residual from (28). I then use these estimates to quantify the extent to which within-firm heterogeneity matters, as well as examine the impact of import competition on within-firm specialization.

6.1 Product-Level Heterogeneity Correlations

I begin by presenting some simple correlations between the various sources of product-level heterogeneity. Table 6, below, shows the correlation between estimated quality, TFPQ, log prices, and markups, after being demeaned within their respective product code and year so I will not be making apples to oranges comparisons. Many of these correlations have the expected sign, with prices being

[77] See Klette and Griliches (1996).

negatively correlated with TFPQ but positively correlated with quality, and markups being positively correlated with quality.

Table 6: Variety-Level Correlations

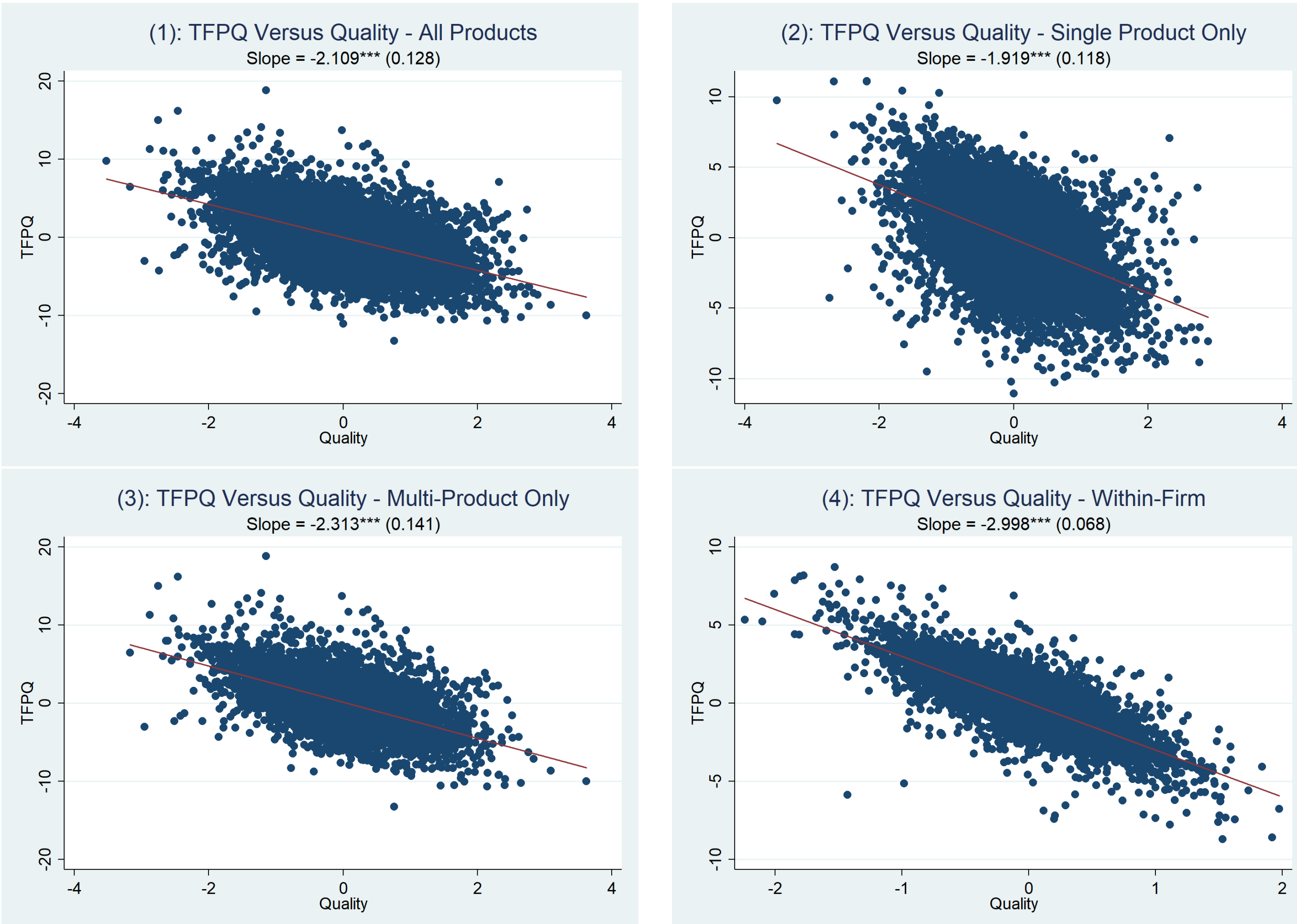
Variables	Quality	TFPQ	$\ln\left(P_{it}^{jg}\right)$	Markup
Quality $\left(\eta_{it}^{jg}\right)$	1.000			
TFPQ $\left(\omega_{it}^{jg}\right)$	-0.532***	1.000		
Price $\left(\ln\left(P_{it}^{jg}\right)\right)$	0.874***	-0.596***	1.000	
Markups $\left(\frac{P_{it}^{jg}}{MC_{it}^{jg}}\right)$	0.258***	-0.034***	0.087***	1.000

Notes: All variables demeaned within their respective product code and year. Quality refers to the estimated residual from (28), and TFPQ refers to the estimated residual from (31). Markups defined as the ratio of price to marginal costs, where marginal costs are estimated estimated using the marginal cost inversion described in Appendix C2. *p<0.1; **p<0.05; ***p<0.01.

One potentially surprising finding in Table 6 is that *quality and TFPQ are negatively correlated.* Note that this is a fairly strong negative correlation, and holds both across single-product firms, as well as across products *within* multi-product firms, as one can see from Figure 1, where I report a number of scatter plots between TFPQ and quality at the product level. Note that this sort of correlation appears to contradict the intuition, considered in Kugler and Verhoogen (2012) as well as Garcia-Marin and Voigtländer (2017), that TFPQ and quality are *complements*, in the sense that high TFPQ firms will wish to produce high quality products. Rather, it appears that *high TFPQ products tend to be lower quality.* A similar negative correlation between cost-side advantages and demand side advantages has also been found by Forlani et al. (2016), and Jaumandreu and Yin (2016), who consider the separate estimation of TFPQ and quality, primarily in a single-product firm context. These correlations could imply that these different sources of product advantages are simply substitutes, in the sense that high quality producers do not have an incentive to cut costs, since overall demand for high quality products is high enough to cover high costs. On the other hand, this correlation could also be due to technological constraints, where high-quality goods simply require more inputs (e.g., worker hours), and hence, will tend to have lower measured TFPQ.

While determining the exact reason for this negative correlation is beyond the scope of this paper, note that negatively correlated quality and TFPQ has important implications for how we think about firm-heterogeneity and specialization at the plant-level. In particular, there are, roughly speaking, two classes of products within this market: *cheap-goods*, which are low quality but easy to produce (i.e. high TFPQ), and *high-quality goods*, which generate more demand for a given price, but generally are costly to produce and therefore have relatively low TFPQ levels. In the following section, I more closely examine how plant-level specialization responds to import competition, and whether plants choose to focus on the quality dimension, or the productivity dimension.

Figure 1: Negative Correlation between TFPQ and Quality



Notes: Standard errors of slopes, adjusted for two-way clustering by plant and product code, reported in parentheses. Quality refers to the year-product code demeaned residual from (28), while TFPQ refers to the year-product code demeaned residual from (31). Panel (1) includes all products, Panel (2) only includes single-product plants, and Panels (3) and (4) only includes products produced by multi-product plants. Panel (4) plots within-firm demeaned values of TFPQ and quality.

6.2 Within-Plant Allocative Efficiency and Import Competition

Recent theoretical contributions by Eckel and Neary (2010), Bernard et al. (2011), Mayer et al. (2014) and Mayer et al. (2016) have shown that productivity gains from trade can operate through a *within-firm* margin. Specifically, multi-product firms that face greater import competition should be more specialized, either by dropping low quality or low efficiency varieties, or by focusing more of their resources towards their highest performing products. In this section, I obtain new evidence for these within-firm reallocative gains using the tools for measuring within-firm heterogeneity developed in this paper.

While previous papers have also examined empirical evidence for within-firm allocative gains from trade, either by examining product-dropping in response to trade liberalization, as in Bernard et al. (2011), or by examining revenue-skewness within plants in markets with different levels of trade exposure, as in Mayer et al. (2014) and Mayer et al. (2016), this the first study to examine within-firm allocative gains from trade using *direct* measures of within firm productivity and quality heterogeneity. This is important, as without estimates of the *magnitude* of within firm heterogeneity, it is impossible to determine whether within firm reallocations are economically important.[78]

As a first-step to determining whether within-firm reallocations matter, I first decompose the variance of product-level productivity and quality into across-plant and within-plant variation, in Table 7, below.[79] Note that the magnitude of within plant variance in productivity and quality is approximately half the magnitude of across plant variation, implying that there is less heterogeneity within plants than across plants. On the other hand, the portion of the variation explained by the within plant component is sizeable, accounting for more than a third of the overall variation for both dimensions of product level heterogeneity.

Table 7: Multi-product Plant Variance Decompositions

	Across Firm	Within Firm	Total
TFPQ			
Variance	4.288	2.379	6.667
Percentage	64 %	36 %	100 %
Quality			
Variance	0.250	0.168	0.418
Percentage	60 %	40 %	100 %
Observations	11,172	11,172	11,172

Notes: Only products produced by multi-product firms included in the sample. TFPQ corresponds to log quantity TFP, ω_{it}^{j}, from the main text, demeaned at the 5-digit product-year level. Quality corresponds to the estimated residual from (28), demeaned at the 5-digit product-year level. Across firm variance corresponds to variance of estimated $\left(\frac{1}{J_{it}}\sum_{k\in\mathbb{Y}_{it}} x_{it}^{k}\right)$, where x is the outcome of interest, while the within-firm variance corresponds to the estimated variance of $\left(x_{it}^{j}-\frac{1}{J_{it}}\sum_{k\in\mathbb{Y}_{it}} x_{it}^{k}\right)$.

Since Table 7 documents sizeable within-plant heterogeneity, it is not yet apparent how much within-plant heterogeneity *matters*, in the sense that within-plant reallocations can lead to within-plant efficiency improvements. While Mayer et al. (2014) get at these magnitudes by examining the relationship

[78]Mayer et al. (2014) get at the economic magnitude of within firm reallocations by showing that their model implies a linear relationship between a product's rank in the within firm productivity ladder and export sales, where the slope of this line depends on the degree of across and within firm heterogeneity. Since they can quantify the magnitude of across firm heterogeneity using estimates from other studies that use the same data, they use this relationship to estimate the magnitude of within firm heterogeneity that would rationalize the export sales distribution across products, given the structure of their model. Note that this approach relies heavily on the parametric assumptions they make concerning across firm heterogeneity (Pareto), as well as their approach to modelling within-firm heterogeneity (as one moves away from a firm's core product, marginal costs are geometrically increasing at a constant rate). My approach does not require any of these restrictions, as well as allows for multiple dimensions of within-firm heterogeneity (quality and productivity).

[79]These variance decompositions are based on the identity $x_{it}^{j}=\left(\frac{1}{J_{it}}\sum_{k\in\mathbb{Y}_{it}} x_{it}^{k}\right)+\left(x_{it}^{j}-\frac{1}{J_{it}}\sum_{k\in\mathbb{Y}_{it}} x_{it}^{k}\right)$, where the two terms in brackets are uncorrelated by construction. The variance of the first term corresponds to the across plant variation, while the second component corresponds to the within-plant component.

between within-firm reallocations and firm-level *productivity* gains, simply examining the productivity implications of within-firm reallocations would be problematic given the negative correlation between TFPQ and quality discussed above. In particular, *within-firm productivity improvements may come at the expense of quality*, which may hurt overall consumer welfare. Hence, as a first step towards quantifying the extent to which within-plant reallocations may matter, I first need a metric for product-level performance, or efficiency, that captures both dimensions of product-level heterogeneity.

We can find such a metric of efficiency by examining the residual of a *revenue production function*, following Klette and Griliches (1996) and De Loecker (2011). As I show in Appendix E, one can use the structure of the demand and production function to show that log revenues at the product level are related to input use as follows:

$$\ln(R_{it}^{jg}) = \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\left(\beta_l l_{it}^{jg} + \beta_k k_{it}^{jg} + \beta_m m_{it}^{jg}\right) + \underbrace{\widetilde{\omega}_{it}^{jg} + \widetilde{\eta}_{it}^{jg}}_{\equiv h_{it}^{jg}}, \tag{35}$$

where:

$$\begin{aligned}\widetilde{\eta}_{it}^{jg} \equiv\ & \frac{\eta_{it}^{j}}{\alpha+\sigma} + \frac{1}{1+\frac{\alpha}{\sigma}}\ln\left(E_t^h\right) \\ & - \frac{(1-\sigma)}{1+\frac{\alpha}{\sigma}}\ln\left(\sum_{k\in\Lambda_t^g}\exp\left(\frac{\delta_{mt}^{kg}}{\sigma}\right)\right) \\ & - \frac{1}{1+\frac{\alpha}{\sigma}}\ln\left(\sum_{l\in\Omega_t^h}\left(\sum_{k\in\Lambda_t^l}\exp\left(\frac{\delta_{mt}^{kl}}{\sigma}\right)\right)^{\sigma}\right),\end{aligned} \tag{36}$$

$$\widetilde{\omega}_{it}^{jg} \equiv \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\omega_{it}^{jg}, \tag{37}$$

Note that the production function coefficients in (35) are each scaled down by $\frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}$, since increases in input use will require that prices fall for quantity produced to still equal quantity demanded, resulting in smaller returns to scale than one would find when examining the quantity production function (31). More importantly, the residual from this rescaled production function, $h_{it}^{jg} \equiv \widetilde{\omega}_{it}^{jg} + \widetilde{\eta}_{it}^{jg}$, is a weighted sum of productivity shocks, $\widetilde{\omega}_{it}^{jg}$, and demand shifters, $\widetilde{\eta}_{it}^{jg}$, which incorporate variety-level quality differences, as well as product-code level variation in overall demand, markups, as well as product space congestion effects.[80]

Taking h_{it}^{jg}, or product-level TFPR, as a straightforward metric for product-level performance, I then define *plant-level* TFPR as the sum of plant-product level TFPR, weighted by plant-product level *input shares* S_{it}^{jg}:

[80]In particular, note that $\frac{(1-\sigma)}{\alpha+\sigma}\ln\left(\sum_{k\in\Lambda_t^g}\exp\left(\frac{\delta_{it}^{kg}}{\sigma}\right)\right)$ and $\frac{1}{\alpha+\sigma}\ln\left(\sum_{l\in\Omega_t^h}\left(\sum_{k\in\Lambda_t^l}\exp\left(\frac{\delta_{it}^{kl}}{\sigma}\right)\right)^{\sigma}\right)$, will increase whenever there is entry of new competitors, leading to an overall *decrease* in $\widetilde{\eta}_{it}^{jg}$ due to product space congestion effects. On the other hand, product-code level increases in markups tend to *decrease* $\frac{(1-\sigma)}{\alpha+\sigma}\ln\left(\sum_{k\in\Lambda_t^g}\exp\left(\frac{\delta_{it}^{kg}}{\sigma}\right)\right)$ and $\frac{1}{\alpha+\sigma}\ln\left(\sum_{l\in\Omega_t^h}\left(\sum_{k\in\Lambda_t^l}\exp\left(\frac{\delta_{it}^{kl}}{\sigma}\right)\right)^{\sigma}\right)$, since increases in prices decrease each δ_{it}^{jg} terms, and therefore tend to *increase* $\widetilde{\eta}_{it}^{jg}$.

$$h_{it} \equiv \sum_{j \in \mathbb{Y}_{it}} S_{it}^{jg} h_{it}^{jg}. \tag{38}$$

Following Olley and Pakes (1996), I can then decompose overall plant-level TFPR into *average* and *allocative efficiency* components:[81]

$$\begin{aligned} h_{it} &= \frac{1}{J_{it}} \sum_{j \in \mathbb{Y}_{it}} h_{it}^{jg} + \sum_{j \in \mathbb{Y}_{it}} \left(h_{it}^{jg} - \frac{1}{J_{it}} \sum_{j \in \mathbb{Y}_{it}} h_{it}^{jg} \right) \left(S_{it}^{jg} - \frac{1}{J_{it}} \sum_{j \in \mathbb{Y}_{it}} S_{it}^{jg} \right) \\ &= \underbrace{\overline{h}_{it}}_{\text{Average Performance}} + \underbrace{OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}\right)}_{\text{Allocative Efficiency}}. \end{aligned} \tag{39}$$

The second term in (39), sometimes referred to as the "OP" covariance after Olley and Pakes (1996), increases as a plant allocates a greater portion of its input towards high performance products, and therefore functions as a measure of allocative efficiency within a plant.[82] Since this object is generated by a simple revenue-productivity decomposition, this means that increases in this plant-level allocative efficiency are measured in revenue productivity units, thereby generating a single dimensional metric to quantify within-plant allocative gains.[83]

While h_{it}^{jg} provides a useful index of product-level performance, it will also prove useful to decompose h_{it}^{jg} further into within and across product code variation in TFPQ, as well as quality, so one can isolate the exact channels leading to within plant efficiency improvements. In particular, note that by construction:

$$h_{it}^{jg} = \widetilde{\omega}_t^g + \widetilde{\omega}_{it}^{j|g} + \widetilde{\eta}_t^g + \widetilde{\eta}_{it}^{j|g}, \tag{40}$$

where $x_t^g \equiv \frac{1}{|\Lambda_t^g|} \sum_{k \in \Lambda_t^g} x_{it}^{kg}$ is average performance by 5-digit product code, and $x_{it}^{j|g} = x_{it}^{jg} - x_t^g$ for $x \in \{\widetilde{\omega}_{it}^{jg}, \widetilde{\eta}_{it}^{jg}\}$ is the variety specific deviation from that average. As a result, one can decompose overall plant-level allocative efficiency as follows:

$$\begin{aligned} OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}\right) = & \underbrace{OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\omega}}_t^g\right)}_{\text{“Cheap” product codes}} + \underbrace{OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}}_t^g\right)}_{\text{“High demand” product codes}} \\ & + \underbrace{OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\omega}}_{it}^{j|g}\right)}_{\text{Relatively high TFPQ varieties}} + \underbrace{OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}}_{it}^{j|g}\right)}_{\text{Relatively high quality varieties}}. \end{aligned} \tag{41}$$

The decomposition in (41) shows that there are four ways in which a plant can increase within-plant allocative efficiency. In particular, they can specialize in high TFPQ, or high demand *product codes*, in the sense that these product codes generally require less inputs than other product codes,

[81] Derivation exactly follows the derivation provided in Olley and Pakes (1996).

[82] See, for example, Bartelsman et al. (2013) and Melitz and Polanec (2015), who use this terminology in the context of across plant reallocations.

[83] In fact, most productivity studies are in revenue-productivity (TFPR) units, rather than TFPQ units - see De Loecker and Goldberg (2014). The only novelty here is measuring plant-level revenue productivity as the input-weighted sums of product-level log TFPR, mirroring the industry-level revenue productivity decomposition of Olley and Pakes (1996).

or generate more revenues for a given log-price, respectively. On the other hand, a plant may also specialize in products that it is *relatively* efficient at producing, in the sense that the product-level TFPQ or "demand shifters" are larger when compared with other firms producing the same product code.

Since, as I show in Appendix E, all the within product-code variation in $\widetilde{\eta}_{it}^{jg}$ is due to the demand residual from (28), I refer to increases in $OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}}_{it}^{j|g}\right)$ as increased "quality specialization", while I refer to increases in $OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\omega}}_{it}^{j|g}\right)$ as "relative TFPQ" specialization. Variation in these terms will be the key focus of my subsequent analysis of firm-level performance, since increased relative quality and relative TFPQ will generally indicate increased efficiency. Note, however, that variation in $OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\omega}}_{t}^{g}\right)$ and $OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}}_{t}^{g}\right)$ cannot always be given an increased efficiency interpretation, as different product codes are measured in different units, and therefore some high TFPQ product codes may simply appear "cheap" when compared to other products because they are measured in larger units. This is simply a reflection of the fact that making across product-code comparisons of productivity (and quality) are not always meaningful, and therefore interpreting the product-code level specialization terms will generally be difficult.[84]

Bearing these caveats in mind concerning the interpretation of the across product code specialization terms, I consider aggregate trends in plant-level efficiency in Figure 2, below. In panel (1), I plot the growth rate of each allocative efficiency term in (41), *relative* to the average value of firm-level performance, $\overline{h}_t \equiv \frac{1}{N_t}\sum_{i=1}^{N_t} h_{it}$, in 2000, similar to the industry-wide decompositions in Pavcnik (2002). In particular for each dimension of product-level heterogeneity, $X \in (\widetilde{\omega}_t^g, \widetilde{\eta}_t^g, \widetilde{\omega}_{it}^{j|g}, \widetilde{\eta}_{it}^{j|g})$, I plot:

$$(\text{Plant-Level RTFP Growth Due to Specialization in X})_t = \frac{\overline{OP}_t\left(\vec{S}_{it}, \vec{X}\right) - \overline{OP}_{2000}\left(\vec{S}_{i,2000}, \vec{X}\right)}{\overline{h}_{2000}}, \tag{42}$$

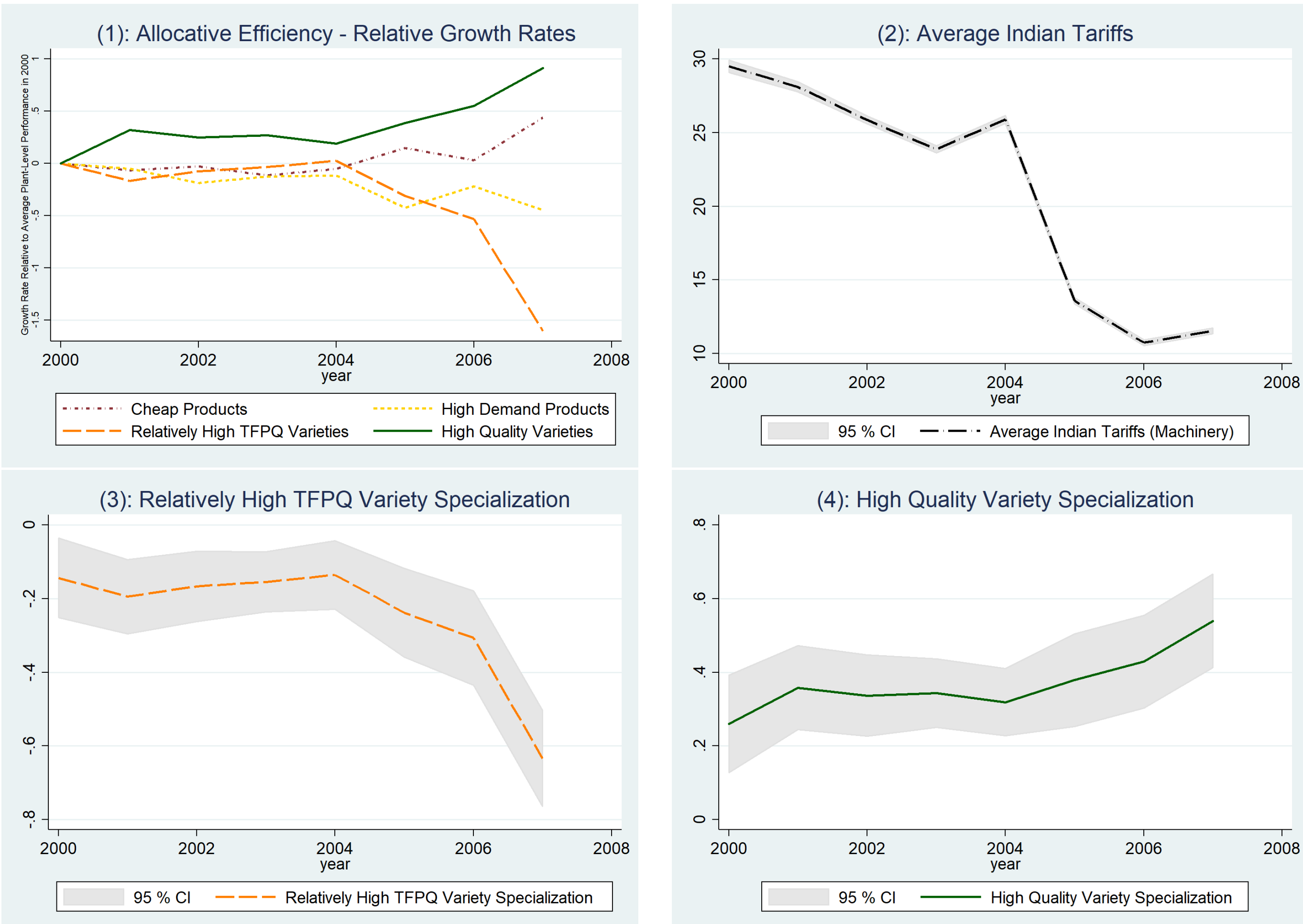
where $\overline{OP}_t\left(\vec{S}_{it}, \vec{X}\right) \equiv \frac{1}{N_t}\sum_{i=1}^{N_t} OP_{it}\left(\vec{S}_{it}, \vec{X}\right)$. As a result, the yearly value of each series can be interpreted as the proportion of average firm-level performance growth attributable to that form of product-level specialization.

While the average values of the cross-product specialization terms do not change all that much over the sample period, post-2004 there is a very large increase in the average degree of within product code specialization. In particular, *plants appear to be specializing much more in relatively high quality varieties, at the expense of their relatively high TFPQ varieties*. Note that this tendency for plants to specialize in high quality varieties at the expense of their relatively high TFPQ varieties can also be seen in the average levels of the within product code allocative efficiency terms, which I plot in panels (3) and (4) in Figure 2. In particular, $\overline{OP}_t\left(\vec{S}_{it}, \vec{\widetilde{\omega}}_{it}^{j|g}\right)$ is always negative, while $\overline{OP}_t\left(\vec{S}_{it}, \vec{\widetilde{\eta}}_{it}^{j|g}\right)$ is always positive, implying that plants tend to allocate more inputs to relatively high-quality varieties, and less

[84] In particular, note that prices will also be measured in different units across product codes, and thus the average level of the demand shifter within a product-code suffers from the same problems as across product code TFPQ comparisons. Note, however, that the *sum* of these two terms will be in (log) revenue TFP units, which does not suffer from the units problem these two terms suffer from on their own.

inputs to cheap varieties.

Figure 2: Aggregate Trends in Within-Firm Allocative Efficiency and Tariffs



Notes: Panel (1) plots the relative growth rates of $\overline{OP}_t\left(\vec{S}_{it},\vec{\tilde{\omega}}^g_t\right)$ (Specialization in Cheap Products), $\overline{OP}_t\left(\vec{S}_{it},\vec{\tilde{\eta}}^g_t\right)$ (Specialization in High Demand Products), $\overline{OP}_t\left(\vec{S}_{it},\vec{\tilde{\omega}}^{j|g}_{it}\right)$ (Specialization in Relatively High TFPQ Varieties) and $\overline{OP}_t\left(\vec{S}_{it},\vec{\tilde{\eta}}^{j|g}_{it}\right)$ (Specialization in High Quality Varieties). Note that growth rates in Panel (1) are normalized relative to the average value of plant-level performance, $\overline{h}_t$, in the year 2000- See Equation (42). Panel (2) plots the average tariff rate for all products within the Machinery, Equipment, and Parts Industry over time. Panel (3) plots average value of plant-level specialization in relatively high TFPQ varieties over time, $\overline{OP}_t\left(\vec{S}_{it},\vec{\tilde{\omega}}^{j|g}_{it}\right)$), while Panel (4) plots the average value of plant-level specialization in relatively high quality varieties, $\overline{OP}_t\left(\vec{S}_{it},\vec{\tilde{\eta}}^{j|g}_{it}\right)$. Shaded area in panels (2), (3), and (4) plot the 95% confidence intervals, taking the demand and production function estimates as data.

Note that the magnitude of the shift toward quality (and away from TFPQ) is sizeable. In particular, the growth rates in Figure 2 imply that the shift toward quality from 2000-2007 *would have increased average plant level-efficiency by close to 100 %*, if it were possible to shift resources into high quality goods without also implying a shift *away* from high TFPQ goods. Note, however, since high quality varieties tend to have low TFPQ, as per Figure 1, the shift towards quality generates allocative efficiency *losses* that actually dominate the TFPR gains generated by the increase in quality specialization.

These overall trends in specialization over time naturally leads to the question - is quality specialization driven by trade liberalization? As Panel (2) of Figure 2 demonstrates, to a first approximation this appears to be the case, as the dramatic increase in quality specialization appears to occur around a round of dramatic tariff cuts passed in India in 2004, with average tariffs falling by more that 65 %. To obtain more convincing evidence on whether tariff cuts are driving the within-firm shifts towards quality, I regress these measures of within-plant specialization on the average of the tariff rates for each 5-digit product produced by the plant and a series of year and plant fixed effects. These results are reported in Table 8, below.

Table 8: Tariffs And Within-Firm Efficiency

	(1) Quality Specialization	(2) Relative TFPQ Specialization	(3) Relative TFPR Specialization
Tariffs_{it}	-0.0138*	0.0133*	-0.000511
	(0.00758)	(0.00689)	(0.00595)
Observations	1,747	1,747	1,747

Standard errors clustered by plant

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

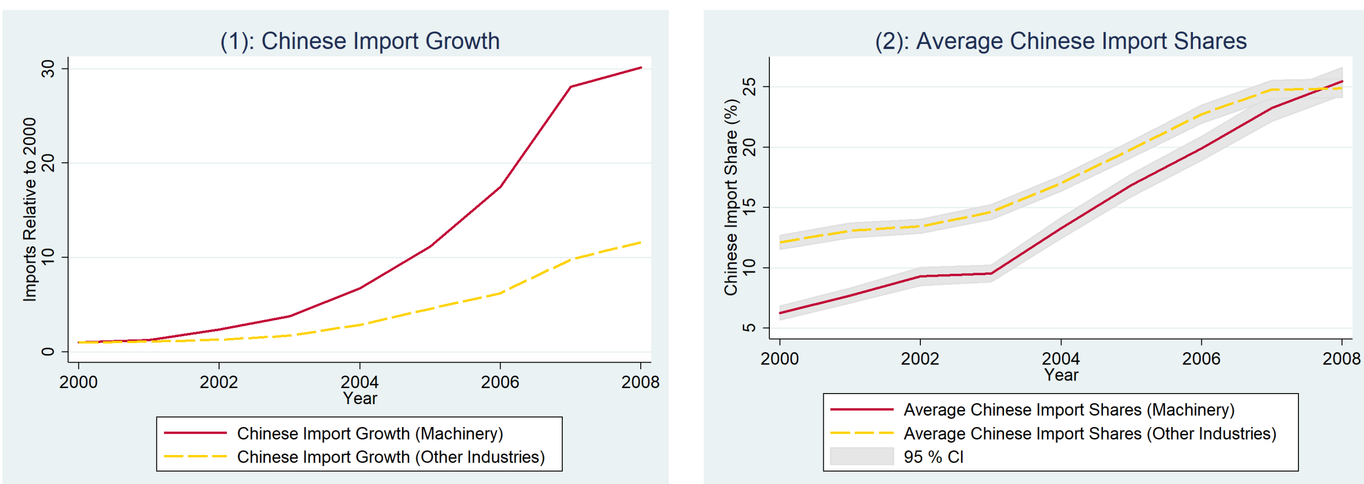
Notes: Each column corresponds to an OLS regression of the listed outcome variable on firm-level average tariffs as described in the main text.All regressions include year and plant fixed effects. Quality Specialization is measured by $OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}}_{it}^{j|g}\right)$, while Relative TFPQ Specialization is measured by $OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\omega}}_{it}^{j|g}\right)$ - see equation (41). Relative TFPR specialization measured by $OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}^{j|g}\right)$. Tariffs_{it} measured as firm-level average of tariffs applied to each product-code produced within a plant. Standard errors treat demand and production function estimates as data.

The first two columns of Table 8 are consistent with tariff cuts leading to increased quality specialization at the plant-level, while decreasing specialization in relatively high TFPQ goods. We can determine the overall effect on plant level efficiency in the third column, where I report a similar regression for plant-level TFPR specialization, or $OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}^{j|g}\right)$. Since $h_{it}^{j|g} = \widetilde{\omega}_{it}^{j|g} + \widetilde{\eta}_{it}^{j|g}$, the tariff coefficient on this regression will equal the sum of the coefficients in columns 1 and 2. In general, these point estimates imply that the shift towards quality induced by tariff cuts was *revenue productivity neutral*, in the sense that it did not appear to significantly change overall plant-level performance as measured by $h_{it} = \sum_{j \in \mathbb{Y}_{it}} S_{it}^{jg} h_{it}^{jg}$.

Note, however, that the standard errors in Table 8 are somewhat large, leading to point estimates for quality and TFPQ specialization that are only statistically significant at the 10 % level. This is largely due to the fact that there is not a lot of cross-product code variation in tariff levels, as one can see from the extremely tight confidence intervals for average Indian tariffs plotted in Figure 2. Moreover, issues related to statistical inference aside, there are two further problems related to using tariffs to measure the degree of import competition. First, tariff levels will not directly capture differences in the degree of foreign competition across markets, since different product codes, conditional on a tariff level, will have different numbers of foreign competitors. Secondly, tariff levels are likely to be chosen in India *in* *response* to product-level performance, perhaps due to political economy considerations, as discussed in Trefler (1993), as well as Topalova and Khandelwal (2011) who specifically look at tariff changes in response to industry performance in India. As a result, these estimates may not reflect a truly causal effect of tariff changes on firm-level efficiency.

To deal with these concerns, I take advantage of an alternative source of variation in import competition within India: the emergence of China during the 2000s as a major trading partner. As discussed in Autor et al. (2013), the emergence of China as a major player in international markets following their ascension to the WTO in 2001 was largely driven by efficiency improvements *within China*, implying that a significant portion of Chinese export growth is plausibly exogenous to industry performance in India. More importantly, as I show in Figure 3, the increase in Chinese import competition experienced within the Machinery, Equipment, and Parts industry in India was extremely large. While India imported just over 5% of their machinery, equipment, and parts, from China in the early 2000s, by the the end of my sample period close to 25 % of all imports in this industry come from China. Moreover, Chinese imports increased by approximately a factor of 30 in this industry, which is roughly three times the increase in Chinese imports experienced in other Indian industries.[85]

Figure 3: Chinese Imports in India: Machinery, Equipment, and Parts vs. Other Industries



Notes: Chinese import growth rates are deflated by the Indian wholesale price index. Average Chinese import shares calculated as the simple average of import shares over each 5-digit ASICC codes. See Data Appendix B3 for details on the construction of imports by product-code.

To determine whether this significant increase in Chinese import competition led to increased plant-level allocative efficiency, I estimate the following model:

$$\text{Specialization}_{it} = \beta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right) + \alpha_t + \gamma_i + \epsilon_{it}, \tag{43}$$

where $\text{Specialization}_{it}$ is a measure of plant-level specialization or allocative efficiency, $\text{IM}^{g}_{\text{IND,CHN},t}$ de-

[85]While these rates of import growth are enormous, one should bear in mind that both China and India experienced massive GDP growth over this time period, with India's GDP more than doubling and China's increasing by a factor of 4, while more than half of Chinese export growth in the 2000s was driven by machinery exports (Berger and Martin (2013)).

notes total Indian imports from China in product code g at time t, Λ_{it} denotes the set of product codes produced by firm i, G_{it} is the number of unique 5-digit product codes produced by firm i, and α_t and γ_i denote year and plant fixed effects, respectively.

Note that simply estimating (43) by OLS is likely to *underestimate* the effect of Chinese import exposure on specialization, as one would expect Chinese imports to rise *more* for product codes where domestic producers are not performing as well, potentially due to lower degree of specialization.[86] To deal with this source of bias, I estimate (43) by 2SLS, instrumenting $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$ with total imports from China in other low to middle income countries *besides* India.[87] More formally, for each plant I construct the instrument:

$$Z_{it}^{\text{CHN}} = \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\sum_{k\in\mathbb{C}^{\text{No India}}}\text{IM}^{g}_{\text{k,CHN},t}\right), \tag{44}$$

where $\mathbb{C}^{\text{No India}}$ denotes the set of low to middle income countries *excluding India*.[88] Note that Z_{it}^{CHN} should function as a valid instrument as long as Chinese import growth in other countries is not driven by factors determining plant-level specialization in India.[89]

The OLS and IV results for this estimation strategy can be found in Tables 9 and 10, below, where I consider as outcomes the various measures of within-product code measures of plant-level allocative efficiency described earlier. Overall, the qualitative story told in Tables 9 and 10 is similar: Plants tend to specialize more in the production of relatively high quality outputs, at the expense of their relatively high TFPQ varieties. Overall, the net effect of plant-level performance, as measured by columns (3) and (4), respectively, indicates that this pattern of specialization led to small loses in plant-level performance, although these changes are not statistically different from zero. Comparing the IV and OLS estimates, we see that accounting for the endogeneity of imports appears to be correcting the expected downward bias, as the point estimates for all measures of plant-level specialization are substantially larger than the OLS estimates. In columns (4) through (6) of Table 9 and columns (6) to (8) of Table 10, I examine whether this change in specialization is driven by changes in the the number of individual varieties produced, or number of distinct product codes produced, by including

[86]Note, of course, that since I include plant fixed effects in this specification, for the bias to work in this direction it would have to be that lower performing product codes had lower *growth rates* in specialization, compared to other industries.

[87]I only include low to middle income countries in this set as Chinese trade to high income countries is likely to depend on very different factors than those determining trade between other low and middle income countries, which would tend to decrease instrument strength. This is analogous to the strategy in Autor et al. (2013), who when constructing a similar instrument for Chinese import exposure in the United States, only include Chinese exports to other high income countries. The list of WTO low and middle income countries can be found at http://wits.worldbank.org/referencedata.html.

[88]When mapping Chinese imports in other countries from 4-digit HS codes to 5-digit ASICC codes, I do not allocate imports across ASICC codes using the import weights discussed in Appendix B3, since using these import weights would mechanically generate higher Chinese export growth in AISCC codes that tend to be imported the most *in India*, which is exactly the sort of endogenous variation in Chinese imports to India that I am trying to avoid with this instrument. Hence, each 4-digit HS code that maps to a 5-digit ASICC code is given equal weight in the construction of the instrument.

[89]Note that I allow the set of products produced by the firm to vary with my instrument. This is because the level of Chinese exports in newly products added products may directly affect the degree of specialization within the plant. I return to this point in Section 6.4 and Appendix G1.

number of variety and number of product code fixed effects. Note that the inclusion of these controls barely changes the estimated magnitudes, implying that the changes in specialization I am observing are generally not due to systematic differences in the degree of specialization for firms that either produce more varieties or distinct product codes.

Table 9: Import Competition And Within Plant Efficiency (OLS)

	(1) Quality Specialization	(2) Relative TFPQ Specialization	(3) Relative TFPR Specialization	(4) Quality Specialization	(5) Relative TFPQ Specialization	(6) Relative TFPR Specialization
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^g_{\text{IND,CHN},t}\right)$	0.0166	-0.0306***	-0.0140	0.0105	-0.0252**	-0.0147
	(0.0130)	(0.0113)	(0.00900)	(0.0126)	(0.0113)	(0.00932)
Product and Variety FE?	No	No	No	Yes	Yes	Yes
Observations	1,712	1,712	1,712	1,710	1,710	1,710

Standard errors clustered by plant

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

Notes: Each column corresponds to an OLS regression of the listed outcome variable on firm-level import competition as described in the main text. All regressions include year and plant fixed effects, while columns (4) through (6) also include number of product code and number of variety fixed effects. Quality Specialization is measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\eta}^{\vec{j}|g}_{it}\right)$, while Relative TFPQ Specialization is measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\omega}^{\vec{j}|g}_{it}\right)$ - see equation (41). Relative TFPR Specialization measured by $OP_{it}\left(\vec{S}_{it}, h^{\vec{j}|g}_{it}\right)$. Standard errors treat demand and production function estimates as data.

Table 10: Import Competition And Within Plant Efficiency (First Stage and IV)

	(1) $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^g_{\text{IND,CHN},t}\right)$	(2) Quality Specialization	(3) Relative TFPQ Specialization	(4) Relative TFPR Specialization	(5) $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^g_{\text{IND,CHN},t}\right)$	(6) Quality Specialization	(7) Relative TFPQ Specialization	(8) Relative TFPR Specialization
Z^{CHN}_{it}	0.794***				0.767***			
	(0.103)				(0.103)			
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^g_{\text{IND,CHN},t}\right)$		0.0856**	-0.119***	-0.0335		0.0890**	-0.117***	-0.0276
		(0.0381)	(0.0365)	(0.0259)		(0.0398)	(0.0386)	(0.0272)
Product and Variety FE?	No	No	No	No	Yes	Yes	Yes	Yes
Observations	1,712	1,712	1,712	1,712	1,710	1,710	1,710	1,710

Standard errors clustered by plant

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

Notes: Columns (1) and (4) report the first stage regression for the instrumental variable regressions discussed in the main text. AP F-statistic for excluded instrument equals 58.95 for column (1) , and 55.31 in column (4). Columns (2) through (4) and (6) through (8) correspond to an instrumental variables regression on the listed outcome variable, where $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^g_{\text{IND,CHN},t}\right)$ is instrumented by Z^{CHN}_{it} , defined by equation (44). All regressions include year and plant fixed effects, while columns (5) through (8) also include number of product code and number of variety fixed effects. Quality Specialization is measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\eta}^{\vec{j}|g}_{it}\right)$, while Relative TFPQ Specialization is measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\omega}^{\vec{j}|g}_{it}\right)$ - see equation (41). Relative TFPR Specialization measured by $OP_{it}\left(\vec{S}_{it}, h^{\vec{j}|g}_{it}\right)$. Standard errors treat demand and production function estimates as data.

Note that the magnitude of the IV estimates in Table 10 are quite large. In particular, the point estimates in column (6) of Table 10 imply that a doubling of Chinese imports would increase high quality specialization by around 0.062 units. Since increases in quality specialization directly increase plant-level TFPR by the same amount according to equations (39) and (41), this implies an increase in plant-level TFPR due to quality specialization of around 20 % of the average value of plant-level TFPR in 2000.[90] These magnitudes are such that the increases in Chinese import competition observed in the data may actually explain the general trends in quality specialization documented in Figure 2. In particular, since Chinese imports rose over the sample period by more than a factor of 25 (See Figure 3), this would imply a change of allocative quality efficiency of around 0.286 units which slightly larger than the observed increase allocative quality efficiency documented in Figure 2 from 2000 to 2007.

Note, however, that since quality is negatively correlated with TFPQ, these reallocations will also tend decrease plant-level TFPR, as more inputs are also moved to high TFPQ production lines. In particular, according to column (7) of Table 10 a doubling of Chinese imports also tends to decrease plant-level TFPR by 0.081 units, or 26 % of the average value of average plant-level TFPR in 2000, through decreased specialization in relatively high TFPQ varieties. Note, however, that the net effect of the reallocations, as described by column (4), is not statistically different from zero, implying that overall these reallocations are approximately *TFPR neutral*.

Note, however, that the estimates described in Table 10 only consider the *within-product code* effect of reallocations on plant-level TFPR. While multi-product plants tended to reallocate their inputs towards product lines with slightly lower TFPR, compared to other plants producing varieties within the same product code, note that most within-plant reallocations will occur *across product codes*, and therefore the overall effect of reallocations on plant level TFPR will depend on whether the product codes a plant is specializes in have relatively higher or lower values of TFPR.

In Table 11, below, I consider the overall effect of these within-plant reallocations on plant-level TFPR. Column (1) examines product-code level specialization, by changing the outcome variable to $OP_{it}\left(\vec{S}_{it}, \vec{h}_t^g\right) = OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\omega}}_t^g\right) + OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}}_t^g\right)$.[91] This regression indicates that plants reallocated their inputs towards product codes with higher TFPR, on average, in response to Chinese import competition. In column (2), I examine the net effect of these reallocations on plant-level TFPR by changing the outcome variable to $OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}\right)$. These results, by construction, will equal the *sum* of the effects in column (4) of Table 10 and column (1) of Table 11 (see equation (41)). Overall, these reallocations appear to have slightly increased TFPR, although these changes are not statistically significant.

The general picture obtained from the effect of plant-level reallocations is they are generally *TFPR-neutral*, although this net effect masks the huge changes in the gross *composition* of a firm's output set described in columns (6) and (7) of Table 10. Note, however, that the estimates in column (2) only account for plant-level changes in efficiency due to reallocations. Overall plant-level TFPR gains will also depend on changes in the unweighted average value of plant-level TFPR, $\overline{h}_{it}$, as described by equation (39). Hence, in column (3), I change the outcome variable to the unweighted average of plant-level TFPR, finding that Chinese import competition did not lead to any statistically significant

[90]Note that, following Pavcnik (2002), all product-level TFPR values are measured *relative* to the unweighted average of TFPR across products in the first year of the sample.

[91]I consider the sum of the TFPQ and demand specialization effects because $h_t^g = \widetilde{\omega}_t^g + \widetilde{\eta}_t^g$ is measured in TFPR units (see equation (35)), and therefore one can easily make across product code comparisons. On the other hand, since different product codes are measured in different units (e.g. pounds versus units sold), differences in $\widetilde{\omega}_t^g$ or $\widetilde{\eta}_t^g$ can be driven by simple differences in units, making the interpretation of these terms problematic on their own.

changes in this variable- in fact, the point estimates are actually consistent with TFPR decreases.[92] The overall effect of Chinese import competition on plant-level TFPR is then described in column (4), where the outcome variable is simply the sum of the outcome variables in columns (4) and (5), i.e. $h_{it} = \overline{h}_{it} + OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}\right)$.

Table 11: Decomposing Impact of Chinese Imports of Firm Performance

	(1) Product Code Specialization	(2) TFPR Specialization	(3) Average Performance	(4) Total Performance	(5) Standard Firm TFPR	(6) Log Total Plant Revenue
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.0525**	0.0190	-0.0900	-0.0710	-0.0928	0.00693
	(0.0220)	(0.0159)	(0.0867)	(0.0808)	(0.0596)	(0.0281)
Observations	1,712	1,712	1,712	1,712	1,712	1,712

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the listed outcome variable, where $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by Z^{CHN}_{it} , defined by equation (44). All regressions include year, plant, number of product code and number of variety fixed effects. Product code specialization is measured by $OP_{it}\left(\vec{S}_{it}, \vec{h^{g}_{t}}\right) = OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\omega}^{g}_{t}}\right) + OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}^{g}_{t}}\right)$, while total specialization is measured by $OP_{it}\left(\vec{S}_{it}, \vec{h_{it}}\right) = OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}^{j|g}_{it}}\right) + OP_{it}\left(\vec{S}_{it}, \vec{\widetilde{\eta}^{j|g}_{it}}\right) + OP_{it}\left(\vec{S}_{it}, \vec{h^{g}_{t}}\right)$. Average TFPR is the plant-level unweighted average of variety-level TFPR, $\overline{h}_{it}$, while total TFPR is given by $h_{it} = \overline{h}_{it} + OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}\right)$. Standard Firm TFPR is measured as $\ln\left(\sum_{j\in\mathbb{Y}_{it}} P^{j}_{it}Q^{j}_{it}\right) - \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\left(\beta_L l_{it} + \beta_K k_{it} + \beta_M m_{it}\right)$. Standard errors treat demand and production function estimates as data.

Overall, *there is no statistically significant change in plant-level TFPR*. Hence, standard approaches to examining changes in plant-level performance, may conclude Chinese competition has no effect, or potentially decreases, plant-level performance. To see this, note that one would actually find small within-plant productivity losses using standard methods, as I show in column (5), where I change the outcome variable to the standard measure of plant-level TFPR one would use if one did not have plant-product level information, i.e. Plant TFPR $\equiv \ln\left(\sum_{j\in\mathbb{Y}_{it}} P^{j}_{it}Q^{j}_{it}\right) - \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\left(\beta_L l_{it} + \beta_K k_{it} + \beta_M m_{it}\right)$.[93] Here, I obtain a decrease in standard plant-level TFPR with a similar magnitude - although the estimate becomes statistically significant at the 10% level. As a result, with standard methods, one may conclude that import competition decreases plant-level performance. This would misleading, however, as we know from columns (1) and (2) of Table 10 that plants are generally improving their overall output quality bundle, by allocating more resources to their relatively highest quality goods. However, producing an

[92] Note that one obtains very similar results even if one estimates the production function after allowing trade to endogenously change the productivity process. See Appendix F.

[93] Note that standard measures of plant-level TFPR will implicitly incorporate within-plant specialization effects. To see this, note that since $S^{jg}_{it} = \frac{X^{jg}_{it}}{X_{it}}$ for $X = (L, K, M)$, then one can use the revenue production function, (76), to show that $\frac{P^{jg}_{it}Q^{jg}_{it}}{\left(L^{\beta_L}_{it} M^{\beta_K}_{it} M^{\beta_M}_{it}\right)^{\frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}}} = \left(S^{jg}_{it}\right)^{\frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}(\beta_L+\beta_K+\beta_M)} \exp\left(\widetilde{\omega}^{jg}_{it} + \widetilde{\eta}^{jg}_{it}\right)$. Summing over all $j \in \mathbb{Y}_{it}$ and then taking logs yields:

$$\ln\left(\sum_{j\in\mathbb{Y}_{it}} P^{j}_{it}Q^{j}_{it}\right) - \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\left(\beta_L l_{it} + \beta_K k_{it} + \beta_M m_{it}\right) = \ln\left(\sum_{j\in\mathbb{Y}_{it}} \left(S^{jg}_{it}\right)^{\frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}(\beta_L+\beta_K+\beta_M)} \exp\left(\widetilde{\omega}^{jg}_{it} + \widetilde{\eta}^{jg}_{it}\right)\right)$$

output bundle with greater high quality goods generally involves a greater proportion of goods with lower TFPQ, generating a net effect that is approximately TFPR neutral. Interestingly as I show in column (6) import competition shocks appear to have no statistically significant effect on overall plant-level revenues. Since increased Chinese import competition will tend to decrease overall plant level revenues due to competition effects, for revenues to remain constant, it must be the case that these reallocations towards a plant's relatively higher quality goods are allowing the plant to *shield* itself from the increased Chinese import competition.

The key takeaway from these results is that within-plant responses to Chinese import competition in this industry are not so much productivity increasing, as has been emphasized by much of theoretical literature such as Bernard et al. (2011), and Mayer et al. (2014), Mayer et al. (2016), but rather involve reallocations towards a plant's relatively highest quality goods. Since quality reallocations are costly, and therefore partly imply revenue TFP loses due to the fact that quality and TFPQ are negatively correlated, these reallocations generally do not improve plant-level performance using standard TFPR measures. On the other hand, since plant-level revenues are not decreasing in response to the increases in Chinese import competition, these reallocations must be generating some gains for the plants- in particular, the increased quality specialization is likely allowing to maintain approximately constant revenues in the face of increased Chinese import competition.

6.3 Robustness

In this section, I consider the robustness of the result that Chinese import competition leads plants to specialize more in the production of high quality varieties, at the expense of the their high TFPQ varieties. In particular, I show that these results are also robust to including potentially vertically integrated plants, as well as alternative measures of Chinese import competition at the product level. Similarly, I show that these results also are also robust to alternative methods for production function estimation, as well reasonably robust to an alternative approach to estimating within plant reallocations that allows from differences in materials price and wages across production lines.

6.3.1 Potentially Vertically Integrated Plants

Since the approach to estimating multi-product firms developed in this paper is not appropriate for vertically integrated multi-product producers that sell some of their inputs, in the baseline estimates described in Table 10 I drop all plants that produce outputs sets that could potentially indicate vertical integration.[94] Note, however, that this may be throwing away important information. In particular, even though these plants produce output sets that *could* be vertically integrated, they may still be treating these product sets as separate production lines. For example, Atalay et al. (2014) find that only around one half of vertically integrated firms ship upstream inputs to their downstream plants. As a result, simply dropping all potentially vertically integrated plants may be overly conservative. In Table 12, below, I re-estimate the baseline specialization regressions without dropping potentially

[94] For details on the construction of the potentially vertically integrated variable, see Appendix B.

vertically integrated plants.

In general, the qualitative results in Table 12 are not all that different from the baseline estimates in Table 10. While the the magnitude of the IV regressions are slightly smaller with vertically integrated firms, either indicating that this subpopulation is less affected by import competition, or vertical integration itself is leading to measurement error in the degree of specialization. That said, the fact that including this subsample also leads to point estimates consistent with increased quality specialization (and less TFPQ specialization) is reassuring.

Table 12: Import Competition And Within-Firm Efficiency (First Stage and IV)

	(1) Chinese Imports	(2) Quality Specialization	(3) Relative TFPQ Specialization	(4) Relative TFPR Specialization
Z_{it}^{CHN}	0.780***			
	(0.0744)			
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$		0.0581*	-0.0662**	-0.00812
		(0.0326)	(0.0332)	(0.0232)
Observations	2,510	2,510	2,510	2,510

Standard errors clustered by plant

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

Notes: Column (1) reports the first stage regression for the instrumental variable regressions discussed in the main text. AP F-statistic for excluded instrument equals 109.93. Columns (2) through (4) correspond to an instrumental variables regression on the listed outcome variable, where $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by Z_{it}^{CHN} , defined by equation (44). All regressions include year, plant, number of product code and number of variety fixed effects. Quality Specialization is measured by $OP_{it}\left(\vec{S}_{it},\tilde{\eta}_{it}^{\vec{j}|g}\right)$, while Relative TFPQ Specialization is measured by $OP_{it}\left(\vec{S}_{it},\tilde{\omega}_{it}^{\vec{j}|g}\right)$ - see equation (41). Relative TFPR Specialization measured by $OP_{it}\left(\vec{S}_{it},h_{it}^{\vec{j}|g}\right)$. Standard errors treat demand and production function estimates as data

6.3.2 Alternative Measures of Chinese Import Competition

In Table 13, below, I consider two alternative measures of Chinese import competition: the average value of log chinese imports by product, and the the revenue weighted sum of log Chinese imports by product. This is in part to verify that a) there is nothing special about the the particular way I am defining plant-level import exposure that is leading to my results and b) to examine whether there are different effects depending on whether a plant faces more competition is it's high-revenue product lines. Since this measure of plant level import competition is scaled slightly differently, I instead use $Z_{it}^{\text{AltCHN}} = \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}}\ln\left(\sum_{k\in\mathbb{C}^{\text{No India}}}\text{IM}^{g}_{k,CHN,t}\right)$ as an instrument. In general, these results are fairly similar to the baseline, with the revenue weighted measure of import competition leading to slightly larger specialization effects.

Table 13: Import Competition And Within-Firm Efficiency (First Stage and IV)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Quality	Relative TFPQ	Relative TFPR		Quality	Relative TFPQ	Relative TFPR
	$\frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}}\ln\left(\text{IM}^g_{\text{IND,CHN},t}\right)$	Specialization	Specialization	Specialization	$\sum_{j\in\mathbb{Y}_{it}}\alpha^{jg}_{it}\ln\left(\text{IM}^g_{\text{IND,CHN},t}\right)$	Specialization	Specialization	Specialization
Z^{AltCHN}_{it}	0.731***				0.702***			
	(0.0986)				(0.117)			
$\frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}}\ln\left(\text{IM}^g_{\text{IND,CHN},t}\right)$		0.109**	-0.0951**	0.0137				
		(0.0422)	(0.0398)	(0.0316)				
$\sum_{j\in\mathbb{Y}_{it}}\alpha^{jg}_{it}\ln\left(\text{IM}^g_{\text{IND,CHN},t}\right)$						0.113**	-0.0990**	0.0143
						(0.0467)	(0.0432)	(0.0330)
Observations	1,710	1,710	1,710	1,710	1,710	1,710	1,710	1,710

Standard errors clustered by plant

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

Notes: Columns (1) and (5) report the first stage regression for the instrumental variable regressions discussed in the main text. AP F-statistic for excluded instrument equal 54.90 in column (1) and 36.20 in column (5). Columns (2) through (4) and (6) through (8) correspond to an instrumental variables regression on the listed outcome variable, where $Z^{\text{AltCHN}}_{it} = \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}}\ln\left(\sum_{k\in\mathbb{C}^{\text{No India}}}\text{IM}^g_{k,CHN,t}\right)$. α^{jg}_{it} denotes the revenue share of variety $j\in\Lambda^g_{it}$ sold by firm i. All regressions include year, plant, number of product code and number of variety fixed effects. Quality Specialization is measured by $OP_{it}\left(\vec{S}_{it},\widetilde{\eta}^{\vec{j}|g}_{it}\right)$, while Relative TFPQ Specialization is measured by $OP_{it}\left(\vec{S}_{it},\widetilde{\omega}^{\vec{j}|g}_{it}\right)$ - see equation (41). Relative TFPR Specialization measured by $OP_{it}\left(\vec{S}_{it},h^{\vec{j}|g}_{it}\right)$. Standard errors treat demand and production function estimates as data

6.3.3 Alternative Approaches to Production Function Estimation

In this section, I examine whether alternative approaches to estimating the production function modify my baseline results. In particular, a key assumption used in the production function estimation approach outlined in Section 5, is that productivity follows an exogenous AR(1) process, $\omega^j_{it} = \rho^g_0 + \rho\omega^j_{i,t-1} + \xi^j_{it}$, where ρ^g_0 is product-code specific mean, and ρ is the persistent of productivity shocks. Since this approach directly restricts the manner if which TFPQ may vary across firms and time, and therefore could be driving my results, I consider two alternative estimation strategies that relax these restrictions.

Strategy 1- Endogenous Productivity as in De Loecker (2013) and De Loecker et al. (2016): First, I modify the law of motion to include lagged tariffs and lagged Chinese imports, $\omega^{jg}_{it} = \rho^g_0 + \rho\omega^{jg}_{i,t-1} + \beta_0\text{tariff}^g_{t-1} + \beta_1\ln\left(\sum_{g\in\Lambda_{it}}\text{IM}^g_{\text{China},t-1}\right) + \xi^{jg}_{it}$, leading to the augmented estimating equation:

$$\begin{aligned} y^j_{it} = \rho^g_0 + \rho y^j_{i,t-1} + \beta_L\left(\widehat{l}^j_{it} - \rho\widehat{l}^j_{i,t-1}\right) + \beta_K\left(\widehat{k}^j_{it} - \rho\widehat{k}^j_{i,t-1}\right) \\ + \beta_M\left(\widehat{m}^j_{it} - \rho\widehat{m}^j_{i,t-1}\right) + \beta_0\text{tariff}^g_{t-1} + \beta_1\ln\left(\sum_{g\in\Lambda_{i,t-1}}\text{IM}^g_{\text{China},t-1}\right) + \xi^j_{it}, \end{aligned} \tag{45}$$

which, as before, is estimated using nonlinear GMM.[95]

The rationale for including these extra controls in the law of motion for productivity is to control for changes in plant-level productivity that are directly driven by investment decisions within a plant. Since some recent studies, including Bloom et al. (2016), have found that trade liberalization directly leads to

[95] After including these extra variables in the law of motion, the estimation algorithm proceeds exactly as described for the baseline results described in Section 5, I.e. I also include dummies for the number of products to control for potential economy of scope effects, and de-mean the data within 5-digit product code to deal with the product-code specific intercepts.

plant level productivity improvements through R&D, the evolution of productivity may directly *depend* on how exposed plants are to trade. While ideally one could deal with this by including direct controls for innovation effort in the law of motion, as in Doraszelski and Jaumandreu (2013), who observe R&D expenditures, since I do not have this sort of data, I follow De Loecker et al. (2016) and control for these effects indirectly, by including trade controls in the law of motion.

Strategy 2 -Variety Fixed-Effects as in Blundell and Bond (2000): Rather than directly modelling the evolution of the productivity process, one can also relax the implicit restrictions on productivity by allowing each variety to have its *own* product-specific mean in the AR(1) process, i.e. $\omega_{it}^{jg} = \rho_{0i}^{jg} + \rho\omega_{i,t-1}^{jg} + \xi_{it}^{jg}$, where ρ_{0i}^{jg} is the variety-specific fixed effect. This approach requires that the *persistence* of productivity be the same for all products, as do all of the previous methods, although allowing for variety-specific means may allow for richer dimensions of productivity differences across varieties *within* a product code.

To account for variety-specific fixed effects in the productivity process, I follow Blundell and Bond (2000), and estimate the "rho-differenced" production function model after taking first differences, i.e.:

$$\begin{aligned}\Delta y_{it}^{jg} &= \rho\Delta y_{i,t-1}^{jg} + \beta_L\left(\widehat{\Delta} l_{it}^{jg} - \rho\Delta\widehat{l}_{i,t-1}^{jg}\right) + \beta_K\left(\Delta\widehat{k}_{it}^{jg} - \rho\Delta\widehat{k}_{i,t-1}^{jg}\right) \\ &\quad + \beta_M\left(\Delta\widehat{m}_{it}^{jg} - \rho\Delta\widehat{m}_{i,t-1}^{jg}\right) + \Delta\xi_{it}^{jg},\end{aligned} \tag{46}$$

where Δ is the time-differencing operator. The above model can then be estimated by nonlinear GMM using the $y_{i,t-2}^{jg}$, $l_{i,t-2}^{jg}$, $k_{i,t-2}^{jg}$, and $m_{i,t-2}^{jg}$ as instruments.[96]

Since it is well known that lagged levels of inputs tend to be weakly correlated with (quasi) differences, and therefore simply using the second lags as instruments tends to perform poorly due to finite sample bias, I follow Blundell and Bond (1998) and Blundell and Bond (2000) and use a system GMM estimator to estimate the production function parameters, where I augment the first-differenced moments from (46), with the additional level restrictions $\mathbb{E}\left(\Delta x_{i,t-1}\left(\rho_{0i}^{jg} + \xi_{it}^{jg}\right)\right) = 0$, where x is a product-level output or input.[97] [98]

Results: In Figure 4 and Table 14, I plot the correlation between TFPQ and quality implied by these alternative approaches to production function estimation, as well as the estimated impact of Chinese import competition on relative TFPQ specialization, respectively. Note that these results barely differ from the baseline estimates in Figure 1 and Table 10. The predicted slope between TFPQ and quality in Figure 4 differs in by at most 0.1 units compared to the baseline, and the impact of Chinese import competition in Table 14 differs by, at worst, 0.006 units, compared to my baseline results.

[96] Note that one can also include $k_{i,t-1}$ as an instrument as long as plant-level capital is predetermined, which I do in my actual estimation procedure. I also include ΔZ_t^g and ΔZ_{t-1}^g, the current and lagged differences of my input price instruments, in the instrument set, as the moments should help identify the materials input elasticity.

[97] See Blundell and Bond (2000) for a discussion of the validity of these extra moment restrictions, a sufficient condition for which is that output and input use evolve according to a jointly stationary process.

[98] I use a nonlinear GMM estimator, where I directly minimize the GMM criterion function with respect to each of the structural parameters. Note that this differs slightly from Blundell and Bond (2000), who use a two-step method where in the first-step they recover composites of structural parameters using linear system GMM (e.g. $\beta_l \times \rho$, which acts as the "reduced-form" coefficient on the lagged difference of labour input). Since the structural parameters are over-identified in this case, they then recover estimates of the structural parameters using a second-stage minimum-distance estimator.

Table 14: Import Competition and Within-Firm Efficiency: Productivity Process Robustness

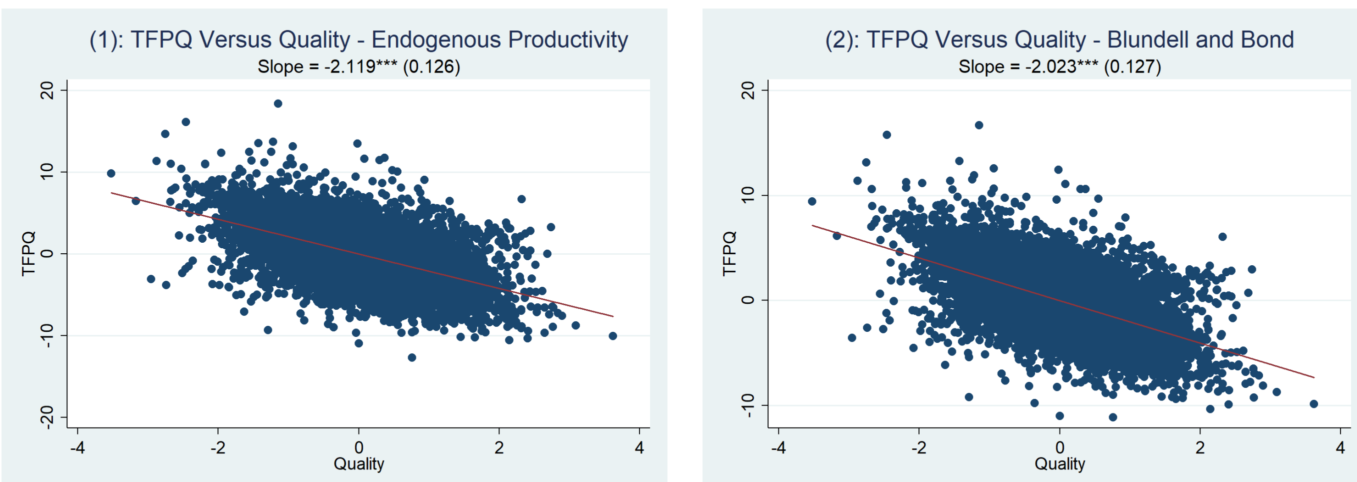
	(Endo Prod) Relative TFPQ Specialization	(Blundell Bond) Relative TFPQ Specialization
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$	-0.115***	-0.110***
	(0.0381)	(0.0357)
Observations	1,710	1,710

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression with $OP_{it}\left(S^{jg}_{it}, \Delta_{gt}\widetilde{\omega}^{jg}_{it}\right)$ as the outcome variable, and $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by Z^{CHN}_{it}, defined by equation (44). All regressions include year, plant, number of product code and number of variety fixed effects. Column (1) uses productivity estimates obtained from the extended endogenous productivity model (45). Column (2) uses productivity estimates obtained from the modified Blundell and Bond (2000) approach to production function estimation described in the main text. Standard errors treat demand and production function estimates as data.

Figure 4: Negative Correlation between TFPQ and Quality: Alternative Methods



Notes: Standard errors of slopes, adjusted for two-way clustering by plant and product code, reported in parentheses. Quality refers to the year-product code demeaned residual from (28), while TFPQ refers to the year-product code demeaned residual from (31). Panel (1) estimates TFPQ using the endogenous productivity model (45), yielding point estimates of $\beta_L = 0.431$, $\beta_K = 0.102$, and $\beta_M = 0.692$. Panel (2) estimates productivity using the modified Blundell and Bond (2000) estimation approach discussed in the main text, yielding point estimates of $\beta_L = 0.671$, $\beta_K = 0.0174$, and $\beta_M = 0.460$. Scatter plots and regressions are at the level of an individual plant-product.

Even though these changes in second stage estimates are fairly small, there are sizeable differences in the production function coefficients, with β_L varying between 0.374 (baseline) to 0.671 (Blundell and Bond), β_k varying between 0.0174 (Blundell and Bond) to 0.102 (endogenous productivity), and β_M varying from 0.460 (Blundell and Bond) to 0.780 (baseline).[99] As a result, it appears that the exact

[99]Note that while the negative point estimate for the Blundell and Bond estimator is problematic from for constructing

method used to identify the production function is not a key driver of my results, as my estimates do not appear to be all that sensitive to estimated production function parameters.

6.3.4 Within-Plant Input Price Variation

Note that my approach to estimating within-plant TFPQ variation and input shares assumes that materials prices are constant across products within a firm. While this assumption is primarily driven by data limitations, note that if unobserved within-plant input price dispersion is driven by differences in *output quality across products*, then I may actually be *overstating* the negative correlation between TFPQ and quality within plants. In particular, recall from Section 3.3.2 that if there is within-plant input price dispersion, then as long as each plant is a price taker in the input markets, then the share of materials going into production line j should be given by:

$$M_{it}^{j} = \frac{\frac{\lambda_{it}^{j} Y_{it}^{j}}{W_{it}^{jM}}}{\sum_{k \in \mathbb{Y}_{it}} \frac{\lambda_{it}^{k} Y_{it}^{k}}{W_{it}^{kM}}} M_{it}. \tag{21}$$

Note that the above implies that input allocation rule (9) will allocate *too many* inputs to production lines with high material input costs, and too few materials to low cost production lines. As a result, this will make TFPQ appear *smaller* in production lines with expensive inputs, and larger in production lines with cheap inputs. If most of this input price variation is driven by quality differences across production lines, in the sense that producing high quality inputs requires more expensive materials, then this unobserved input price variation could largely explain the negative correlation between TFPQ and quality within a plant.

To determine whether this sort of unobserved input price variation is driving my results, I attempt to proxy for the unobserved input price variation within multi-product plants, using information on the input prices charged by single product plants, as proposed in Section 3.3.2. In particular, I use data on output quality, input prices, and location to predict log materials prices in the following OLS regression for the subset of single product plants:

$$\ln(W_{it}^{M}) = \beta_0 \widehat{\eta}_{it}^{j} + \beta_1 \left(\widehat{\eta}_{it}^{j}\right)^2 + \beta_2 \left(\widehat{\eta}_{it}^{j}\right)^3 + \theta_g + \alpha_t + \gamma_s + \epsilon_{it}, \tag{47}$$

where $\widehat{\eta}_{it}^{j}$ is the estimated demand residual from (28) for single product plant i, and θ_g, α_t and γ_s are product code, time, and state fixed effects, respectively.

I then use the out-of-sample predicted values from (22) at the product-level to estimate within-plant inputs prices. Letting $\widehat{W}_{it}^{jM}$ denote the exponentiated predicted price of materials for product $j \in \mathbb{Y}_{it}$, I then allocate materials across production lines using the modified input share formula:

$$\widehat{M}_{it}^{j} = \frac{\frac{\lambda_{it}^{j} Y_{it}^{j}}{\widehat{W}_{it}^{jM}}}{\sum_{k \in \mathbb{Y}_{it}} \frac{\lambda_{it}^{k} Y_{it}^{k}}{\widehat{W}_{it}^{kM}}} M_{it}. \tag{48}$$

TFPQ estimates, the results change very little if I instead treat this as a zero for calculating TFPQ, or as a very small positive number such as 0.01.

Since this changes the quantity of inputs allocated to each production line, I then re-estimate the production function using these new input shares, and calculate the new implied values of product level TFPQ. The modified correlations between quality and TFPQ are reported in columns (3) and (4) of Table 15, below. Unsurprisingly, the negative correlation between TFPQ and quality becomes somewhat smaller, both across products and within-plants. However, the magnitude of the adjustment is not particularly large, as the unadjusted correlation, reported in Columns (1) and (2), only differs from the adjusted correlations by around 0.1 units, which is a difference of around 5% of a standard deviation of TFPQ dispersion.

Table 15: TFPQ and Quality Correlations: Accounting For Within-Firm Price Dispersion

	(1) TFPQ	(2) TFPQ	(3) TFPQ	(4) TFPQ	(5) TFPQ	(6) TFPQ
Quality	-2.313***	-2.908***	-2.274***	-2.794***	-2.260***	-2.766***
	(0.141)	(0.0865)	(0.142)	(0.0783)	(0.142)	(0.0773)
Observations	11,172	11,172	11,172	11,172	11,172	11,172
Plant FE	NO	YES	NO	YES	NO	YES
Materials Price Adjustment	NO	NO	YES	YES	YES	YES
Wage Adjustment	NO	NO	NO	NO	YES	YES

Standard errors clustered separately by plant and product code

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

Notes: Each Column presents an OLS regression of TFPQ on Quality (demeaned within product-code and year), for the subset of products only produced by multi-product plants. Columns (1) and (2) calculates TFPQ using the baseline approach (without accounting for within-plant input price dispersion) described in Section 5. Columns (3) and (4) estimate TFPQ after adjusting material inputs to account for input price dispersion using equation (48). Columns (5) and (6) adjust within firm material and labour inputs to account for input price and wage dispersion. Standard errors treat demand and production function estimates as data.

Since accounting for input price dispersion slightly dampens the negative correlation between TFPQ and quality, it is worth examining whether this adjustment affects my estimates of the impact of Chinese import competition on plant-level specialization. In columns (1) and (2) of Table 16, below, I re-estimate (43) by 2SLS, after re-estimating the production function and modifying the materials input shares to account for within-plant input price dispersion. Note that since materials input shares now *differ* from the labour and capital input shares, which are still given by the baseline formula (9), I can calculate allocative efficiency using either labour/capital input shares, as I do in column (1), or using materials shares, as I do in column (2).

Generally, the qualitative pattern for TFPQ specialization is not all that different from the unadjusted case, although there appears to be a much smaller effect once I measure input shares using materials shares, rather than labour shares.[100] We can see the same pattern in the quality regressions, with the effect becoming smaller once one measures allocative efficiency using materials shares rather

[100] While I lose statistical significance at standard levels for this regression, note that these results are quite close to significant at the 10 % level, with an associated P-value of 0.119

than labour shares, leading to a loss of statistical significance. Since it is not entirely obvious that materials input shares act as a better proxy for efficiency than labour shares, these estimates generally imply that unobserved within-firm price variation is not likely to be a key driver of my baseline results, at least in terms of the allocation of labour and capital across production lines.

Table 16: Import Competition and Within-Firm Efficiency: Input Price Robustness

	(1) Relative TFPQ Specialization	(2) Relative TFPQ Specialization	(3) Relative TFPQ Specialization	(4) Relative TFPQ Specialization	(5) Relative TFPQ Specialization
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	-0.100***	-0.0579	-0.0985***	-0.0560	-0.0961***
	(0.0374)	(0.0377)	(0.0370)	(0.0374)	(0.0364)
	High Quality Specialization	High Quality Specialization	High Quality Specialization	High Quality Specialization	High Quality Specialization
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.0890**	0.0538	0.0890**	0.0538	0.0887**
	(0.0398)	(0.0423)	(0.0398)	(0.0423)	(0.0395)
Observations	1,710	1,710	1,710	1,710	1,710
Materials Price Adjustment	YES	YES	YES	YES	YES
Wage Adjustment	NO	NO	YES	YES	YES
Standard Shares	YES	NO	YES	NO	NO
Materials Shares	NO	YES	NO	YES	NO
Labour Shares	NO	NO	NO	NO	YES

Standard errors clustered by plant

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

Notes: Columns (1) through (5) correspond to an instrumental variables regression on the listed outcome variable, where $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by Z^{CHN}_{it}, defined by equation (44). All regressions include year, plant, number of product code and number of variety fixed effects. Relative TFPQ and High Quality Specialization measured by $OP_{it}\left(\vec{S}^{X}_{it}, \widetilde{\omega}^{\vec{j}|g}_{it}\right)$ and $OP_{it}\left(\vec{S}^{X}_{it}, \widetilde{\eta}^{\vec{j}|g}_{it}\right)$, respectively, where $\vec{S}^{X}_{it}$ denotes the vector input shares across production lines within plant i for input X. Columns (1) and (3) use the share of capital input for the share terms, columns (2) and (4) use the share of materials inputs, while column (5) uses the share of labour inputs. Columns (1) and (2) calculate $\omega^{j|g}_{it}$ after adjusting material inputs to account for material input price dispersion using equation (48). Columns (3) to (5) adjust within plant material and labour inputs to account for material input price and wage dispersion. Standard errors treat demand and production function estimates as data.

As a further robustness check, I also allow for within plant wage dispersion, modifying the above procedure to account for the the plant-level information I have on wages.[101] Implicitly, using this approach to determine within-plant wage dispersion requires that labour be a *static* input. Although this is unlikely to hold exactly in practice, and therefore is an imperfect approach, it is still worth considering whether heterogeneous wages across production lines may explain my results.

[101] Wages are in terms of rupees per manday worked.

However, as one can see in columns (5) and (6) of Table 15, also allowing for within plant wage dispersion barely changes the TFPQ and quality correlations, relative to the case where I only accounted for materials price dispersion. Similarly, the Chinese import competition results reported in columns (3) to (5) of Table 16 are not all that different from the first two columns. Measuring allocative efficiency using materials shares leads to the smaller, and generally statistically insignificant effects, as was the case when I only accounted for materials price dispersion, while using labour or capital shares does not change the estimates all the much, relative to the baseline estimates in Table 10. In general, it appears that accounting for input price dispersion is a second-order concern for the baseline results.

6.4 Extensive Versus Intensive Margin Adjustments

In this section, I examine whether the reallocations towards quality are primarily driven by *intensive margin adjustments*, i.e. the reallocation of inputs between the same set of product codes, or *extensive margin adjustments*, i.e. reallocation of inputs through product adding and dropping. For this purpose, I partition the set of products produced by each firm at time t and $t-1$ into three non overlapping categories: 1) *Dropped products*, denoted by $D_{it} \subset \mathbb{Y}_{i,t-1}$, and corresponding to the set of products that were produced in $t-1$ but dropped at time t, 2) *New products*, denoted by $N_{it} \subset \mathbb{Y}_{it}$, and corresponding to the set products that were produced at time t but not produced at time $t-1$, and 3) *Constant products*, denoted by C_{it}, and corresponding to the set of products that were produced in both $t-1$ and t.[102] Further define the within-category $X \in (N, D, C)$ share as $S_{it}^{jg|X} \equiv \frac{S_{it}^{jg}}{S_{it}^{X}}$, where S_{it}^{X} denotes the total input share of products belonging to category $X \in (N, D, C)$ at time t. I then make use of the following decomposition, similar to that in Melitz and Polanec (2015), for changes in the plant-level OP covariance, the derivation for which can be found in Appendix G:

$$\begin{aligned}\Delta OP_{it}\left(\vec{S}_{it}, \vec{x}_{it}\right) = & \underbrace{\sum_{j \in C_{it}} S_{it}^{jg|C}\, \widehat{x}_{it}^{jg} - \sum_{j \in C_{it}} S_{i,t-1}^{jg|C}\, \widehat{x}_{i,t-1}^{jg}}_{(I)\text{: Intensive Margin}} \\ & \underbrace{+ S_{it}^{D}\left(\sum_{j \in C_{it}} S_{i,t-1}^{jg|C}\, \widehat{x}_{i,t-1}^{jg} - \sum_{j \in D_{it}} S_{i,t-1}^{jg|D}\, \widehat{x}_{i,t-1}^{jg}\right)}_{(D)\text{: Product Dropping}} \\ & \underbrace{+ S_{it}^{N}\left(\sum_{j \in N_{it}} S_{it}^{jg|N}\, \widehat{x}_{it}^{jg} - \sum_{j \in C_{it}} S_{it}^{jg|C}\, \widehat{x}_{it}^{jg}\right)}_{(N)\text{: New Products}},\end{aligned} \tag{49}$$

[102]When a plant drops *varieties* of the same product code g over time, I"drop" the variety with the *most different* value of $\omega_{i,t-1}^{jg}$, when compared with the values of ω_{it}^{jg} observed within product code g in the following period. More formally, if a plant drops V varieties of the same product code g, I calculate the *minimum* pairwise difference between $\omega_{i,t-1}^{jg}$ of each variety $j \in \Lambda_{i,t-1}^{g}$, and the estimated value of ω_{it}^{jg} of each variety $j \in \Lambda_{it}^{g}$. The set of varieties with the V largest values of this minimum pairwise difference are then classified as the "dropped" products. Similarly, if a plant adds N varieties of the same product code, I consider the N varieties with the largest minimum pairwise differences as the new products at time t.

where $\widehat{x}_{it}^{jg} = \left(x_{it}^{jg} - \frac{1}{J_{it}} \sum_{j \in \mathbb{Y}_{it}} x_{it}^{jg} \right)$ for x_{it}^{j} is a particular product-level performance measure, e.g. $x_{it}^{jg} \in (\widetilde{\omega}_t^g, \widetilde{\eta}_t^g, \widetilde{\omega}_{it}^{j|g}, \widetilde{\eta}_{it}^{j|g})$

The intensive margin term (I) in equation (49) is simply an estimate of the difference between the OP covariance term that would have occurred if there was no product adding or dropping. To account for changes in specialization due product switching, the product dropping term (D) then increases as a plants drop products that are relatively low performing compared to their products that are sold in both periods, while the new product term (N) increases as plants add new products that are higher performing than constant product sets. Conditional on dropping a set of relatively low performing products or adding a set of relatively high performing goods, note that that the extensive margin terms (D) and (N) are also increasing in the share of inputs allocated to dropped products and new products, respectively. Together, these two terms capture the portion of the change in the OP covariance that is due to changes in the plant's set of produced products.

I decompose the overall change in plant-level specialization in quality and relatively high TFPQ varieties due to Chinese import competition into extensive and intensive margin components in Tables 17 and 18, below. First, in column (1) I re-estimate (43) in first-differences, rather than with plant-fixed effects, to obtain an estimate of the full effect of Chinese import competition on plant level specialization.[103] In columns (2) through (4), I then change the outcome variable to the intensive margin term (I), the product dropping term (D), and the new product product (N) from (49), respectively. Note that by construction, the point estimates in columns (2) through (4) will sum to the overall effect estimated in column (1). I also provide the same set of regressions with the Chinese import competition variable modified to the average of log imports, in columns (5) through (8), respectively, to see whether slight modifications to the way import competition is measured change the weights put on intensive versus extensive margin effects.

Regardless of the particular method used to measure import competition, Tables 17 and 18 indicate that the vast majority of the reallocations towards quality, and away from relatively high TFPQ products, are due to *product dropping* effects, where plants facing larger increases in import competition are more likely to drop lower quality goods, or drop lower quality goods that previously used a larger share of resources. On the other hand, note that product dropping accounts for a smaller share of the reallocations *away* from high TFPQ varieties, which would be consistent with plants dropping products based on their relatively low position in quality space, rather their relatively high position in TFPQ space. As a result, while the TFPQ reallocations incorporate some sizeable intensive margin effects, accounting for around 30 to 60 % of the total re-allocative effect across the two specifications, there appears to be a much smaller role for intensive margin adjustments in quality specialization. In particular, the estimates in column (5) of Table 17 imply that intensive margin adjustments account

[103] For sample comparability with Table 10, which includes the randomly sampled plants that are not always observed in consecutive periods, the first differences in Tables 17 and 18 are with respect to *plant time*, i.e. for each plant i, $t-1$ refers to the most recent year I observe plant i before time t. To appropriately control for time trends in this specification, I also include current and lagged calendar year fixed effects. A standard first-differences specification (i.e. differences with respect to calendar time), can be found in Appendix G1, where plants with "holes" are dropped. Note that the qualitative results are quite similar across specifications, although the standard first differences estimates in Appendix G1 display larger specialization effects.

for just under 25 % of the reallocations towards quality, while the estimates in column (2) imply that only around 5% of the change in quality specialization is due to intensive margin adjustments.

Table 17: Plant-Level Changes in Quality Specialization: Intensive versus Extensive Margin

Quality	(1) Total Specialization	(2) Intensive Margin	(3) Product Dropping	(4) New Products	(5) Total Specialization	(6) Intensive Margin	(7) Product Dropping	(8) New Products
$\Delta \ln \left(\frac{1}{G_{it}} \sum_{g \in \Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.0815**	0.00552	0.0829**	-0.00685				
	(0.0363)	(0.0349)	(0.0358)	(0.0310)				
$\Delta \frac{1}{J_{it}} \sum_{(j,g) \in \mathbb{Y}_{it}} \ln \left(\text{IM}^{g}_{\text{IND,CHN},t}\right)$					0.0765*	0.0180	0.0867**	-0.0282
					(0.0393)	(0.0334)	(0.0356)	(0.0346)
Observations	989	989	989	989	989	989	989	989

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the described outcome, where $\Delta \ln \left(\frac{1}{G_{it}} \sum_{g \in \Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by $\Delta Z^{\text{CHN}}_{it}$, defined by equation (44), and $\Delta \frac{1}{J_{it}} \sum_{(j,g) \in \mathbb{Y}_{it}} \ln \left(\text{IM}^{g}_{\text{China},t}\right)$ is instrumented by $\Delta Z^{\text{AltCHN}}_{it} = \Delta \frac{1}{J_{it}} \sum_{(j,g) \in \mathbb{Y}_{it}} \ln \left(\sum_{k \in \mathbb{C}^{\text{No India}}} \text{IM}^{g}_{k,CHN,t}\right)$, where Δ is the time differencing operator according to *plant time* (See footnote 103). All regressions include fixed effects for current and lagged calendar year. AP F-statistic for excluded instrument equals 63.85 for columns (1) through (4) and equals 62.19 for columns (5) through (8). Total Specialization refers to $\Delta OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\eta}}^{j|g}_{it}\right)$, while Intensive Margin, Product Dropping, and New Product refer to terms (I), (D) and (N) in equation (49), respectively, with $x^{jg}_{it} = \tilde{\eta}^{j|g}_{it}$. Standard errors treat demand and production function estimates as data.

Table 18: Plant-Level Changes in Relative TFPQ Specialization: Intensive versus Extensive Margin

Relative TFPQ	(1) Total Specialization	(2) Intensive Margin	(3) Product Dropping	(4) New Products	(5) Total Specialization	(6) Intensive Margin	(7) Product Dropping	(8) New Products
$\Delta \ln \left(\frac{1}{G_{it}} \sum_{g \in \Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	-0.108***	-0.0363	-0.0548**	-0.0167				
	(0.0368)	(0.0284)	(0.0275)	(0.0326)				
$\Delta \frac{1}{J_{it}} \sum_{(j,g) \in \mathbb{Y}_{it}} \ln \left(\text{IM}^{g}_{\text{IND,CHN},t}\right)$					-0.0810**	-0.0494*	-0.0511*	0.0195
					(0.0393)	(0.0282)	(0.0274)	(0.0353)
Observations	989	989	989	989	989	989	989	989

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the described outcome, where $\Delta \ln \left(\frac{1}{G_{it}} \sum_{g \in \Lambda_{it}} \text{IM}^{g}_{\text{China},t}\right)$ is instrumented by $\Delta Z^{\text{CHN}}_{it}$, defined by equation (44), and $\Delta \frac{1}{J_{it}} \sum_{(j,g) \in \mathbb{Y}_{it}} \ln \left(\text{IM}^{g}_{\text{China},t}\right)$ is instrumented by $\Delta Z^{\text{AltCHN}}_{it} = \Delta \frac{1}{J_{it}} \sum_{(j,g) \in \mathbb{Y}_{it}} \ln \left(\sum_{k \in \mathbb{C}^{\text{No India}}} \text{IM}^{g}_{k,CHN,t}\right)$, where Δ is the time differencing operator according to *plant time* (See footnote 103). All regressions include fixed effects for current and lagged calendar year. AP F-statistic for excluded instrument equals 63.85 for columns (1) through (4) and equals 62.19 for columns (5) through (8). Total Specialization refers to $\Delta OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\omega}}^{j|g}_{it}\right)$, while Intensive Margin, Product Dropping, and New Product refer to terms (I), (D) and (N) in equation (49), respectively, with $x^{jg}_{it} = \tilde{\omega}^{j|g}_{it}$. Standard errors treat demand and production function estimates as data.

Note that columns (4) and (8) of Tables 17 and 18 indicate that adding new products may have played very little role in these reallocations towards quality. Note, however, that this is not entirely true, as the quality reallocations are partly driven by *product switching* rather than *product dropping*. To see this, in Table 19, below, I change the dependent variable in these regression to changes in the number of varieties produced by a firm. While the point estimates for the effect of Chinese import competition on the number of products are generally negative, these regressions are not statistically significant, implying that many of the plants that are dropping products must also be adding new products.[104] As a result, the product dropping term (D) in (49) may also be increasing due to plants dropping products that previously used larger shares of their resources. In columns (4) and (7) of Tables 19, I find that this is generally the case, by changing the outcome variable to the share of inputs previously allocated to dropped products. In particular, when I measure import competition using the average of log imports within the plant, I find statistically significant increases in the share of inputs previously allocated to dropped products in both columns (4) and (7) of Table 19.[105]

Table 19: Dropping Products versus Product Switching

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Number of Varieties	ΔS^C_{it}	S^N_{it}	$S^D_{i,t-1}$	ΔS^C_{it}	S^N_{it}	$S^D_{i,t-1}$
$\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^g_{\text{IND,CHN},t}\right)$	-0.000375	0.00511	0.0297*	0.0272	0.000895	0.0253*	0.0298*
	(0.0590)	(0.0109)	(0.0177)	(0.0184)	(0.0109)	(0.0150)	(0.0159)
	Number of Varieties	ΔS^C_{it}	S^N_{it}	$S^D_{i,t-1}$	ΔS^C_{it}	S^N_{it}	$S^D_{i,t-1}$
$\Delta \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}} \ln\left(\text{IM}^g_{\text{IND,CHN},t}\right)$	-0.0649	0.000261	0.0386**	0.0425**	0.000688	0.0334**	0.0328**
	(0.0540)	(0.0121)	(0.0170)	(0.0180)	(0.0119)	(0.0143)	(0.0151)
Drop at least one product ?	No	No	No	No	Yes	No	Yes
Add at least one product ?	No	No	No	No	Yes	Yes	No
Observations	989	989	989	989	664	634	628

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the described outcome, where $\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^g_{\text{IND,CHN},t}\right)$ is instrumented by $\Delta Z^{\text{CHN}}_{it}$, defined by equation (44), where Δ is the time differencing operator according to *plant time*. All regressions include fixed effects for current and lagged calendar year. AP F-statistic for excluded instrument equals 63.85. Columns (1) through (4) include all plants, column (5) only includes plants who either drop a product or add a drop, column (6) only includes plants that add a product, and column (7) only includes plants that drop a product. Standard errors treat demand and production function estimates as data.

Finally, Table 19 also considers regressions where the outcome variable is the *change* in inputs

[104]Note that there is a lot of product switching in the data, with 60 % of the observations in these regressions involving plants that both added and dropped products.

[105]While I do not obtain statistical significance at conventional levels in column (4) when I measure import competition using the natural log of average import competition, note that I do obtain statistically significant effects in column (7) (at the 10 % level) when I restrict the sample to only contain plant-years where at least one product is dropped, and the magnitude of the effect in this column is not all that different from that displayed in column (4).

allocated to old products over time (ΔS_{it}^{C}), as well as the share of inputs allocated to new products (S_{it}^{N}), to examine whether the resources freed up by product dropping are being reallocated towards old or new products. In general, I find that increased import competition generally does not change the share of resources allocated to old products, but rather, tends to increase the share of inputs allocated to new products. Interestingly, the share of inputs being allocated towards new goods in response to Chinese import shocks is approximately the same magnitude as the share of inputs previously used in dropped products, implying that the increased quality specialization through product dropping is partly driven by the reallocation of inputs towards *new* product lines, rather than old product lines. Interestingly, this result is largely driven by the level of import competition in the *new* products, rather than changes in import competition for old products. In particular, as I show in Appendix G1, if I measure changes in Chinese import competition holding the product set fixed at its $t-1$ level, i.e. only measure changes in import competition for the dropped and constant product sets, then I no longer find that import competition increases specialization in quality. As a result, it must be that the reallocation towards quality is largely occurring through increased Chinese import competition in new varieties. Note, however, that since these reallocations are generally occurring through *product switching*, this makes sense. In particular, when a plant faces more import competition for its new varieties, it may want to drop one of its other product lines, since this will free up dynamic inputs (e.g. labour and capital) for use in the new production line. This will decrease its marginal costs in the new production line, thereby allowing the plant to price the new product more competitively.

6.5 Alternative Approaches to Recovering Input Shares

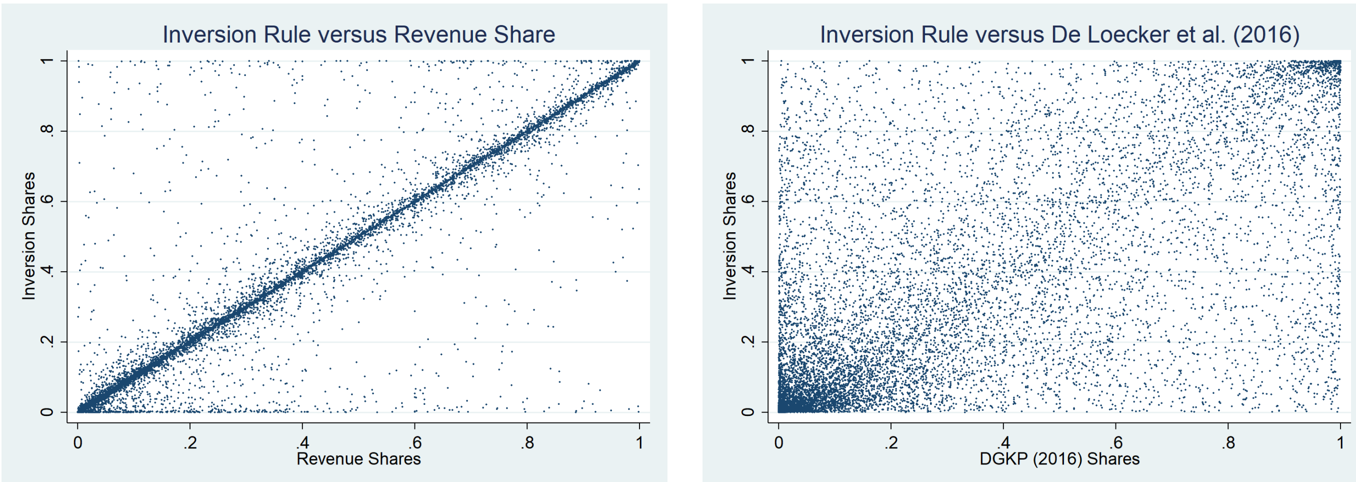
In this section, I compare alternative approaches to recovering input shares to the approach proposed in this paper. This comprises two key approaches - using revenue shares to approximate input shares, as in Foster et al. (2008), or assuming that there is no productivity dispersion within a plant and using this restriction to recover the input shares, as in De Loecker et al. (2016). Note that if one restricts attention to homogeneous of degree $\phi > 0$ production technologies as I have in this paper, then using the approach described by De Loecker et al. (2016) to determine input shares has a simple closed form expression: $S_{it}^{j} = \frac{\left(Y_{it}^{j}\right)^{\frac{1}{\phi}}}{\sum_{k\in\mathbb{Y}_{it}}\left(Y_{it}^{k}\right)^{\frac{1}{\phi}}}$.[106] For easier comparability across methods, I hold production function parameters constant during this excericse, taking ϕ for use in this formula as the sum of β_L, β_K and β_M in Table 5.

Figure 5, below, produces the scatter plot of the inputs shares as recovered using the marginal cost inversion described in this paper, versus those obtained from revenue shares, and those obtained from the De Loecker et al. (2016) approach. Note that revenue shares are much more strongly correlated with the shares obtained in this paper, than the shares obtained from De Loecker et al. (2016). On the other hand, both approaches are still positively correlated with the shares obtained in this paper, with revenue shares and the input inversion shares having a correlation coefficient of around 0.93, while

[106] One obtains this formula by substituting $\frac{X_{it}^{j}}{X_{it}} = S_{it}^{j}$ for each X and j and $\omega_{it}^{j} = \omega_{it}$ for each j into the production technology equation (2), summing over all $j \in \mathbb{Y}_{it}$, then dividing (2) by this sum.

the De Loecker et al. (2016) and input inversion shares have a correlation coefficient of approximately 0.65.[107]

Figure 5: Comparison of Alternative Shares



Notes: Inversion shares are the input shares determined using the marginal cost inversion as described in Appendix C2. DGKP (2016) shares are given by $S_{it}^{j} = \frac{\left(Y_{it}^{j}\right)^{\frac{1}{\phi}}}{\sum_{k \in \mathbb{Y}_{it}} \left(Y_{it}^{k}\right)^{\frac{1}{\phi}}}$.

The fact that revenue shares and the De Loecker et al. (2016) shares do not lead to the exact same estimates of input shares naturally leads to the question of whether these alternative approaches to recovering input shares lead to differences in productivity predictions. In Figure 6, below, I plot the estimated kernel density plot of TFPQ estimated using the three different input share rules, demeaned within 5-digit ASICC code, for all products products by multi-product plants.[108] Surprisingly, the differences in inputs shares matter very little in terms of determining the exact magnitude of TFPQ dispersion.

At face value, this may mean that the exact method used to determine input shares may not matter too much in practice. Note, however, that this is not necessarily true for all questions of substantive interest. In particular, while the approach described in De Loecker et al. (2016) cannot deal directly within changes in plant-level TFPQ specialization, since dispersion in TFPQ is ruled out by construction, one could still use their approach to examine whether there are any changes in *quality* specialization following a trade shock.

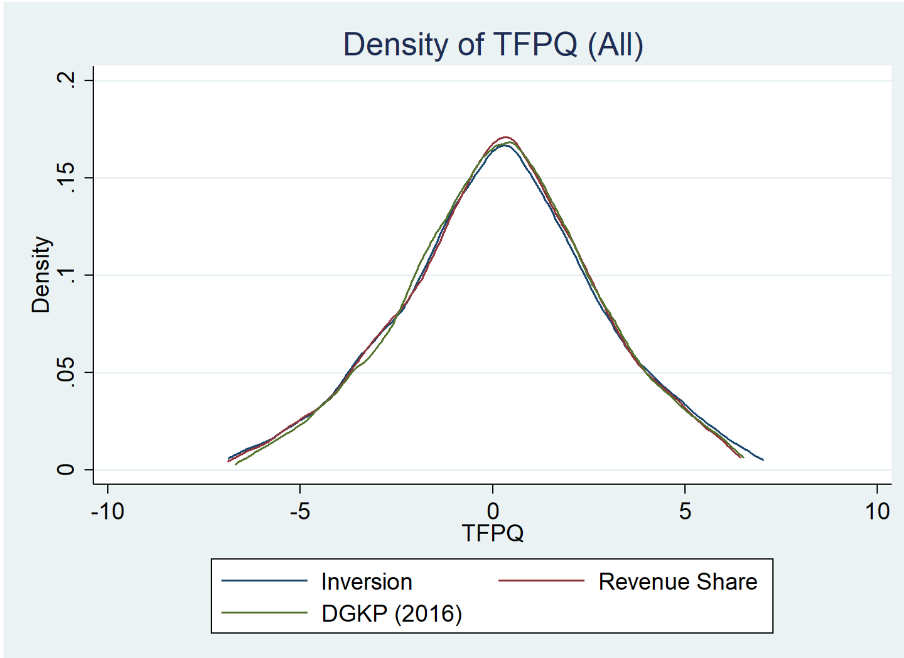
I consider this question in Table 20, below, where in column 1, I consider the same baseline quality specialization regression described in Table 10, with the De Loecker et al. (2016) input shares replacing the marginal cost inversion input shares. Interestingly, I find no significant effect of Chinese import competition on quality specialization in this specification, implying that for some questions, the exact technique used to recover the input shares can indeed matter.[109]

[107] Only multi-product plants are used to construct these statistics, as the correlation coefficient will always be one for the subset of single product plants.

[108] I trim the 1st and 99th percentile of each distribution. Product codes that are only observed once are dropped when estimating the densities.

[109] Note, however, that even if one were to find an effect of quality specialization using this approach, since this approach

Figure 6: TFPQ dispersion: Alternative Shares



Notes: Kernel density plots of TFPQ by product, or $\omega_{it}^{j} \equiv \ln(y_{it}^{j}) - \ln(F(\vec{X}_{it}^{j}))$. Inversion shares are the input shares determined using the marginal cost inversion as described in Appendix C2. De Loecker et al. (2016) shares are given by $S_{it}^{j} = \frac{\left(Y_{it}^{j}\right)^{\frac{1}{\phi}}}{\sum_{k\in\mathbb{Y}_{it}}(Y_{it}^{k})^{\frac{1}{\phi}}}$. TFPQ is demeaned within 5-digit product code. Only multi-product plants are included in the kernel density plot.

Table 20: Plant-Level Changes in Relative Specialization: Alternative Shares

	(IV) Quality Specialization (DGKP Shr)	(IV) Quality Specialization (Rev Shr)	(IV) Relative TFPQ Specialization (Rev Shr)	(IV) Relative TFPR Specialization (Rev Shr)
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.0294	0.107***	-0.0992***	0.00740
	(0.0426)	(0.0378)	(0.0305)	(0.0250)
Observations	1,712	1,712	1,712	

Standard errors clustered by plant

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

Notes: Columns (1) and (4) report the first stage regression for the instrumental variable regressions discussed in the main text. Columns (1) through (3) construct input shares and well as TFPQ using revenue shares. Column (4) determines input shares according to $S_{it}^{j} = \frac{\left(Y_{it}^{j}\right)^{\frac{1}{\phi}}}{\sum_{k\in\mathbb{Y}_{it}}(Y_{it}^{k})^{\frac{1}{\phi}}}$. See the notes on Table 10 for more details.

In columns (2) through (4), I consider the same exercise for quality, TFPQ, and TFPR specialization,

has to abstract from TFPQ dispersion within plant, this approach will like undervalue the *costs* of these reallocations, given the robust negative correlation between TFPQ and quality within plants.

where input shares are estimated using revenue shares, since this approach allows for TFPQ dispersion. In this case, I find very similar effects to the baseline effects described in Table 10, likely due to the fact that revenue shares are much more strongly correlated with the marginal cost inversion shares than the De Loecker et al. (2016) shares. Overall, the general picture is that for some questions, such as the overall magnitude of TFPQ disperson, the exact approach used to recover the input shares does not matter much, although it can matter for some important questions, such as picking up trade induced quality specialization within a plant.

7 Conclusion

In this this paper, I have developed a new approach to estimating product-level TFP for multi-product firms in data sets where the allocation of inputs across production lines is unknown. The approach, based on a standard differentiated goods model of price competition, only requires information on firm-product level prices and quantities to be implemented. Since this information is becoming more widely available in many firm and plant level datasets, the techniques developed here should be able to help researchers interested in firm level productivity tackle a wider class of problems.

I then applied the approach to a panel of manufacturing plants in India from 2000-2007, finding new evidence for the within-firm allocative gains from trade, emphasized by Eckel and Neary (2010), Bernard et al. (2011), Mayer et al. (2014) and Mayer et al. (2016). In particular, I found that increased Chinese import competition generated increased specialization in quality, where firms reallocated a greater portion or their inputs towards relatively higher quality goods, but *away* from their relatively high TFPQ goods. These findings provide new evidence for the quality upgrading effects of import competition, previously explored by Amiti and Khandelwal (2013), although through a new margin: the within-plant reallocation of inputs across high and low quality production line. Since these reallocations towards quality are costly, in the sense that TFPQ is relatively *lower* for these varieties, overall these reallocations tended to decrease plant level TFPQ. As a result, while these reallocations did not generate sizeable productivity gains, they generally implied large gross changes in the composition of a plant's output towards higher quality goods. Given that I also found there to be statistically significant effect of Chinese import competition on overall plant-level revenues, it must be that these quality reallocations in part allowed the plant to shield itself from the increase in Chinese import competition.

The findings of this paper point towards a number of interesting paths for future research. Since reallocations towards quality in this industry are primarily driven by product dropping, it would be useful in future work to obtain a better understanding of what determines a plant's product sets. This would be particularly useful as I also find that product dropping is largely due to increased import competition in *new* varieties, rather than direct import competition in the particular varieties being dropped. This particular channel for generating quality upgrading within the plant, which relies on across-product linkages in the particular product sets, a plant chooses to produce, is largely unstudied. As a result, further explorations along these lines would be valuable.

Moreover the negative correlation between quality and TFPQ uncovered in this paper, which is also documented in Forlani et al. (2016) and Jaumandreu and Yin (2016), is a relatively surprising finding worthy of further analysis. Whether this reflects an explicit choice by the plants, i.e. explicitly aiming to cut the costs of products without widespread consumer appeal, or a technological constraint, such as high quality goods simply requiring more inputs, has implications for the patterns of specialization I observe at the plant level. In particular, if the pattern reflects an explicit choice by plants, it would be useful to determine if there are ways to change firms incentives to engage in both quality upgrading *and* productivity upgrading. These sorts of policies could then augment the quality-focusing gains from trade uncovered in this paper, through subsequent efficiency gains. On the other hand, if this relationship is entirely technological in nature, within-firm allocative gains, as well as the across-firm allocative gains uncovered in other papers, may be limited by the quality-cost tradeoff uncovered in this paper.

Similarly, while the negative correlation between TFPQ and quality was uncovered using a very simple demand system, it would be useful to examine if one can uncover this relationship with more flexible demand structures, that allow different types of consumers to value quality and low-prices differently (e.g. low versus high income consumers). While the demand system considered in this paper cannot allow for these sort of effects, the observed movement towards specialization in quality, could have something to do with producers trying to capture different subsets of the market, and may have implications for income inequality that are worthy of further analysis.

References

Ackerberg, Daniel, Benkard Lanier, Steven Berry, and Ariel Pakes (2007), "Econometric tools for analyzing market outcomes." In *Handbook of Econometrics* (J. J. Heckman and E.E. Leamer, eds.), volume 6, Elsevier.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer (2015), "Identification properties of recent production function estimators." *Econometrica*, 83, 2411–2451.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D O'Connell (2016), "How do electricity shortages affect industry? Evidence from India." *American Economic Review*, 106, 587–624.

Amiti, Mary and Amit K Khandelwal (2013), "Import competition and quality upgrading." *Review of Economics and Statistics*, 95, 476–490.

Anderson, Simon P, André De Palma, and Jacques-Francois Thisse (1987), "The CES is a discrete choice model?" *Economics Letters*, 24, 139–140.

Anderson, Theodore Wilbur and Cheng Hsiao (1982), "Formulation and estimation of dynamic models using panel data." *Journal of Econometrics*, 18, 47–82.

Arellano, Manuel and Stephen Bond (1991), "Some tests of specification for panel data: Monte carlo evidence and an application to employment equations." *Review of Economic Studies*, 58, 277–297.

Atalay, Enghin (2014), "Materials prices and productivity." *Journal of the European Economic Association*, 12, 575–611.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson (2014), "Vertical integration and input flows." *The American Economic Review*, 104, 1120–1148.

Autor, David H., David Dorn, and Gordon H Hanson (2013), "The China syndrome: Local labor market effects of import competition in the United States." *American Economic Review*, 103, 2121–2168.

Balat, Jorge, Irene Brambilla, and Yuya Sasaki (2016), "Heterogeneous firms: Skilled-labor productivity and the destination of exports." *Working Paper*.

Baldwin, John and Wulong Gu (2009), "The impact of trade on plant scale, production-run length and diversification." In *Producer Dynamics: New Evidence from Micro Data*, 557–592, University of Chicago Press.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta (2013), "Cross-country differences in productivity: The role of allocation and selection." *The American Economic Review*, 103, 305–334.

Berger, Brett and Robert F Martin (2013), "The Chinese export boom: an examination of the detailed trade data." *China & World Economy*, 21, 64–90.

Bernard, Andrew B. and J Bradford Jensen (2004), "Exporting and productivity in the USA." *Oxford Review of Economic Policy*, 20, 343–357.

Bernard, Andrew B, Stephen J Redding, and Peter K Schott (2010), "Multiple-product firms and product switching." *The American Economic Review*, 100, 70–97.

Bernard, Andrew B, Stephen J Redding, and Peter K Schott (2011), "Multiproduct firms and trade liberalization." *Quarterly Journal of Economics*, 126, 1271–1318.

Berry, Steven, Amit Gandhi, and Philip Haile (2013), "Connected substitutes and invertibility of demand." *Econometrica*, 81, 2087–2111.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium." *Econometrica*, 63, 242–262.

Berry, Steven T. (1994), "Estimating discrete-choice models of product differentiation." *The RAND Journal of Economics*, 25, 242–262.

Berry, Steven T and Philip A Haile (2014), "Identification in differentiated products markets using market level data." *Econometrica*, 82, 1749–1797.

Besley, Timothy and Robin Burgess (2004), "Can labor regulation hinder economic performance? Evidence from India." *Quarterly Journal of Economics*, 119, 91–134.

Björnerstedt, Jonas and Frank Verboven (2013), "Does merger simulation work? Evidence from the Swedish analgesics market." *American Economic Journal: Applied Economics*.

Bloom, Nicholas, Mirko Draca, and John Van Reenen (2016), "Trade induced technical change? The impact of chinese imports on innovation, IT and productivity." *The Review of Economic Studies*, 83, 87–117.

Blum, Bernardo S, Sebastian Claro, Ignatius Horstmann, and David A Rivers (2018), "The abcs of firm heterogeneity: The effects of demand and cost differences on exporting." *Working Paper*.

Blundell, Richard and Stephen Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics*, 87, 115–143.

Blundell, Richard and Stephen Bond (2000), "GMM estimation with persistent panel data: An application to production functions." *Econometric Reviews*, 19, 321–340.

Bond, Stephen and Måns Söderbom (2005), "Adjustment costs and the identification of Cobb Douglas production functions." *Oxford Economics Department Working Paper*.

Brandt, Loren, Feitao Jiang, Yao Luo, and Yingjun Su (2017), "Ownership and productivity in vertically-integrated firms: Evidence from the Chinese steel industry." *Working Paper*.

Collard-Wexler, Allan and Jan De Loecker (2015), "Reallocation and technology: Evidence from the US steel industry." *American Economic Review*, 105, 131–171.

De Loecker, Jan (2007), "Do exports generate higher productivity? Evidence from Slovenia." *Journal of International Economics*, 73, 69–98.

De Loecker, Jan (2011), "Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity." *Econometrica*, 79, 1407–1451.

De Loecker, Jan (2013), "Detecting learning by exporting." *American Economic Journal: Microeconomics*, 5, 1–21.

De Loecker, Jan, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik (2016), "Prices, markups, and trade reform." *Econometrica*, 84, 445–510.

De Loecker, Jan and Pinelopi Koujianou Goldberg (2014), "Firm performance in a global market." *The Annual Review of Economics*.

De Loecker, Jan and Frederic Warzynski (2012), "Markups and firm-level export status." *American Economic Review*, 102, 2437–2471.

Dhyne, Emmanuel, Amil Petrin, Valerie Smeets, and Frederic Warzynski (2017), "Multiproduct firms, import competition, and the evolution of firm product technical efficiencies." *NBER Working Paper*.

Diewert, W. Erwin (1973), "Functional forms for profit and transformation functions." *Journal of Economic Theory*, 6, 284–316.

Doraszelski, Ulrich and Jordi Jaumandreu (2013), "R&D and productivity: Estimating endogenous productivity." *Review of Economic Studies*, 80, 1338–1383.

Eckel, Carsten, Leonardo Iacovone, Beata Javorcik, and J Peter Neary (2015), "Multi-product firms at home and away: Cost-versus quality-based competence." *Journal of International Economics*, 95, 216–232.

Eckel, Carsten and J Peter Neary (2010), "Multi-product firms and flexible manufacturing in the global economy." *Review of Economic Studies*, 77, 188–217.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler (2004), "The effects of structural reforms on productivity and profitability enhancing reallocation: Evidence from Colombia." *Journal of Development Economics*, 75, 333–371.

Feenstra, Robert and Hong Ma (2007), "Optimal choice of product scope for multiproduct firms under monopolistic competition." Technical report, National Bureau of Economic Research.

Forlani, Emanuele, Ralf Martin, Giordiano Mion, and Mirabelle Muûls (2016), "Unraveling firms: Demand, productivity and markups heterogeneity." *Working Paper*.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2008), "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?" *American Economic Review*, 98, 394–425.

Gandhi, Amit, Salvador Navarro, and David Rivers (2016), "On the identification of production functions: How heterogeneous is productivity?" *Working Paper*.

Garcia-Marin, Alvaro and Nico Voigtländer (2013), "Exporting and plant-level efficiency gains: It's in the measure." Technical report, National Bureau of Economic Research.

Garcia-Marin, Alvaro and Nico Voigtländer (2017), "Product-level efficiency and core competence in multi-product plants." *Working Paper*.

Goldberg, Pinelopi K., Amit Khandelwal, Nina Pavcnik, and Petia Topalova (2010a), "Imported intermediate inputs and domestic product growth: Evidence from India." *The Quarterly Journal of Economics*, 125, 1727–1767.

Goldberg, Pinelopi K., Amit Khandelwal, Nina Pavcnik, and Petia Topalova (2010b), "Multiproduct firms and product turnover in the developing world: Evidence from India." *Review of Economics and Statistics*, 92, 1042–1049.

Griliches, Zvi and Jacques Mairesse (1995), "Production functions: The search for identification." Technical report, National Bureau of Economic Research.

Hall, Robert E (1973), "The specification of technology with several kinds of output." *Journal of Political Economy*, 81, 878–892.

Hottman, Colin J, Stephen J Redding, and David E Weinstein (2016), "Quantifying the sources of firm heterogeneity." *Quarterly Journal of Economics*, 131, 1291–1364.

Hsieh, Chang-tai and Peter J. Klenow (2009), "Misallocation and manufacturing TFP in China and India." *Quarterly Journal of Economics*, 1403.

Jaumandreu, Jordi and Heng Yin (2016), "Cost and product advantages: A firm-level model for the chinese exports and industry growth." Technical report, Boston University mimeo.

Khandelwal, Amit (2010), "The long and short (of) quality ladders." *Review of Economic Studies*, 77, 1450–1476.

Klette, Tor Jakob and Zvi Griliches (1996), "The inconsistency of common scale estimators when output prices are unobserved and endogenous." *Journal of Applied Econometrics*, 343–361.

Kugler, Maurice and Eric Verhoogen (2012), "Prices, plant size, and product quality." *Review of Economic Studies*, 79, 307–339.

Levinsohn, James and Amil Petrin (2003), "Estimating production functions using inputs to control for unobservables." *Review of Economic Studies*, 70, 317–341.

Liu, Runjuan (2010), "Import competition and firm refocusing." *Canadian Journal of Economics*, 43, 440–466.

Martin, Leslie A, Shanthi Nataraj, and Ann E Harrison (2017), "In with the big, out with the small: Removing small-scale reservations in India." *The American Economic Review*, 107, 354–386.

Mayer, Thierry, Marc J. Melitz, and Gianmarco IP Ottaviano (2014), "Market size, competition, and the product mix of exporters." *American Economic Review*, 104, 495–536.

Mayer, Thierry, Marc J. Melitz, and Gianmarco IP Ottaviano (2016), "Product mix and firm productivity responses to trade competition." Technical report, National Bureau of Economic Research.

McFadden, Daniel (1978), "Modeling the choice of residential location." *Transportation Research Record*.

McFadden, Daniel et al. (1978), "Cost, revenue, and profit functions." *History of Economic Thought Chapters*, 1.

Medina, Pamela (2017), "Import competition, quality upgrading and exporting: Evidence from the peruvian apparel industry." Technical report, University of Toronto.

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

Melitz, Marc J and Sašo Polanec (2015), "Dynamic Olley-Pakes productivity decomposition with entry and exit." *The Rand journal of economics*, 46, 362–375.

Melitz, Marc J and Daniel Trefler (2012), "Gains from trade when firms matter." *The Journal of Economic Perspectives*, 26, 91–118.

Morlacco, Monica (2018), "Market power in input markets: Theory and evidence from french manufacturing." *Working Paper*.

Nevo, Aviv (1998), "Identification of the oligopoly solution concept in a differentiated-products industry." *Economics Letters*, 59, 391–395.

Olley, G. Stephen and Ariel Pakes (1996), "The dynamics of productivity in the telecommunications equipment industry." *Econometrica*, 64, 1263–1297.

Pavcnik, Nina (2002), "Trade liberalization, exit, and productivity improvements: Evidence from chilean plants." *Review of Economic Studies*, 69, 245–276.

Restuccia, Diego and Richard Rogerson (2008), "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic dynamics*, 11, 707–720.

Rosse, James N. (1970), "Estimating cost function parameters without using cost data: Illustrated methodology." *Econometrica*, 256–275.

Smeets, Valerie and Frederic Warzynski (2013), "Estimating productivity with multi-product firms, pricing heterogeneity and the role of international trade." *Journal of International Economics*, 90, 237–244.

Topalova, Petia and Amit Khandelwal (2011), "Trade liberalization and firm productivity: The case of India." *Review of Economics and Statistics*, 93, 995–1009.

Train, Kenneth E (2009), *Discrete choice methods with simulation*. Cambridge university press.

Trefler, Daniel (1993), "Trade liberalization and the theory of endogenous protection: An econometric study of us import policy." *Journal of Political Economy*, 101, 138–160.

Trefler, Daniel (2004), "The long and short of the Canada-U. S. free trade agreement." *American Economic Review*, 94, 870–895.

Valmari, Nelli (2016), "Estimating production functions of multiproduct firms." Technical report, Research Institute of the Finnish Economy.

Verboven, Frank (1996), "The nested logit model and representative consumer theory." *Economics Letters*, 50, 57–63.

Verhoogen, Eric A (2008), "Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector." *Quarterly Journal of Economics*, 123, 489–530.

8 Appendix

8.1 Appendix A: Proof of Lemma 1

Consider the following conditional cost-minimization problem (CM) for firms possessing an arbitrary production technology $Y_{it}^j = \exp(\omega_{it}^j)F\left(\vec{X}_{it}^j\right)$:

$$\begin{aligned} \underset{\mathbb{X}_{it}}{\text{Min}} \quad & \sum_{M\in\mathbb{M}}\sum_{j\in\mathbb{Y}_{it}} \mathbf{W}^M\left(M_{it}, A_{it}^M\right) M_{it}^j \\ \text{subject to} \quad & \exp(\omega_{it}^j)F\left(\vec{X}_{it}^j\right) \geq Y_{it}^j,\ \forall j\in\mathbb{Y}_{it} \\ & \sum_{j\in\mathbb{Y}_{it}} X_{it}^j = X_{it},\ \forall X\in\mathbb{K}, \end{aligned} \tag{CM}$$

where $M_{it} = \sum_{j\in\mathbb{Y}_{it}} M_{it}^j$.

Any solution to (CM), will consist of an *aggregate input vector* for static inputs $\vec{M}_{it}$, as well as an *input-share allocation* for each input $j\in\mathbb{Y}_{it}$, $\vec{S}_{it}^j$, where S_{it}^j is a $1\times P$ vector of input shares going into production line j. Let $\vec{S}_{it}$ denote the matrix of all input shares, with typical element (j,X) of $\vec{S}_{it}$ being given by S_{it}^{jX}. Note that using this notation, $\vec{X}_{it}^j = \vec{S}_{it}^j\circ\vec{X}_{it}$, where $\circ$ denotes element-by-element multiplication of two vectors.

To prove Lemma 1, it is easier to break down the above problem into two parts. First consider the following transformation function problem (TF):

$$\begin{aligned} \underset{\vec{S}_{it}^{-r}}{\text{maximize}} \quad & Y_{it}^r = \exp(\omega_{it}^r)F\left(\left(\mathbf{1} - \sum_{j\in\mathbb{Y}_{it}^{-r}}\vec{S}_{it}^j\right)\circ\vec{X}_{it}\right) \\ \text{subject to} \quad & \exp(\omega_{it}^j)F\left(\vec{S}_{it}^j\circ\vec{X}_{it}\right) \geq Y_{it}^j,\ \forall j\in\mathbb{Y}_{it}^{-r} \\ & 0\leq S_{it}^{JX}\leq 1\ \forall j\in\mathbb{Y}_{it}^{-r}\ \&\ \forall X\in(\mathbb{K},\mathbb{M}), \end{aligned} \tag{TF}$$

where $\mathbf{1}$ is a $1\times P$ vector of ones, $r\in\mathbb{Y}_{it}$ is an arbitrary good within firm i's output set, $\mathbb{Y}_{it}^{-r}\equiv\mathbb{Y}_{it}\setminus\{r\}$, and $\vec{S}_{it}^{-r}$ is the matrix of input shares obtained after excluding the row corresponding to output r.

The objective function of (TF) evaluated at an optimal solution yields the *transformation function for output* r, $Y_{it}^r = \widetilde{T}(\vec{Y}_{it}^{-r}, \vec{X}_{it}, \vec{\omega}_{it})$, which corresponds to the maximal level of output firm i can obtain, given a vector of aggregate inputs $\vec{X}_{it}$ and some predetermined level of other outputs that firm i wishes to produce, $\vec{Y}_{it}^{-r}$.[110]

Rather than directly solving (CM) to obtain the conditional cost function, we may instead obtain the cost function by first solving (TF), and the following problem:

$$\begin{aligned} \underset{\vec{M}_{it}}{\text{Min}} \quad & \sum_{M\in\mathbb{M}} \mathbf{W}^M\left(M_{it}, A_{it}^M\right) M_{it} \\ \text{subject to} \quad & Y_{it}^r = \widetilde{T}(\vec{Y}_{it}^{-r}, \vec{X}_{it}, \vec{\omega}_{it}). \end{aligned} \tag{CM2}$$

[110] Note that the transformation function essentially describes the set of efficient aggregate-input and output vectors that firm i may produce with the TFP vector $\vec{\omega}_{it}$: see Diewert (1973).

With this in mind, the input shares implied by (TF) will solve (CM) as long as the level of static inputs in (TF) is chosen to solve (CM2). I use this property to prove Lemma 1:

Lemma 1. *If Assumptions 1 through 6 hold, then there exists a solution to the firm's conditional cost minimization problem satisfying* $X^j_{it} = S^j_{it} X_{it} \quad \forall X \in (\mathbb{K}, \mathbb{M})$*, where* $S^j_{it} \in [0,1]$ *and* $\sum_{j \in \mathbb{Y}_{it}} S^j_{it} = 1$.

Proof. Consider the Lagrangian associated with (TF)

$$L = \exp(\omega^r_{it}) F\left(\left(\mathbf{1} - \sum_{j \in \mathbb{Y}^r_{it}} \vec{S}^j_{it}\right) \circ \vec{X}_{it}\right) + \sum_{j \in \mathbb{Y}^r_{it}} \mu^j_{it}\left(\exp(\omega^j_{it}) F\left(\vec{S}^j_{it} \circ \vec{X}_{it}\right) - Y^j_{it}\right). \tag{50}$$

Suppose that the vector of static inputs, $\vec{M}_{it}$, is chosen to solve (CM2), which guarantees that at an optimum $\mathbf{1} - \sum_{j \in \mathbb{Y}^r_{it}} \vec{S}^j_{it} > \mathbf{0}$. The set of first-order necessary conditions for a solution to (TF) where all production function constraints bind ($\mu^j_{it} > 0 \; \forall j \in \mathbb{Y}^r_{it}$) are then given by:

$$FOC^{jX}_{it}\left(\vec{S_{it}}, \vec{\mu_{it}}\right) \equiv \exp(\omega^r_{it}) \frac{\partial F(\vec{X}^r_{it})}{\partial X} X_{it} - \mu^j_{it} \exp(\omega^j_{it}) \frac{\partial F(\vec{X}^j_{it})}{\partial X} X_{it} = 0, \tag{51}$$

$$P^j\left(\vec{S_{it}}, \vec{\mu_{it}}\right) \equiv Y^j_{it} - \exp(\omega^j_{it}) F\left(\vec{X}^j_{it}\right) = 0, \tag{52}$$

where I am making the substitution $\vec{X}^j_{it} = \vec{S}^j_{it} \circ \vec{X}_{it}$. Note that the set of first-order conditions are satisfied if $FOC^{jX}_{it}\left(\vec{S_{it}}, \vec{\mu_{it}}\right) = 0 \; \forall j \in \mathbb{Y}_{it}$ and $\forall X \in (\mathbb{K}, \mathbb{M})$, and $P^j\left(\vec{S_{it}}, \vec{\mu_{it}}\right) = 0 \; \forall j \in \mathbb{Y}_{it}$.

Consider the following allocation $\vec{S}^*_{it}$ and set of Lagrangian multipliers, $\vec{\mu}^*_{it}$:

$$S^{jX*}_{it} = S^{j*}_{it} = \frac{\left(\mu^{j*}_{it} \exp(\omega^j_{it})\right)^{\frac{1}{1-\phi}}}{\left(\exp(\omega^r_{it})\right)^{\frac{1}{1-\phi}} + \sum_{k \in \mathbb{Y}^r_{it}} \left(\mu^{k*}_{it} \exp(\omega^k_{it})\right)^{\frac{1}{1-\phi}}} \quad \forall j \in \mathbb{Y}^r_{it}, \forall X \in (\mathbb{K}, \mathbb{M}), \tag{53}$$

$$\mu^{j*}_{it} = \exp(\omega^r_{it} - \omega^j_{it}) \left(\frac{\left(\frac{Y^j_{it}}{\exp(\omega^j_{it})}\right)^{\frac{1}{\phi}}}{\left(\left(F(\vec{X}_{it})\right)^{\frac{1}{\phi}} - \sum_{k \in \mathbb{Y}^r_{it}} \left(\frac{Y^k_{it}}{\exp(\omega^k_{it})}\right)^{\frac{1}{\phi}}\right)}\right)^{1-\phi} \quad \forall j \in \mathbb{Y}^r_{it}. \tag{54}$$

I now show that $FOC^{jX}_{it}\left(\vec{S}^*_{it}, \vec{\mu}^*_{it}\right) = P^j\left(\vec{S}^*_{it}, \vec{\mu}^*_{it}\right) = 0 \; \forall j \in \mathbb{Y}^r_{it}$ and $\forall X \in (\mathbb{K}, \mathbb{M})$.

Since F is homogenous of degree ϕ, $\frac{\partial F}{\partial X}$ is homogenous of degree $\phi - 1$. Moreover, note that (53) implies $S^{jX*}_{it} = S^{j*}_{it} \quad \forall X \in (\mathbb{K}, \mathbb{M})$, which implies $\vec{X}^j_{it} = S^{j*}_{it} \vec{X}_{it} \; \forall j \in \mathbb{Y}_{it}$.[111] Substituting these expressions into (51) yields:

$$FOC^{jX}_{it}\left(\vec{S}^*_{it}, \vec{\mu}^*_{it}\right) = \exp(\omega^r_{it}) \left(S^{r*}_{it}\right)^{\phi-1} \frac{\partial F(\vec{X}_{it})}{\partial X} X_{it} - \mu^{j*}_{it} \exp(\omega^j_{it}) \left(S^{j*}_{it}\right)^{\phi-1} \frac{\partial F(\vec{X}_{it})}{\partial X} X_{it}, \tag{55}$$

where S^{j*}_{it} has been taken out of $\frac{\partial F}{\partial X}$ since $\frac{\partial F}{\partial X}$ is homogeneous of degree $\phi - 1$. We can then write:

[111] In particular, note that (53) implies $S^{rX*}_{it} = S^{r*}_{it} = \frac{\left(\exp(\omega^r_{it})\right)^{\frac{1}{1-\phi}}}{\left(\exp(\omega^r_{it})\right)^{\frac{1}{1-\phi}} + \sum_{k \in \mathbb{Y}^r_{it}} \left(\mu^{k*}_{it} \exp(\omega^k_{it})\right)^{\frac{1}{1-\phi}}} \quad \forall X \in (\mathbb{K}, \mathbb{M})$.

$$FOC_{it}^{jX}\left(\vec{S_{it}^*},\vec{\mu_{it}^*}\right)=\frac{\partial F(\vec{X}_{it})}{\partial X}X_{it}\left(\exp(\omega_{it}^r)\left(S_{it}^{r*}\right)^{\phi-1}-\mu_{it}^{j*}\exp(\omega_{it}^j)\left(S_{it}^{j*}\right)^{\phi-1}\right). \tag{56}$$

Note that (53) implies that $s_{it}^{rX*}=s_{it}^{r*}=\frac{\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}}{\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}+\sum_{k\in\mathbb{Y}_{it}^r}\left(\mu_{it}^{k*}\exp(\omega_{it}^k)\right)^{\frac{1}{1-\phi}}}\quad\forall X\in(\mathbb{K},\mathbb{M})$. Substituting this expression and (53) into (56) yields, after cancelling and collecting like terms:

$$FOC_{it}^{jX}\left(\vec{S}_{it}^*,\vec{\mu}_{it}^*\right)=\frac{\partial F(\vec{X}_{it})}{\partial X}X_{it}\left(\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}+\sum_{k\in\mathbb{Y}_{it}^r}\left(\mu_{it}^{k*}\exp(\omega_{it}^k)\right)^{\frac{1}{1-\phi}}\right)^{1-\phi}(1-1)=0.$$

Hence, $FOC_{it}^{jX}\left(\vec{S}_{it}^*,\vec{\mu}_{it}^*\right)=0\ \forall j\in\mathbb{Y}_{it}^r$ and $\forall X\in(\mathbb{K},\mathbb{M})$.

Next I show the same holds for all $P^j\left(\vec{S}_{it}^*,\vec{\mu}_{it}^*\right)=0$. Substitute $\vec{X}_{it}^j=S_{it}^{j*}\vec{X}_{it}$ into (52), and use the fact that F is homogeneous of degree ϕ, yielding:

$$P^j\left(\vec{S}_{it},\vec{\mu}_{it}\right)=Y_{it}^j-\exp(\omega_{it}^j)\left(S_{it}^{j*}\right)^{\phi}F\left(\vec{X}_{it}\right). \tag{57}$$

To obtain an expression for S_{it}^{j*} in terms of exogenous parameters, multiply (54) by $\exp(\omega_{it}^j)$, and then take this equation to the power of $\frac{1}{1-\phi}$, yielding:

$$\left(\mu_{it}^{j*}\exp(\omega_{it}^j)\right)^{\frac{1}{1-\phi}}=\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}\frac{\left(\frac{Y_{it}^j}{\exp(\omega_{it}^j)}\right)^{\frac{1}{\phi}}}{\left(\left(F(\vec{X}_{it})\right)^{\frac{1}{\phi}}-\sum_{k\in\mathbb{Y}_{it}^r}\left(\frac{Y_{it}^k}{\exp(\omega_{it}^k)}\right)^{\frac{1}{\phi}}\right)}. \tag{58}$$

Sum (58) over all $k\in\mathbb{Y}_{it}^r$, yielding:

$$\sum_{k\in\mathbb{Y}_{it}^r}\left(\mu_{it}^{k*}\exp(\omega_{it}^k)\right)^{\frac{1}{1-\phi}}=\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}\frac{\sum_{k\in\mathbb{Y}_{it}^r}\left(\frac{Y_{it}^k}{\exp(\omega_{it}^j)}\right)^{\frac{1}{\phi}}}{\left(\left(F(\vec{X}_{it})\right)^{\frac{1}{\phi}}-\sum_{k\in\mathbb{Y}_{it}^r}\left(\frac{Y_{it}^k}{\exp(\omega_{it}^k)}\right)^{\frac{1}{\phi}}\right)}. \tag{59}$$

Adding $\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}$ to both sides of (59) yields:

$$\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}+\sum_{k\in\mathbb{Y}_{it}^r}\left(\mu_{it}^{k*}\exp(\omega_{it}^k)\right)^{\frac{1}{1-\phi}}=\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}\frac{\left(F(\vec{X}_{it})\right)^{\frac{1}{\phi}}}{\left(\left(F(\vec{X}_{it})\right)^{\frac{1}{\phi}}-\sum_{k\in\mathbb{Y}_{it}^r}\left(\frac{Y_{it}^k}{\exp(\omega_{it}^k)}\right)^{\frac{1}{\phi}}\right)}. \tag{60}$$

Dividing (58) by (60) yields:

$$\frac{\left(\mu_{it}^{j*}\exp(\omega_{it}^j)\right)^{\frac{1}{1-\phi}}}{\left(\exp(\omega_{it}^r)\right)^{\frac{1}{1-\phi}}+\sum_{k\in\mathbb{Y}_{it}^r}\left(\mu_{it}^{k*}\exp(\omega_{it}^k)\right)^{\frac{1}{1-\phi}}}=\left(\frac{Y_{it}^j}{\exp(\omega_{it}^j)F(\vec{X}_{it})}\right)^{\frac{1}{\phi}}. \tag{61}$$

Since $\frac{\left(\mu_{it}^{j*}\exp(\omega_{it}^{j})\right)^{\frac{1}{1-\phi}}}{\left(\exp(\omega_{it}^{r})\right)^{\frac{1}{1-\phi}}+\sum_{k\in\mathbb{Y}_{it}^{r}}\left(\mu_{it}^{k*}\exp(\omega_{it}^{k})\right)^{\frac{1}{1-\phi}}} = S_{it}^{j*}$, we can substitute (61) into (57), yielding:

$$P^{j}\left(\vec{S}_{it},\vec{\mu}_{it}\right) = Y_{it}^{j} - \exp(\omega_{it}^{j})\left(\left(\frac{Y_{it}^{j}}{\exp(\omega_{it}^{j})F(\vec{X}_{it})}\right)^{\frac{1}{\phi}}\right)^{\phi} F\left(\vec{X}_{it}\right) = Y_{it}^{j} - Y_{it}^{j} = 0.$$

Hence, $FOC_{it}^{jX}\left(\vec{S}_{it}^{*},\vec{\mu}_{it}^{*}\right) = P^{k}\left(\vec{S}_{it}^{*},\vec{\mu}_{it}^{*}\right) = 0\ \forall j \in \mathbb{Y}_{it}^{r}$ and $\forall X \in (\mathbb{K},\mathbb{M})$, which means $\vec{S}_{it}^{*}$ and $\vec{\mu}_{it}^{*}$ satisfy the first-order necessary conditions for (TF). Moreover, since F is quasi-concave, the first-order necessary conditions are sufficient. Therefore $\vec{S}_{it}^{*}$ and $\vec{\mu}_{it}^{*}$ are a solution to (TF). The proposition follows by noting $\vec{S}_{it}^{*}$ satisfies $S_{it}^{jX} = S_{it}^{j}\quad \forall X \in (\mathbb{K},\mathbb{M})$ by construction, and the shares implied by (TF) will also solve (CM) when evaluated at the level of static inputs implied by (CM2).

□

8.2 Appendix B: Data Appendix

8.2.1 Appendix B1: Sample Selection and Industry Details

The core sample used in this paper is all firms producing products belonging to the two-digit ASICC codes 74-78, which I refer to as the Machinery, Equipment and Parts industry.[112] Details on these product codes are described in Table 21, below. This choice of codes was in part driven by observed similarities (see Table 21), as well as data driven. In particular, while I regard ASICC 77: ELECTRICAL & ELECTRONIC MACHINERY & EQUIPMENT INCL PARTS as the "core" industry, note that 45 % of the multi-product observations for ASICC 77 are produced by plants that produce products belonging to multiple 2-digit AISICC codes. Since focussing only of ASICC 77 firms would require that I drop close to half of the multi-product observations, I also add ASICC codes 74, 75, 76, and 78 to the sample, because they are relatively similar categories, and together account for more than 80 % of the multi-product variety-year observations for firms producing products belonging to ASICC 77.

As discussed in the main text, this industry was also chosen in large part because vertical integration is a smaller problem than in other industries. In particular, as described in Table 22, below, the Machinery, Equipment, and Parts industry has a relatively low level of vertical integration.

To determine firms which may be vertically integrated, I construct an input-output table, using information on inputs purchased by single product firms. For each output 5-digit ASICC code j, I collect the set of 5-digit ASICC codes k that are listed as inputs for single-product firms that produce j. I consider a multi-product firm *potentially vertically integrated* if they produce a pair of outputs, j and k, where at least 10 % of single-product producers of j also buy input k. While the 10% threshold is somewhat arbitrary, note more restrictive rules may result in many non-vertically integrated firms being classified as vertically integrated. For example, if one instead considers a firm potentially vertically integrated if *any* producer of j is observed using k, then more than 75 % of multi-product firms in the ASI, i.e. including all industries, are considered vertically integrated. Since it seems, at face value,

[112] Two digit ASICC codes from the 2009 version of the ASICC codes, which I match to the observed 2000-2008 ASICC codes in my data.

implausible that this many firms are actually producing multiple outputs for vertical integration reasons, I restrict attention to input-output pairs that are somewhat common, to remove the influence of outlier (rare) production sets. The 10% threshold results in just over 40% of all ASI multiproduct firms being classified as potentially vertically integrated. While this approach still classifies a fairly large percentage of multi-product firms as potentially vertically integrated, note that more lenient thresholds (e.g 20%) resulted in some obvious examples of vertical integration (e.g. batteries and battery plates) not being classified as vertically integrated.

As discussed in the main text, unless otherwise stated, I drop potentially vertically integrated plants in any regressions requiring information on the allocation of inputs across production lines, since these are unlikely to be measured correctly. I do, however, include these firms in the demand estimation regressions, since input shares do not need to be known for their prices do be information. Similarly, while I include multi-industry firms in the demand regressions, since they can aid in identification, I only consider single industry firms in all subsequent regressions, i.e. firms who *only* produce products from ASICC codes 74-78, since these regressions generally require estimates of within-firm input allocations, which I cannot know without knowledge of the production function used in other industries. These regressions also do not include plants with *ad-valorem* tax rates below 0 or above 1, as they are likely plagued by data entry errors. Similarly, all regressions drop plants with negative or missing prices and quantities, who do not record any values for any of the inputs used in production function estimation, who did not operate for 12 months, and whose entries appeared to have some major data entry errors, including a small number of firms who listed their initial production year as being past the end of the sample period, reported input quantity codes that did not equal the standard quantity code for the given ASICC code, as well some firms who had entries for a particular firm-type that was not included in any of the ASI code books.

Table 21: Machinery, Equipment, and Parts

ASICC	Description	Observations	Multi-product Revenue Share (%)	Multi-product Observation Share (%)	Example 5-digit Products
74	MISC. MANUFACTURE OF BASE METALS, N.E.C	10,035	66	73	-Cylinders -Wheels -Rims
75	NON-ELECTRICAL MACHINE TOOLS & GENERAL PURPOSE MACHINERIES AND COMPONENTS AND PARTS THEREOF	14,903	74	77	-Gears -Ball Bearings -Valves
76	NON-ELECTRICAL INDUSTRY SPECIFIC EQUIPMENT / MACHINERIES INCL PARTS THEREOF	12,317	80	77	-Agriculture Implements -Textile Machinery -Drilling Machines
77	ELECTRICAL & ELECTRONIC MACHINERY & EQUIPMENT INCL PARTS (EXCL MEDICAL & NON-CONVENTIONAL ENERGY EQUIPMENT)	22,032	71	76	-Transformers -Control Equipment -Batteries
78	ELECTRONICS EQUIPMENT & PARTS EXCL BIOMEDICAL EQUIPMENT	4,822	67	77	-Printed Circuit Plate -T.V. Set -Personal Computer

Notes: Observations refers to number of variety-year observations.

Table 22: Potentially Vertically Integrated Firms: Top 10 2-Digit ASICC Codes

ASICC	Description	Observations	Potentially Vertically Integrated Revenue Share (%)	Potentially Vertically Integrated Observation Share (%)
15	BEVERAGES, TOBACCO AND PAN MASALA	8,068	20	22
64	SYNTHETIC (MAN-MADE) AND MIXED TEXTILES	9,592	52	29
74	MISC. MANUFACTURE OF BASE METALS, N.E.C [NO SUB-GROUP FORMATION]	10,035	28	18
76	NON-ELECTRICAL INDUSTRY SPECIFIC EQUIPMENT/MACHINERIES INCL PARTS THEREOF	12,317	12	14
42	PLASTIC, PVC ARTICLES INCL PACKAGING PRODUCTS AND FOOTWEAR PLASTIC OR PVC	12,545	43	20
75	NON-ELECTRICAL MACHINE TOOLS & GENERAL PURPOSE MACHINERIES AND COMPONENTS AND PARTS THEREOF	14,903	26	19
77	ELECTRICAL & ELECTRONIC MACHINERY & EQUIPMENT INCL PARTS (EXCL MEDICAL & NON-CONVENTIONAL ENERGY EQUIPMENT)	22,032	31	31
63	COTTON, COTTON YARN AND FABRICS	26,565	36	36
71	IRON & STEEL (INCL STAINLESS STEEL) & ARTICLES THEREOF	32,619	55	29
12	FRUITS, VEGETABLES, CEREALS & PULSES AND OTHER VEGETABLE PRODUCES LIKE LAC, GUM ETC. AND PREPARATION THEREOF	34,616	57	62

Notes: Observations refers to number of variety-year observations. Potentially vertically integrated firms are firms that produce at least one 5-digit ASICC code that is classified as a potential input for another product they produce. Potential inputs are identified for each product code by examining all product codes used by single product firms. Product code j is classified as a potential input for product code k is at least 10% of observed single product k producers list item code j as an input.

8.2.2 Appendix B2: 3-digit ASICC codes

3-digit ASICC codes

ASICC code	Description
740	MISC. MANUFACTURE OF BASE METALS, N.E.C
751	NON-ELECTRICAL MACHINE TOOLS & GENERAL PURPOSE MACHINERIES AND COMPONENTS AND PARTS THEREOF
761	AGRICULTURAL & FORESTRY MACHINERIES/PARTS THEREOF
762	FOOD, BEVERAGES & TOBACCO PROCESSING MACHINERIES & PARTS
763	MININGS, QUARRYING & METALLURGICAL MACHINERIES/PARTS
764	CONSTRUCTION/CEMENT MACHINERIES & PARTS
765	TEXTILE, LEATHER & RUBBER PROCESSING, PAPER PRINTING MACHINERIES & PARTS THEREOF
766	NON-ELECTRICAL DOMESTIC/OFFICE APPLIANCES & PARTS
767	CHEMICAL/PLASTIC/GLASS/WEAPON/AMMUNITION MACHINERIES AND PARTS THEREOF
768	LIFT AND LIFTING EQUIPMENT, FIXED OR MOBILE & PARTS THEREOF
769	MISC NON-ELECTRICAL MACHINERIES AND PARTS THEREOF, N.E.C
771	ELECTRICAL MACHINERY/EQUIPMENT
772	ELECTRICAL MOTORS, GENERATORS, TRANSFORMER, POWER PACK [THIS INCL PUMP SET FITTED WITH ELECTRIC MOTOR]
773	SWITCH, SWITCH-GEAR, CONTROL PANEL, CIRCUIT BREAKERS ETC AND PARTS THEREOF
774	LAMP, FILAMENT, ELECTRODES/ANODES/CONNECTORS, FITTINGS & PARTS
775	MEASURING/CONTROLLING/REGULATING INSTRUMENTS
776	BATTERY, ACCUMULATORS, CELLS AND PARTS THEREOF
777	DOMESTIC AND OFFICE ELECTRICAL EQUIPMENT
778	ELECTRO MAGNET, FANS, ARMATURE, COILS & ELECTRO-MAGNETIC EQUIPMENT
779	ELECTRICAL EQUIPMENT, PARTS AND ACCESSORIES, N.E.C
781	TELEPHONE/TELECOMMUNICATION/TRANSMISSION EQUIPMENT
782	AUDIO/VIDEO/SOUND APPARATUS & PARTS
783	COMPUTER & COMPUTING EQUIPMENT & PERIPHERALS & PARTS
784	ELECTRONIC VALVES/TUBES & COMPONENTS
785	ELECTRONIC CARDS & ITS COMPONENTS
789	OTHER ELECTRONIC COMPONENTS & PARTS

Notes: All 3-digit ASICC codes at the from 2009 ASICC. I map three digit ASICC codes to 5-digit ASICC codes from 2001-2008 using a custom made concordance that matches product descriptions at the 5-digit level.

8.2.3 Appendix B3: Trade Data Appendix

As described in the main text, I use data on imports and tariffs from UN Comtrade and TRAINS at the 4-digit HS-96 level to construct estimates of imports and tariffs for each 5-digit ASICC code. To generate the concordance between 5-digit ASICC codes, I first create a concordance between the observed

ASICC codes from 2001 to 2008, to the 2009 5-digit ASICC codes.[113] I then match 5-digit ASICC codes to the 7-digit 2011 National Product Classification for Manufacturing Sector (NPCMS) codes using a concordance created by the Indian Ministry of Statistics.[114] Since the first five digits of NPCMS-2011 are equivalent to the international Central Product Classification 2.0 (CPC2), I then match 5-digit NPCMS-2011/CPC2 codes to 4-digit HS-2007 codes using a United Nations Statistics Division created concordance.[115] Finally, I match 4-digit HS-2007 to 4-digit HS-96 using the concordance which can be found at https://unstats.un.org/unsd/trade/classifications/correspondence-tables.asp.

This process generates a crosswalk between 5-digit ASICC codes and 4-digit HS-96 codes.[116] Note, however, that many HS codes map to multiple ASICC codes. This does not generate any real difficulties for calculating tariff levels across products, as I can simply use the average tariff rate across all matching HS codes to obtain a estimate of the average rate of protection. However, for imports this can generate problems due to double counting imports across items. To deal with this, I allocate imports from each HS-96 code that links to multiple ASI codes proportionally to the overall value of observed materials imports in the ASI production data. Specifically, since each plant records the value of imported inputs for up to five 5-digit ASICC codes each year, I take the total observed value of imports for each ASICC code g, I^g_{ASICC}, and allocate observed imports from HS code h, I^h_{HSt}, to ASICC code g according the formula $\text{Imports}^g_t = \frac{I^g_{ASICC}}{\sum_{l \in \mathbb{H}^h} I^l_{ASICC}} I^h_{HSt}$, where Imports^g_t is value of imports for product code g during time t that I use in my empirical model, and $\mathbb{H}^h$ is the set of 5-digit ASICC codes that map to HS code h.

8.3 Appendix C: Demand Appendix

8.3.1 Appendix C1: Derivation of Demand System from Underlying Preferences

The formulation of consumer preferences follows the "constant expenditure share" formulation of the discrete choice demand system described in of Björnerstedt and Verboven (2013).[117] During each market-period t, a mass S_t of consumers decide whether they wish to buy a *single* item $j \in \Lambda^g_t \subset \Omega_t$. Conditional on choosing to consume j, utility of consumer c is given by:

$$U_{ct}(j, Q^j_{cit}, \epsilon^{jg}_{cit}) = \left(Q^{jg}_{cit} \exp\left(\frac{\eta^{jg}_{it} + \epsilon^{jg}_{cit}}{\alpha} \right) \right)^{\beta_I} (Q_{c0})^{1-\beta_I}, \tag{62}$$

where Q^{jg}_{cit} is the quantity of good $j \in \Lambda^g_t$, which is produced by firm i, currently bought by consumer c, Q_{c0} is the quantity of some composite outside good purchased by consumer c, while η^{jg}_{it} is the mean

[113]I create this concordance by matching on the given item-code descriptions between ASICC 2001-2008 and 20009, which results in a match for the vast majority of the codes.

[114]This can be found at http://www.csoisw.gov.in/cms/En/1027-npcms-national-product-classification-for-manufacturing-sector.aspx

[115]This can be found at which can be found at https://unstats.un.org/unsd/cr/registry/cpc-2.asp.

[116]A small number of ASICC codes (13) could not be matched to trade data. These account for less than one percent of the overall sample, and are thus dropped, as I cannot guarantee to be measuring market size correctly for these observations.

[117]While Björnerstedt and Verboven (2013) work with the indirect utility function directly, one can start from the utility function as a do here, similar to Anderson et al. (1987).

quality of good $j \in \Lambda_t^g$, ϵ_{cit}^{jg} is an idiosyncratic taste shock for product j, and $\alpha > 0$ and $\beta_I \in (0,1)$ are preference parameters.

I assume that conditional on j, consumers choose Q_{ct}^j and Q_{c0} to maximize their utility, given the unit price of good j, P_t^j, and the price of the composite composite commodity, which I normalize to 1, subject to some exogenous income level Y_{ct}.[118] Since (62) is a standard Cobb-Douglas utility function, this leads to equilibrium conditional consumer demands $Q_{ct}^{j*g} = \beta_I \frac{Y_{ct}}{P_{it}^{jg}}$ and $Q_{c0}^{*} = (1-\beta_I)Y_{ct}$. Substituting these expressions into (62) yields the conditional indirect utility function:

$$V_{ct}(j, \epsilon_{cit}^{jg}, Y_{ct}) = C_{ct} Y_{ct} \left(\exp\left(\frac{\beta_I}{\alpha} \eta_{it}^{jg} - \beta_I p_{it}^{jg} + \frac{\beta_I}{\alpha} \epsilon_{cit}^{jg} \right) \right), \tag{63}$$

where $C_{ct} \equiv (\beta_I)^{\beta_I}(1-\beta_I)^{1-\beta_I}$ is an arbitrary constant, and $p_t^j \equiv \ln(P_t^j)$.

Given the conditional indirect utility function (63), consumers will choose the $j \in \Omega_t$ that provides the largest indirect utility. Note, however, that it will prove useful to rescale the indirect utility function by first taking it's natural log, and then multiplying by $\frac{\alpha}{\beta_I}$, yielding:

$$\widetilde{V}_{ct}(j, \epsilon_{cit}^{jg}, Y_{ct}) = \widetilde{c}_{ct} + \frac{\alpha}{\beta_I} y_{ct} - \alpha p_t^j + \eta_t^j + \epsilon_{ct}^j, \tag{64}$$

where $y_{ct} \equiv \ln(Y_{ct})$, and $\widetilde{c}_{ct}$ is an arbitrary constant.

Since positive monotonic transformations of (63) will not affect the ordinal ranking of the products in Ω_t, the consumers problem can instead be formulated as choosing the $j \in \Omega_t$ that maximizes the transformed indirect utility function (64). As is standard in discrete choice models of demand, following McFadden (1978), instead of solving this problem on a consumer by consumer basis, one models the distribution of consumer taste shocks, ϵ_{cit}^{jg}, so that one may determine the *probability* that each product $j \in \Omega_t$ is chosen by a randomly chosen consumer within the general population. In particular, I assume the vector of idiosyncratic taste terms for each consumer c, $\vec{\epsilon}_{ct} = \left(\epsilon_{ct}^{1g}, \epsilon_{ct}^{2g}, .., \epsilon_{ct}^{1l}, \epsilon_{ct}^{2l}, ...\right)$, can be modeled as an i.i.d. draw from a generalized extreme-value distribution with cdf:

$$F(\vec{\epsilon}_{ct}) = \exp\left(- \sum_{g=0}^{G} \left(\sum_{j \in \Lambda_t^g} \exp\left(- \frac{\epsilon_{cit}^{jg}}{\sigma} \right) \right)^{\sigma} \right), \tag{65}$$

where G is the number of product groups (i.e. 5-digit product codes) on the market.

Note that σ governs the within-product code correlation of the idiosyncratic taste terms. In the special case where $\sigma = 1$, (65), there is no correlation of the idiosyncratic taste shocks within a product group, and (65) simplifies to a standard Type-1 extreme value distribution. However, as σ decreases, the correlation between the taste shocks within a product group increases-see Train (2009).

[118] Since in my empirical application, "consumers" of Machinery, Equipment, and Parts, are perhaps more appropriately thought of as downstream producers, I can adapt this framework to such a setting by regarding (62) as a Cobb-Douglas materials aggregator function used by a particular downstream producer c. In this case, producers choose quantities to maximize effective materials per rupee spent, which is an implication of cost-minimization. While the portion of expenditure allocated to material expenditures, Y_{ct}, is endogenous in the downstream firms overall profit maximization problem, note that as long as upstream producers take the downstream output-pricing game as *given*, total expenditure on materials, Y_{ct} is pinned down by downstream output prices and markups, as I discuss in footnote 120, below.

It is well known that if each $\vec{\epsilon}_{ct}$ is an i.i.d. draw from (65), then, given the indirect utility function described by (64), the probability a randomly chosen consumer chooses to consume item $j \in \Lambda_t^g$ has the following closed-form expression:

$$Prob_{it}^{jg} = \frac{\exp\left(\frac{\delta_{it}^{jg}}{\sigma}\right)\left(\sum_{k\in\Lambda_t^g}\exp\left(\frac{\delta_{it}^{kg}}{\sigma}\right)\right)^{\sigma-1}}{\sum_{l=0}^{G}\left(\sum_{k\in\Lambda_t^l}\exp\left(\frac{\delta_{it}^{kl}}{\sigma}\right)\right)^{\sigma}}, \tag{66}$$

where $\delta_{it}^{jg} \equiv \eta_{it}^{jg} - \alpha p_{it}^{jg}$ is the mean utility for choosing to consume variety $j \in \Lambda_t^g$.

Since there are a continuum of consumers within the market, each of whom, with probability $Prob_{it}^{jg}$, purchases $Q_{cit}^{jg*} = \beta_I \frac{Y_{ct}}{P_{it}^{jg}}$ units of item j, the total aggregate demand for product $j \in \Lambda_t^g$, *produced by firm* i, is given by:

$$Q_{it}^{jg} = \frac{E_t}{P_{it}^{jg}} \frac{\exp\left(\frac{\delta_{it}^{jg}}{\sigma}\right)\left(\sum_{k\in\Lambda_t^g}\exp\left(\frac{\delta_{mt}^{kg}}{\sigma}\right)\right)^{\sigma-1}}{\sum_{l=0}^{G}\left(\sum_{k\in\Lambda_t^l}\exp\left(\frac{\delta_{mt}^{kl}}{\sigma}\right)\right)^{\sigma}}, \tag{67}$$

where $E_t = \mathbb{E}\left(Y_{ct}\right)$ is expected expenditure by all consumers on the items within this market.

Note that in the special case where $\sigma = 1$, (67) simplifies to $Q_{it}^{jg} = \frac{E_t \exp\left(\eta_{it}^{j}\right)\left(P_{it}^{jg}\right)^{-(1+\alpha)}}{\sum_{l=1}^{G}\sum_{k\in\Lambda_t^l}\exp\left(\eta_{mt}^{kl}\right)\left(P_{mt}^{kl}\right)^{-\alpha}}$, which is a quality-adjusted variation of the CES demand system that is commonly used in the international trade literature. While this specification of the demand function is often used in applied work for it's simplicity, note that setting $\sigma = 1$ results in conditional choice probabilities that satisfy the independence of irrelevant alternatives (IIA) property, which means that following an increase in the price of any item j, choice probabilities for all varieties $k \neq j$ will increase by the exact same proportion.[119] Given the Cobb-Douglas form of the consumer-level conditional demand function, this also implies a proportionate increase in overall demand for all items $k \neq j$. While these substitution patterns may be appropriate for sufficiently similar products, note that most firms produce products that belong to multiple product classes, such as, for example, rubber insulated cables (ASICC 77475) and PVC insulated Cables (ASICC 77471) . As a result, to use this simplified demand system, one must be willing to assume that cross-price elasticities between these two types of cables are the same as the cross-price elasticities between different *varieties* of a particular class of cable, which is unlikely to hold in practice.

To get around the unrealistic cross-price elasticities generated by the IIA property, one introduces correlation in the individual-level tastes shocks by allowing $\sigma \in (0, 1)$. As long as taste shocks are correlated within product groups, this means that, in the aggregate, some consumers generally prefer rubber cables to PVC cables, and vice-versa. This then leads to cross-price elasticities that are *larger* within product groups than across product groups, since a change in the price of a particular variety of PVC cable will primarily lead to product substitution by the consumers who prefer PVC cables, rather than rubber cables.

[119]To see this, note that relative choice probabilities, $\frac{Prob_t^j}{Prob_t^k}$, only depend on the relative payoffs between the two items, rather than all items in the consumers choice set, as $\frac{Prob_t^j}{Prob_t^k} = \exp\left(\delta_t^j - \delta_t^k\right)$.

8.3.2 Appendix C2: Input Share Inversion Details

To obtain the mapping between demand-side observables and marginal costs for the demand system used in the main text, I need the demand derivatives for (26). These derivatives are given by:[120]

$$\frac{\partial Q_{it}^{jg}}{\partial P_{it}^{jg}} = -\frac{Q_{it}^{jg}}{P_{it}^{jg}}\left(1+\frac{\alpha}{\sigma}-\frac{(1-\sigma)\,\alpha RS_{it}^{j|g}}{\sigma}-\alpha RS_{it}^{jg}\right) \qquad \text{if } j\in\Lambda_t^g, \tag{68}$$

$$\frac{\partial Q_{it}^{jg}}{\partial P_{it}^{kl}} = \begin{cases} \frac{Q_{it}^{jg}}{P_{it}^{kg}}\left(\frac{(1-\sigma)\alpha RS_{it}^{k|g}}{\sigma}+\alpha RS_{it}^{kg}\right) & \text{if } (j,k)\in\Lambda_t^g \\ \frac{Q_{it}^{jg}}{P_{it}^{kl}}\alpha RS_{it}^{kl} & \text{if } j\in\Lambda_{gt},\ k\in\Lambda_t^l, l\neq g, \end{cases} \tag{69}$$

where $RS_{it}^{jg} \equiv \frac{P_{it}^{jg}Q_{it}^{jg}}{E_t^h}$ is the overall revenue market share of product $j\in\Lambda_t^g\subset\Omega_t^h$, while $RS_{it}^{j|g}\equiv\frac{P_{it}^{jg}Q_{it}^{jg}}{\sum_{k\in\Lambda_t^g}P_{it}^{kg}Q_{it}^{kg}}$ is the within-nest revenue share of product j.

I then use these demand derivatives, combined with the firm's pricing first-order conditions, to find the marginal cost inversion necessary to obtain the within-firm input allocations. Note, however, that to account for product specific taxes, which I observe in my data, I need to make a small adjustment to the baseline model discussed in the main text. Let τ_{it}^j denote the *ad-valorem* tax rate on product $j\in\mathbb{Y}_{it}$, which I assume takes the form of sales tax that consumers must pay on top of the listed price. The per-unit price that consumers must pay for some good j overall is then $P_{it}^{jC}=P_{it}^j(1+\tau_{it}^j)$, where P_{it}^j is the price set (and received) at the plant level. Since $P_{it}^{jC}\neq P_{it}^j$, the firm's first order pricing conditions become:

$$Q_{it}^j+(1+\tau_{it}^j)\sum_{k\in\mathbb{Y}_{it}}\frac{\partial Q_{it}^k}{\partial P_{it}^{jC}}\left(P_{it}^k-MC_{it}^k\right)=0, \tag{70}$$

where the $(1+\tau_{it}^j)$ in (70) reflects the fact that $\frac{\partial Q_{it}^k}{\partial P_{it}^j}=\frac{\partial Q_{it}^k}{\partial P_{it}^{jC}}\frac{\partial P_{it}^{jC}}{\partial P_{it}^j}=\frac{\partial Q_{it}^k}{\partial P_{it}^{jC}}(1+\tau_{it}^j)$.

Stacking the above in vector notation:

$$\vec{\tilde{Q}}_{it}+\Delta_{it}\left(\vec{P}_{it}-\vec{MC}_{it}\right)=0, \tag{71}$$

[120] Note that if we interpret consumers as downstream purchases of aggregate materials, as described in footnote 118, then these demand derivatives assume that Machinery, Equipment, and Parts producers take downstream expenditure levels on materials as exogenous. This can be justified by treating the downstream output market equilibrium as *given*. In particular, suppose that downstream production is done using a (potentially unknown) Cobb Douglas production function, with (63) corresponding to the Cobb-Douglas materials aggregator for some downstream firm c. Further suppose that downstream firms also compete in the manner described by (4) in the main text. Then one can use the cost-minizing condition for materials, (6), to show that $W_{ct}^M M_{ct}^{jD}=\beta_M^D\frac{P_{ct}^{jD}Q_{ct}^{jD}}{\mu_{ct}^{jD}}$, where β_M^D is the Cobb-Douglas materials coefficient for downstream firms, $P_{ct}^{jD}Q_{ct}^{jD}$ is total revenue earned by some downstream firm c in product line j, μ_{ct}^{jD} is the markup they charge on this product line, i.e. $\left(\frac{P_{ct}^{jD}}{MC_{ct}^{jD}}\right)$, and $W_{ct}^M M_{ct}^{jD}$ is total expenditure on aggregate materials for product line j in the downstream firm c. Note that total income spent on materials is pinned down by the prices, quantities, and markups charged by the downstream firms in the output market. Hence, as long as upstream firms in the Machinery, Equipment, and Parts industry take the downstream output market equilibirum as given, then each individual producer's "income", i.e. $W_{ct}^M M_{ct}^{jD}$, is exogenously given from their point of view.

where $\vec{P}_{it}$, and $\vec{MC}_{it}$ are column vectors of prices and marginal costs, $\vec{\tilde{Q}}_{it}$ is the vector of quantity sold over $1+\tau_{it}^{j}$, i.e. $\frac{Q_{it}^{j}}{1+\tau_{it}^{j}}$, and Δ_{it} is a square matrix of own and cross-price demand derivatives for all the $J_{it} = |\mathbb{Y}_{it}|$ products produced by firm i, i.e.:

$$\Delta_{it} = \begin{bmatrix} \frac{\partial Q_{it}^{i_1}}{\partial P_{it}^{i_1}} & \frac{\partial Q_{it}^{i_2}}{\partial P_{it}^{i_1}} & \cdots & \frac{\partial Q_{it}^{i_{J_{it}}}}{\partial P_{it}^{i_1}} \\ \frac{\partial Q_{it}^{i_1}}{\partial P_{it}^{i_2}} & \frac{\partial Q_{it}^{i_2}}{\partial P_{it}^{i_2}} & \cdots & \frac{\partial Q_{it}^{i_{J_{it}}}}{\partial P_{it}^{i_2}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial Q_{it}^{i_1}}{\partial P_{it}^{i_{J_{it}}}} & \frac{\partial Q_{it}^{i_2}}{\partial P_{it}^{i_{J_{it}}}} & \cdots & \frac{\partial Q_{it}^{i_{J_{it}}}}{\partial P_{it}^{i_{J_{it}}}} \end{bmatrix}, \tag{72}$$

where i_j is the jth element of $\mathbb{Y}_{it}$.

Since (26) satisfies Assumption 7, and therefore Δ_{it} is invertible, (71) can be rewritten as:

$$\vec{MC}_{it} = (\Delta_{it})^{-1} \vec{\tilde{Q}}_{it} + \vec{P}_{it} = g\left(\vec{P}_{it}, \vec{\tilde{Q}}_{it}, \vec{RS}_{it}, \vec{RS}_{it}^{g}, \alpha, \sigma\right), \tag{73}$$

where the second equality follows from (68) and (69). Hence, using (73), one can rewrite (9) as:

$$X_{it}^{j} = \frac{g_{it}^{j}\left(\vec{P}_{it}, \vec{\tilde{Q}}_{it}, \vec{RS}_{it}, \vec{RS}_{it}^{g}, \alpha, \sigma\right) Y_{it}^{j}}{\sum_{k\in\mathbb{Y}_{it}} g_{it}^{k}\left(\vec{P}_{it}, \vec{\tilde{Q}}_{it}, \vec{RS}_{it}, \vec{RS}_{it}^{g}, \alpha, \sigma\right) Y_{it}^{k}} X_{it}, \tag{74}$$

where the equilibrium value of each g_{it}^{j} function can be determined from (73), given (α, σ).

8.4 Appendix D: Revenue TFP

In this section I show how one can use the structure of the production function and demand function to generate a rescaled production function whose residual is a combination of productivity and demand shifters, following Klette and Griliches (1996) and De Loecker (2011).

Let $\exp\left(\widetilde{\eta}_{it}^{jg}\right) \equiv \left(\frac{E_t^h \exp\left(\frac{\eta_{it}^{jg}}{\sigma}\right)}{\left(\sum_{k\in\Lambda_t^g} \exp\left(\frac{\delta_{mt}^{kg}}{\sigma}\right)\right)^{1-\sigma} \sum_{l\in\Omega_t^h}\left(\sum_{k\in\Lambda_t^l} \exp\left(\frac{\delta_{mt}^{kl}}{\sigma}\right)\right)^{\sigma}}\right)^{\frac{1}{1+\frac{\alpha}{\sigma}}}$, and substitute this expression into (26), yielding $Q_{it}^{jg} = \exp(\widetilde{\eta}_{it}^{jg}) \left(P_{it}^{jg}\right)^{-\left(1+\frac{\alpha}{\sigma}\right)}$. One can then use this expression to write product-level revenues as:

$$R_{it}^{jg} = \exp\left(\widetilde{\eta}_{it}^{jg}\right) \left(Q_{it}^{jg}\right)^{-\frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}}. \tag{75}$$

Substituting $Q_{it}^{jg} = Y_{it}^{j} = \exp\left(\omega_{it}^{jg}\right) \left(L_{it}^{jg}\right)^{\beta_L} \left(K_{it}^{jg}\right)^{\beta_K} \left(M_{it}^{jg}\right)^{\beta_M}$ into (75), taking logs and rearranging yields:[121]

[121] Note that this formula does not take into accounts changes in inventories, although can be adapted to account for this.

$$\frac{R_{it}^{jg}}{\left(L_{it}^{jg}\right)^{\widetilde{\beta}_L}\left(K_{it}^{jg}\right)^{\widetilde{\beta}_K}\left(M_{it}^{jg}\right)^{\widetilde{\beta}_M}} = \exp\left(\widetilde{\eta}_{it}^{jg} + \frac{\frac{\alpha}{\sigma}\omega_{it}^{jg}}{1+\frac{\alpha}{\sigma}}\right) \equiv TFPR_{it}^{jg}, \tag{76}$$

where $\widetilde{\beta}_X \equiv \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\beta_X$ for $X \in (L, K, M)$. Note that $TFPR_{it}^{jg}$ varies across products due to both productivity variation, $\widetilde{\omega}_{it}^{jg} \equiv \frac{\frac{\alpha}{\sigma}\omega_{it}^{jg}}{1+\frac{\alpha}{\sigma}}$, as well demand variation, $\widetilde{\eta}_{it}^{jg}$, which is driven by differences in quality, average product-appeal at the 5-digit ASICC code, as well as product-space congestion effects.[122]

Note, however, that TFPR variation within a 5-digit product code and year is simply a weighted sum of the quality and TFPQ residuals from (28) and (31), respectively. To see this, take logs and demean (76) by product-code and year, yielding:

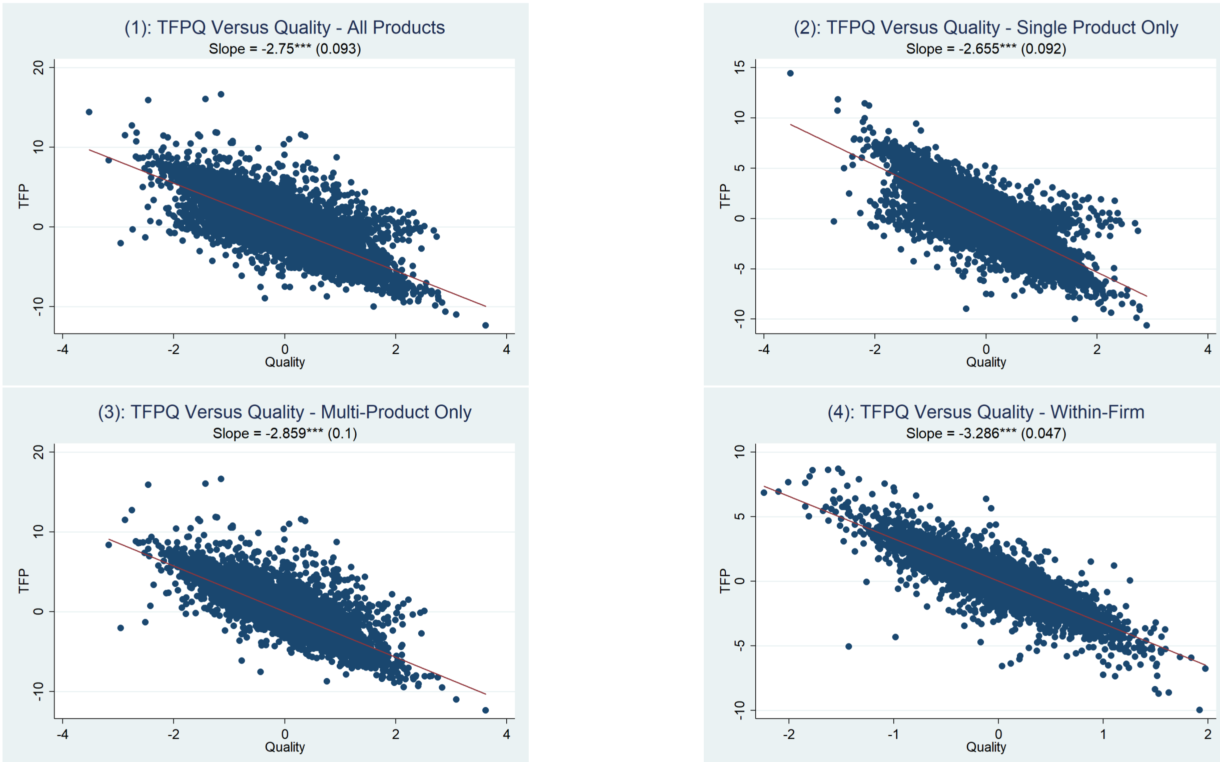
$$\ln\left(TFPR_{it}^{jg}\right) - \ln\left(\overline{TFPR}_t^g\right) = \frac{1}{\sigma+\alpha}\left(\eta_{it}^{jg} - \overline{\eta}_t^g\right) + \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\left(\omega_{it}^{jg} - \overline{\omega}_t^g\right) - \frac{1}{1+\frac{\alpha}{\sigma}}RC_t^g, \tag{77}$$

where $RC_t^g \equiv \ln\left(E_t^h\right) - (1-\sigma)\ln\left(\sum_{k\in\Lambda_t^g}\exp\left(\frac{\delta_{it}^{kg}}{\sigma}\right)\right) - \ln\left(\sum_{l=1}^{G}\left(\sum_{k\in\Lambda_t^l}\exp\left(\frac{\delta_{it}^{kg}}{\sigma}\right)\right)^{\sigma}\right)$, while $\overline{\eta}_t^g$ and $\overline{\omega}_t^g$ are the average values of quality and TFPQ, respectively, within a product code-year.

[122]Note that the demand shifter $\widetilde{\eta}_{it}^{jg}$ also incorporates product-space congestion effects, since increases in $\left(\sum_{k\in\Lambda_t^g}\exp\left(\frac{\delta_{it}^{kg}}{\sigma}\right)\right)^{1-\sigma}$ and $\sum_{l\in\Omega_t^h}\left(\sum_{k\in\Lambda_t^l}\exp\left(\frac{\delta_{it}^{kg}}{\sigma}\right)\right)^{\sigma}$, caused by an increase in the number of competitors at the 5-digit and 3-digit ASICC code, respectively, will scale down $\widetilde{\eta}_{it}^{j}$.

8.5 Appendix E: Results with Materials Measured in Value Units

Figure 7: Negative Correlation between TFPQ and Quality: Materials in Value Units



Notes: Standard errors of slopes, adjusted for two-way clustering by plant and product code, reported in parentheses. Quality refers to the year-product code demeaned residual from (28), while TFPQ refers to the year-product code demeaned residual from (31), with materials measured in value, rather than quantity, units. Panel (1) includes all products, Panel (2) only includes single-product plants, and Panels (3) and (4) only includes products produced by multi-product plants. Panel (4) plots within-firm demeaned values of TFPQ and quality.

Table 23: Tariffs And Within-Firm Efficiency

	(1) Quality Specialization	(2) Relative TFPQ Specialization	(3) Relative TFPR Specialization
Tariffs_{it}	-0.0141*	0.0161**	0.00199
	(0.00768)	(0.00762)	(0.00478)
Observations	1,745	1,745	1,745

Standard errors clustered by plant

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

Notes: Each column corresponds to an OLS regression of the listed outcome variable on firm-level average tariffs as described in the main text. All regressions include year, plant, number of product code and number of variety fixed effects. Quality Specialization measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\eta}_{it}^{\vec{j}|g}\right)$, while Relative TFPQ Specialization measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\omega}_{it}^{\vec{j}|g}\right)$ - see equation (41). High TFPR efficiency measured by $OP_{it}\left(\vec{S}_{it}, h_{it}^{\vec{j}|g}\right)$. Tariffs_{it} measured as firm-level average of tariffs applied to each product-code produced within a plant. TFPQ calculated by measuring materials in revenue units. Standard errors treat demand and production function estimates as data.

Table 24: Import Competition And Within-Firm Efficiency (OLS)

	(1) High Quality Specialization	(2) Relative TFPQ Specialization	(3) High RTFP Specialization
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.0105	-0.0163*	-0.00583
	(0.00816)	(0.00766)	(0.00751)
Observations	1,710	1,710	1,710

Standard errors clustered by plant

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

Notes: Each column corresponds to an OLS regression of the listed outcome variable on firm-level import competition as described in the main text.All regressions include year, plant, number of product code and number of variety fixed effects. Quality Specialization measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\eta}_{it}^{\vec{j}|g}\right)$, while Relative TFPQ Specialization measured by $OP_{it}\left(\vec{S}_{it}, \tilde{\omega}_{it}^{\vec{j}|g}\right)$ - see equation (41). Relative TFPR specialization measured by $OP_{it}\left(\vec{S}_{it}, h_{it}^{\vec{j}|g}\right)$. TFPQ calculated by measuring materials in revenue units. Standard errors treat demand and production function estimates as data.

Table 25: Import Competition And Within-Firm Efficiency (First Stage and IV)

	(1) $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$	(2) High Quality Specialization	(3) Relative TFPQ Specialization	(4) Relative TFPR Specialization
Z^{China}_{it}	0.767***			
	(0.103)			
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$		0.0890**	-0.104***	-0.0147
		(0.0398)	(0.0386)	(0.0226)
Observations	1,710	1,710	1,710	1,710

Standard errors clustered by plant

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

Notes: Column (1) reports the first stage regression for the instrumental variables regressions discussed in the main text. AP F-statistic for excluded instrument equals 55.31. Columns (2) through (6) correspond to an instrumental variables regression om the listed outcome variable, where $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by Z^{CHN}_{it} , defined by equation (44). All regressions include year, plant, number of product code and number of variety fixed effects. Quality Specialization measured by $OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\eta}}^{j|g}_{it}\right)$, while Relative TFPQ Specialization measured by $OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\omega}}^{j|g}_{it}\right)$ - see equation (41). Relative TFPR efficiency measured by $OP_{it}\left(\vec{S}_{it}, \vec{h}^{j|g}_{it}\right)$. TFPQ calculated by measuring materials in revenue units. Standard errors treat demand and production function estimates as data.

8.6 Appendix F: Robustness - Net Impact of Chinese Import Competition on Plant Performance

Table 26: Decomposing Impact of Chinese Imports of Firm Performance: Endogenous TFPQ

	(1) Product Code Specialization	(2) TFPR Specialization	(3) Average Performance	(4) Total Performance	(5) Standard Plant TFPR
$\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.0407*	0.0148	-0.105	-0.0900	-0.103*
	(0.0220)	(0.0155)	(0.0875)	(0.0809)	(0.0576)
Observations	1,710	1,710	1,710	1,710	1,710

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the listed outcome variable, where $\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by Z^{CHN}_{it} , defined by equation (44). All regressions include year, plant, number of product code and number of variety fixed effects. Product code specialization is measured by $OP_{it}\left(\vec{S}_{it}, \vec{h}^{g}_{t}\right) = OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\omega}}^{g}_{t}\right) + OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\eta}}^{g}_{t}\right)$, while total specialization is measured by $OP_{it}\left(\vec{S}_{it}, \vec{h_{it}}\right) = OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\eta}}^{j|g}_{it}\right) + OP_{it}\left(\vec{S}_{it}, \vec{\tilde{\eta}}^{j|g}_{it}\right) + OP_{it}\left(\vec{S}_{it}, \vec{h}^{g}_{t}\right)$. Average TFPR is the plant-level unweighted average of variety-level TFPR, $\overline{h}_{it}$, while total TFPR is given by $h_{it} = \overline{h}_{it} + OP_{it}\left(\vec{S}_{it}, \vec{h}_{it}\right)$. Standard Plant TFPR is measured as $\ln\left(\sum_{j\in\mathbb{Y}_{it}} P^{j}_{it}Q^{j}_{it}\right) - \frac{\frac{\alpha}{\sigma}}{1+\frac{\alpha}{\sigma}}\left(\beta_L l_{it} + \beta_K k_{it} + \beta_M m_{it}\right)$. Standard errors treat demand and production function estimates as data.

In Table 26, I estimate the same decomposition of the various dimensions of plant-level productivity growth following Chinese import shocks as described in Table 11, after re-estimating the production function parameters using the endogenous productivity specification described in Section 6.3.3.

8.7 Appendix G: OP Covariance Decompositions

For notational convenience, first note that:

$$\begin{aligned} OP_{it}\left(\vec{S}_{it},\vec{x}_{it}\right) &= \sum_{j\in\mathbb{Y}_{it}}\left(x_{it}^{jg}-\frac{1}{J_{it}}\sum_{j\in\mathbb{Y}_{it}}x_{it}^{jg}\right)\left(S_{it}^{jg}-\frac{1}{J_{it}}\sum_{j\in\mathbb{Y}_{it}}S_{it}^{jg}\right)\\ &= \sum_{j\in\mathbb{Y}_{it}}\underbrace{\left(x_{it}^{jg}-\frac{1}{J_{it}}\sum_{j\in\mathbb{Y}_{it}}x_{it}^{jg}\right)}_{\equiv\widehat{x}_{it}^{jg}}S_{it}^{jg}-\underbrace{\frac{1}{J_{it}}\sum_{j\in\mathbb{Y}_{it}}\left(x_{it}^{jg}-\frac{1}{J_{it}}\sum_{j\in\mathbb{Y}_{it}}x_{it}^{jg}\right)}_{=0}\\ &= \sum_{j\in\mathbb{Y}_{it}}S_{it}^{jg}\widehat{x}_{it}^{jg}. \end{aligned} \tag{78}$$

Next, partition the set of products produced at time t and $t-1$ into three non overlapping categories: 1) *Dropped products*, denoted by $D_{it}\subset\mathbb{Y}_{i,t-1}$, and corresponding to the set of products that were produced in $t-1$ but dropped at time t, 2) *New products*, denoted by $N_{it}\subset\mathbb{Y}_{it}$, and corresponding to the set products that were produced at time t but not produced at time $t-1$, and 3) *Constant products*, denoted by C_{it}, and corresponding to the set of products that were produced in both $t-1$ and t. Further define the within-category $X\in(N,D,C)$ share as $S_{it}^{jg|X}\equiv\frac{S_{it}^{jg}}{S_{it}^{X}}$, where S_{it}^{X} denotes the total input share of products belonging to category $X\in(N,D,C)$ at time t. Note that by construction $S_{i,t-1}^{D}+S_{i,t-1}^{C}=1$ and $S_{it}^{N}+S_{it}^{C}=1$. One can then write:

$$OP_{i,t-1}\left(\vec{S}_{i,t-1},\vec{x}_{i,t-1}\right)=S_{i,t-1}^{C}\sum_{j\in C_{it}}S_{i,t-1}^{jg|C}\,\widehat{x}_{i,t-1}^{jg}+S_{i,t-1}^{D}\sum_{j\in D_{it}}S_{i,t-1}^{jg|C}\,\widehat{x}_{i,t-1}^{jg}, \tag{79}$$

and:

$$\begin{aligned} OP_{it}\left(\vec{S}_{it},\vec{x}_{it}\right) &= S_{it}^{C}\sum_{j\in C_{it}}S_{it}^{jg|C}\,\widehat{x}_{it}^{jg}+S_{it}^{N}\sum_{j\in N_{it}}S_{it}^{jg|N}\,\widehat{x}_{it}^{jg}\\ &= S_{i,t-1}^{C}\sum_{j\in C_{it}}S_{it}^{jg|C}\,\widehat{x}_{it}^{jg}+\Delta S_{it}^{C}\sum_{j\in C_{it}}S_{it}^{jg|C}\,\widehat{x}_{it}^{jg}+S_{it}^{N}\sum_{j\in N_{it}}S_{it}^{jg|N}\,\widehat{x}_{it}^{jg}\\ &= S_{i,t-1}^{C}\sum_{j\in C_{it}}S_{it}^{jg|C}\,\widehat{x}_{it}^{jg}+S_{i,t-1}^{D}\sum_{j\in C_{it}}S_{it}^{jg|C}\,\widehat{x}_{it}^{jg}+S_{it}^{N}\left(\sum_{j\in N_{it}}S_{it}^{jg|N}\,\widehat{x}_{it}^{jg}-\sum_{j\in C_{it}}S_{it}^{jg|C}\,\widehat{x}_{it}^{jg}\right), \end{aligned} \tag{80}$$

where $\Delta S_{it}^{C}\equiv S_{it}^{C}-S_{i,t-1}^{C}=S_{i,t-1}^{D}-S_{it}^{N}$.

First differencing (80) and (79) yields:

$$\Delta OP_{it}\left(\vec{S}_{it},\vec{x}_{it}\right) = S^{C}_{i,t-1}\left(\sum_{j\in C_{it}} S^{jg|C}_{it}\,\widehat{x}^{jg}_{it} - \sum_{j\in C_{it}} S^{jg|C}_{i,t-1}\,\widehat{x}^{jg}_{i,t-1}\right)$$
$$+ S^{D}_{i,t-1}\left(\sum_{j\in D_{it}} S^{jg|C}_{it}\,\widehat{x}^{jg}_{it} - \sum_{j\in D_{it}} S^{jg|D}_{i,t-1}\,\widehat{x}^{jg}_{i,t-1}\right) \quad (81)$$
$$+ S^{N}_{it}\left(\sum_{j\in N_{it}} S^{jg|N}_{it}\,\widehat{x}^{jg}_{it} - \sum_{j\in C_{it}} S^{jg|C}_{it}\,\widehat{x}^{jg}_{it}\right).$$

Adding and subtracting $S^{D}_{i,t-1}\left(\sum_{j\in C_{it}} S^{jg|C}_{it}\,\widehat{x}^{jg}_{it} - \sum_{j\in C_{it}} S^{jg|C}_{i,t-1}\,\widehat{x}^{jg}_{i,t-1}\right)$ from (81) then yields, after cancelling out like terms:

$$\Delta OP_{it}\left(\vec{S}_{it},\vec{x}_{it}\right) = \sum_{j\in C_{it}} S^{jg|C}_{it}\,\widehat{x}^{jg}_{it} - \sum_{j\in C_{it}} S^{jg|C}_{i,t-1}\,\widehat{x}^{jg}_{i,t-1}$$
$$+ S^{D}_{it}\left(\sum_{j\in C_{it}} S^{jg|C}_{i,t-1}\,\widehat{x}^{jg}_{i,t-1} - \sum_{j\in D_{it}} S^{jg|D}_{i,t-1}\,\widehat{x}^{jg}_{i,t-1}\right)$$
$$+ S^{N}_{it}\left(\sum_{j\in N_{it}} S^{jg|N}_{it}\,\widehat{x}^{jg}_{it} - \sum_{j\in C_{it}} S^{jg|C}_{it}\,\widehat{x}^{jg}_{it}\right),$$

which corresponds to (49) in the main text.

Appendix G1: Constant Chinese Import Competition Production Sets

Table 27: Plant-Level Changes in Quality Specialization: Intensive versus Extensive Margin

	(1) Total Specialization	(2) Intensive Margin	(3) Product Dropping	(4) New Products	(5) Total Specialization	(6) Intensive Margin	(7) Product Dropping	(8) New Products
$\Delta\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.203	-0.396	0.759	-0.160				
	(0.583)	(0.725)	(1.174)	(0.593)				
$\Delta\ln\left(\sum_{g\in\Lambda^{CD}_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$					-0.0144	0.0425	-0.0656	0.00868
					(0.0441)	(0.0421)	(0.0534)	(0.0470)
Observations	989	989	989	989	985	985	985	985

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the described outcome, where $\Delta\ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}}\text{IM}^{g}_{\text{IND,CHN},t}\right)$ or $\Delta\ln\left(\sum_{g\in\Lambda^{CD}_{it}}\text{IM}^{g}_{\text{China},t}\right)$ is instrumented by $\Delta Z^{\text{ChinaCD}}_{it} = \Delta\ln\left(\sum_{g\in\Lambda^{DC}_{it}}\sum_{k\in\mathbb{C}^{\text{No India}}}\text{IM}^{g}_{\text{k,CHN},t}\right)$, where Δ is the time differencing operator, and Λ^{CD}_{it} is the set of product codes belonging to the dropped and constant product sets, D_{it} and C_{it}. All regressions include current and lagged (according to plant time) fixed effects for the number of product codes, number of varieties, and calendar year. AP F-statistic for excluded instrument equals 0.55 for columns (1) through (4) and 26.80 for columns (5) through (8). Total Specialization refers to $\Delta OP_{it}\left(\vec{S}_{it},\vec{\widetilde{\eta}}^{j|g}_{it}\right)$, while Intensive Margin, Product Dropping, and New Product refer to terms (I), (D) and (N) in equation (49), respectively, with $x^{jg}_{it} = \widetilde{\eta}^{j|g}_{it}$. Standard errors treat demand and production function estimates as data.

Table 28: Plant-Level Changes in Relative TFPQ Specialization: Intensive versus Extensive Margin

	(1) Total Specialization	(2) Intensive Margin	(3) Product Dropping	(4) New Products	(5) Total Specialization	(6) Intensive Margin	(7) Product Dropping	(8) New Products
$\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	-0.310	0.0521	-0.685	0.323				
	(0.715)	(0.533)	(1.071)	(0.741)				
$\Delta \ln\left(\sum_{g\in\Lambda^{CD}_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$					0.0294	-0.0131	0.0675	-0.0250
					(0.0518)	(0.0446)	(0.0505)	(0.0505)
Observations	989	989	989	989	985	985	985	985

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the described outcome, where $\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$ or $\Delta \ln\left(\sum_{g\in\Lambda^{CD}_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by $\Delta Z^{\text{ChinaCD}}_{it} = \Delta \ln\left(\sum_{g\in\Lambda^{DC}_{it}}\sum_{k\in\mathbb{C}^{\text{No India}}} \text{IM}^{g}_{\text{k,CHN},t}\right)$, where Δ is the time differencing operator, and Λ^{CD}_{it} is the set of product codes belonging to the dropped and constant product sets, D_{it} and C_{it}. All regressions include current and lagged (according to plant time) fixed effects for the number of product codes, number of varieties, and calendar year. AP F-statistic for excluded instrument equals 0.55 for columns (1) through (4) and 26.80 for columns (5) through (8). Total Specialization refers to $\Delta OP_{it}\left(\vec{S}_{it}, \widetilde{\omega}^{\vec{j}|g}_{it}\right)$, while Intensive Margin, Product Dropping, and New Product refer to terms (I), (D) and (N) in equation (49), respectively, with $x^{jg}_{it} = \widetilde{\omega}^{j|g}_{it}$. Standard errors treat demand and production function estimates as data.

8.7.1 Appendix G2: First-Difference Robustness

Table 29: Plant-Level Changes in Quality Specialization: Intensive versus Extensive Margin

	(1) Total Specialization	(2) Intensive Margin	(3) Product Dropping	(4) New Products	(5) Total Specialization	(6) Intensive Margin	(7) Product Dropping	(8) New Products
$\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$	0.176**	0.0256	0.122*	0.0287				
	(0.0691)	(0.0519)	(0.0643)	(0.0443)				
$\Delta \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}} \ln\left(\text{IM}^{g}_{\text{IND,CHN},t}\right)$					0.162**	0.0437	0.141**	-0.0230
					(0.0674)	(0.0498)	(0.0597)	(0.0524)
Observations	667	667	667	667	667	667	667	667

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the described outcome, where $\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by $\Delta Z^{\text{CHN}}_{it}$, defined by equation (44), and $\Delta \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}} \ln\left(\text{IM}^{g}_{\text{IND,CHN},t}\right)$ is instrumented by $\Delta Z^{\text{AltCHN}}_{it} = \Delta \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}} \ln\left(\sum_{k\in\mathbb{C}^{\text{No India}}} \text{IM}^{g}_{k,CHN,t}\right)$, where Δ is the time differencing operator according to calendar year. All regressions include year fixed effects. AP F-statistic for excluded instrument equals 24.76 for columns (1) through (4) and equals 27.43 for columns (5) through (8). Total Specialization refers to $\Delta OP_{it}\left(\vec{S}_{it}, \widetilde{\eta}^{\vec{j}|g}_{it}\right)$, while Intensive Margin, Product Dropping, and New Product refer to terms (I), (D) and (N) in equation (49), respectively, with $x^{jg}_{it} = \widetilde{\eta}^{j|g}_{it}$. Standard errors treat demand and production function estimates as data.

Table 30: Plant-Level Changes in Relative TFPQ Specialization: Intensive versus Extensive Margin

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total Specialization	Intensive Margin	Product Dropping	New Products	Total Specialization	Intensive Margin	Product Dropping	New Products
$\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \mathrm{IM}^g_{\mathrm{IND,CHN},t}\right)$	-0.171***	-0.0210	-0.0971**	-0.0528				
	(0.0642)	(0.0419)	(0.0487)	(0.0478)				
$\Delta \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}} \ln\left(\mathrm{IM}^g_{\mathrm{IND,CHN},t}\right)$					-0.137**	-0.0454	-0.0944**	0.00256
					(0.0649)	(0.0401)	(0.0434)	(0.0517)
Observations	667	667	667	667	667	667	667	667

Standard errors clustered by plant

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

Notes: Each column corresponds to an instrumental variables regression on the described outcome, where $\Delta \ln\left(\frac{1}{G_{it}}\sum_{g\in\Lambda_{it}} \mathrm{IM}^g_{\mathrm{IND,CHN},t}\right)$ is instrumented by $\Delta Z^{\mathrm{CHN}}_{it}$, defined by equation (44), and $\Delta \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}} \ln\left(\mathrm{IM}^g_{\mathrm{IND,CHN},t}\right)$ is instrumented by $\Delta Z^{\mathrm{ChinaAlt}}_{it} = \Delta \frac{1}{J_{it}}\sum_{(j,g)\in\mathbb{Y}_{it}} \ln\left(\sum_{k\in\mathbb{C}^{\text{No India}}} \mathrm{IM}^g_{k,t}\right)$, where Δ is the time differencing operator according to calendar year. All regressions include year fixed effects. AP F-statistic for excluded instrument equals 24.76 for columns (1) through (4) and equals 27.43 for columns (5) through (8). Total Specialization refers to $\Delta OP_{it}\left(\vec{S}_{it}, \widetilde{\omega}^{\vec{j}|g}_{it}\right)$, while Intensive Margin, Product Dropping, and New Product refer to terms (I), (D) and (N) in equation (49), respectively, with $x^{jg}_{it} = \widetilde{\omega}^{j|g}_{it}$. Standard errors treat demand and production function estimates as data.

8.8 Appendix H: The Transformation Function and Optimal Input Allocations

In this section, I compare the approach for estimating productivity in multi-product firms developed in this paper, to the transformation function approach, as in Diewert (1973) and more recently Dhyne et al. (2017). In particular, I show that that Assumptions 1 through 6 imply a simple functional form for the transformation function. However, this function depends on $J_{it} = |\mathbb{Y}_{it}|$ unobservable plant-product specific TFP terms. Therefore, without more restrictions, the transformation function cannot feasibly be estimated. I then show that the approach used in this paper can be characterized as a way to solve to "too many unobservables" problem inherent to the structure of the transformation function.

As per Appendix A, one may define the transformation function as the maximal level of output that can be obtained for some reference good $r \in \mathbb{Y}_{it}$, given aggregate input vector $\vec{X}_{it}$, a vector of firm-product specific TFP terms, $\vec{\omega}_{it}$, and some vector or other outputs $\vec{Y}^{-r}_{it}$, i.e. $Y^r_{it} = \widetilde{T}(\vec{Y}^{-r}_{it}, \vec{X}_{it}, \vec{\omega}_{it})$. It turns out that it is very simple to solve for an equivalent representation of the technology, the *symmetric* for the transformation function (Diewert 1973), given by $Y^r_{it} - \widetilde{T}(\vec{Y}^{-r}_{it}, \vec{X}_{it}, \vec{\omega}_{it}) \equiv T(\vec{Y}_{it}, \vec{X}_{it}, \vec{\omega}_{it}) = 0$. To obtain the symmetric transformation function, note that by Lemma 1, as long as the input allocation problem has a unique solution, then input ratios are constant across production lines, and therefore $X^j_{it} = S^j_{it} X_{it}$ for all $X \in (\mathbb{K}, \mathbb{M})$. Substituting this equation into the production function, and using the fact that the production technology is homogenous of degree $\phi > 0$ yields:

$$Y^j_{it} = \exp(\omega^j_{it})(S^j_{it})^{\phi} F(\vec{X}_{it}).$$

Rearranging the above yields:

$$\left(\frac{Y_{it}^j}{\exp(\omega_{it}^j)}\right)^{\frac{1}{\phi}} = S_{it}^j\left(F(\vec{X}_{it})\right)^{\frac{1}{\phi}}.$$

Finally, summing over all $j \in \mathbb{Y}_{it}$ and then taking this expression to the power of ϕ yields:

$$\left(\sum_{j\in\mathbb{Y}_{it}}\left(\frac{Y_{it}^j}{\exp(\omega_{it}^j)}\right)^{\frac{1}{\phi}}\right)^{\phi} = F(\vec{X}_{it}). \tag{82}$$

Note that (82) is a simple representation of the symmetric transformation function, since we can also use (82) to write $T(\vec{Y}_{it}, \vec{X}_{it}, \vec{\omega}_{it}) = \left(\sum_{j\in\mathbb{Y}_{it}}\left(\frac{Y_{it}^j}{\exp(\omega_{it}^j)}\right)^{\frac{1}{\phi}}\right)^{\phi} - F(\vec{X}_{it}) = 0$.

Note that the symmetric transformation function here is *seperable*, in the sense defined by Hall (1973), i.e. we are able to write the transformation function as the sum of two functions, one of which depends on outputs, and another that depends on inputs, $T(\vec{Y}_{it}, \vec{X}_{it}, \vec{\omega}_{it}) = A_{it}(\vec{Y}_{it}) - B(\vec{X}_{it})$. It is worth noting in passing that one of the key findings in Hall (1973), is that that separable transformation functions are almost always the result *joint* technologies, which are multi-product production processes that *cannot* be written as a series of independent production technologies. Since the production process considered in this paper is inherently *non-joint*, as per Hall (1973), the fact that I obtain a seperable transformation function may appear surprising. Note, however, that the only case where a a non-joint production process leads to a separable transformation function is if the production technology across production lines within a firm only differs by a scalar multiple. This is exactly equivalent to my Assumption 2, which is why the transformation function is seperable.

The key challenge for identifying the shape of transformation function empirically is that (82) has $J_{it} \equiv |\mathbb{Y}_{it}|$ unobservables. As a result, estimation via GMM is immediately unavailable, since it is impossible to simply use (82) to solve for each value of ω_{it}^j. There are, of course, restrictions that one could place on the variation of TFP that could lead to identification. The easiest of these is to, for example, follow De Loecker et al. (2016) and assume $\omega_{it}^j = \omega_{it}\ \forall j \in \mathbb{Y}_{it}$. This leads to the simple transformation function:

$$\left(\sum_{j\in\mathbb{Y}_{it}}\left(Y_{it}^j\right)^{\frac{1}{\phi}}\right)^{\phi} = \exp(\omega_{it})F(\vec{X}_{it}). \tag{83}$$

The above simplified transformation function could immediately be used to identify firm-level TFP using standard non-linear GMM techniques, and would have the added advantage of allowing one use to multi-product firms directly in estimation, rather than relying on a selected sample of single product firms to estimate the production technology. Another approach would be to put some parametric structure on the distribution of the unobservable ω_{it}^j terms, and then using maximum likelihood for estimation. This would have the cost of making endogeneity difficult to deal with, as one would also have to model the correlation between $\vec{\omega_{it}}$ and $\vec{Y}_{it}$.

Another approach, which is equivalent to the approach described in this paper, is to use price variation within the firm to identify within-firm TFP dispersion, and then combine that with (82).

Specifically, rewrite (82) as:

$$\left(\sum_{j\in\mathbb{Y}_{it}}\left(Y_{it}^{j}\right)^{\frac{1}{\phi}}\left(\frac{\exp(\omega_{it}^{r})}{\exp(\omega_{it}^{j})}\right)^{\frac{1}{\phi}}\right)^{\phi}=\exp(\omega_{it}^{r})F(\vec{X}_{it}), \tag{84}$$

where ω_{it}^{r} is productivity of some reference good $r\in\mathbb{Y}_{it}$. To make estimation of (84) feasible, we need to identify relative TFP within the firm, $\frac{\exp(\omega_{it}^{r})}{\exp(\omega_{it}^{j})}$. To do this, first consider the firm's conditional cost minimization problem after having solved for the optimal input allocations given $\vec{X}_{it}$, as described in Appendix A. Rewriting their conditional cost minimzation problem in terms of the symmetric transformation function yields:[123]

$$\begin{aligned}&\underset{\vec{M}_{it}}{\text{Min}} && \sum_{M\in\mathbb{M}}\mathbf{W}^{M}\left(M_{it},A_{it}^{M}\right)M_{it}\\ &\text{subject to} && \left(\sum_{j\in\mathbb{Y}_{it}}\left(\frac{Y_{it}^{j}}{\exp(\omega_{it}^{j})}\right)^{\frac{1}{\phi}}\right)^{\phi}=F(\vec{X}_{it}),\end{aligned} \tag{85}$$

with associated Lagrangian:

$$L=\sum_{M\in\mathbb{M}}\mathbf{W}^{M}\left(M_{it},A_{it}^{M}\right)M_{it}+\lambda\left(\left(\sum_{j\in\mathbb{Y}_{it}}\left(\frac{Y_{it}^{j}}{\exp(\omega_{it}^{j})}\right)^{\frac{1}{\phi}}\right)^{\phi}-F(\vec{X}_{it})\right). \tag{86}$$

Letting $C\left(\vec{K}_{it},\vec{Y}_{it},\vec{\omega}_{it},\vec{A}_{it}\right)$ denote the conditional cost function for static inputs, as before, then then by the envelope theorem:

$$MC_{it}^{j}=\frac{\partial C}{\partial Y_{it}^{j}}=\lambda\left(\sum_{j\in\mathbb{Y}_{it}}\left(\frac{Y_{it}^{j}}{\exp(\omega_{it}^{j})}\right)^{\frac{1}{\phi}}\right)^{\phi-1}\left(\frac{Y_{it}^{j}}{\exp(\omega_{it}^{j})}\right)^{\frac{1}{\phi}}\frac{1}{Y_{it}^{j}}. \tag{87}$$

Using the above expression to examine marginal cost ratios between an arbitrary good j and the reference good r yields $\frac{MC_{it}^{j}}{MC_{it}^{r}}$ then yields:

$$\frac{MC_{it}^{j}}{MC_{it}^{r}}=\left(\frac{\frac{Y_{it}^{j}}{\exp(\omega_{it}^{j})}}{\frac{Y_{it}^{r}}{\exp(\omega_{it}^{r})}}\right)^{\frac{1}{\phi}}\frac{Y_{it}^{r}}{Y_{it}^{j}}. \tag{88}$$

Solving for $\left(\frac{\exp(\omega_{it}^{r})}{\exp(\omega_{it}^{j})}\right)^{\frac{1}{\phi}}$ in this expression yields:

$$\left(\frac{\exp(\omega_{it}^{r})}{\exp(\omega_{it}^{j})}\right)^{\frac{1}{\phi}}=\frac{Y_{it}^{j}MC_{it}^{j}}{Y_{it}^{r}MC_{it}^{r}}\left(\frac{Y_{it}^{r}}{Y_{it}^{j}}\right)^{\frac{1}{\phi}}. \tag{89}$$

Substituting (89) back into the transformation function (84) yields:

[123]Recall that the conditional cost minimization problem takes the level of dynamic inputs, such as capital, as given, and as such the level of dynamic input costs are taken as given, and are therefore suppressed in the formulation of costs.

$$\frac{Y_{it}^r}{(Y_{it}^r MC_{it}^r)^\phi}\left(\sum_{j\in\mathbb{Y}_{it}} Y_{it}^j MC_{it}^j\right)^\phi = \exp(\omega_{it}^r)F(\vec{X}_{it}), \tag{90}$$

or, after re-arranging and using the fact that the production function is homogeneous of degree $\phi > 0$

$$Y_{it}^r = \left(\frac{Y_{it}^r MC_{it}^r}{\sum_{j\in\mathbb{Y}_{it}} Y_{it}^j MC_{it}^j}\right)^\phi \exp(\omega_{it}^r)F(\vec{X}_{it}) = \exp(\omega_{it}^r)F\left(\frac{Y_{it}^r MC_{it}^r}{\sum_{j\in\mathbb{Y}_{it}} Y_{it}^j MC_{it}^j}\vec{X}_{it}\right). \tag{91}$$

Since marginal costs can be estimated from price and quantity data, there is only one unobservable in (91), the firm-product level TFP term, and therefore standard GMM techniques can be used. More important, note that (91) equivalent to the firm-product level production function equation after using the input allocation rule (9) to estimate input shares. As a result, the approach used in this paper to estimate productivity in multiproduct firms can also be characterized as using pricing (and therefore marginal cost information) to pin down relative TFP within a firm, to make estimation of the transformation function feasible.