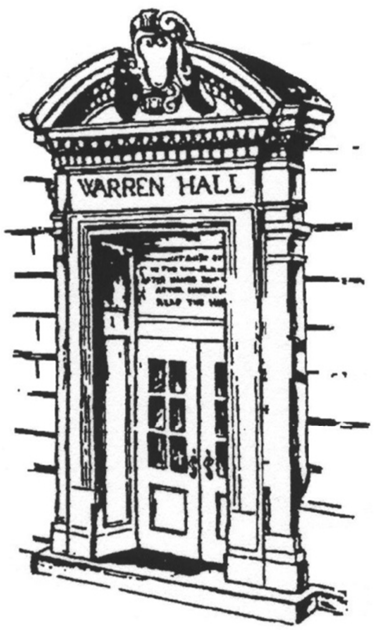
WP 2014-21
November 2014



Working Paper

The Charles H. Dyson School of Applied Economics and Management
Cornell University, Ithaca, New York 14853-7801 USA

Informality among multi-product firms

Dennis Becker

Informality among multi-product firms*

Dennis Becker†
Cornell University

November 14, 2014

Abstract

This paper introduces product-level regulation as a new driver of informality and diversification in a model of heterogeneous multi-product firms and endogenous product choice. Firms face regulations at both the firm- and product-level and may comply with or evade either regulation. The model suggests that firm-level regulation directly causes informality by deterring firm registration. However, the product-level regulation has two effects: it directly drives product informality as evasion of product regulation leading to informality within the formal sector and indirectly deters firms from registering. Further, I demonstrate that the Gini coefficient and Herfindahl index can be implemented in multi-product firm models as revenue-based measures of product diversification. Contrary to the prediction of the commonly used product scope, the revenue-based measures indicate informal firms to be more diversified than formal firms.

JEL-Classification: L2, L5, O17

Keywords: heterogeneous firms; informality; regulations; diversification

*I am grateful to Ravi Kanbur, Nancy Chau, Ankita Patnaik and the 2014 NEUDC participants for helpful comments and suggestions.

†Charles H. Dyson School of Applied Economics and Management, B38 Warren Hall, Cornell University, Ithaca, NY 14853. E-Mail: db653@cornell.edu.

"The transition from informal to formal enterprise status is also gradual; indeed, single firms [...] can carry out some activities informally and others formally at the same time."
de Beer et al. (2013, p. 13)

1 Introduction

The extant theoretical literature employs a dichotomous definition of informality: firms either comply with or evade regulations such as firm registration or taxation.[1] However, in reality firms face a multitude of regulations, and informality is a multi-dimensional concept. A recent study by the World Intellectual Property Organization on the traditional medicine practitioner (TMP) industry in Ghana, where informality accounts for about 40% of the country's GDP (Schneider et al., 2010), highlights this. It notes, "[e]ven though there is a significantly high level of business registration among the TMPs - and this is done in conformity to the legal requirements for practicing - the level of formalization of their business was limited." (Essegbey et al. (2014, p. 15)) More specifically, while 67% of surveyed TMPs had registered their business, only 52% had registered at least one of their products with the Food and Drugs Authority, a legal requirement in Ghana.

The empirical literature provides evidence on four drivers of these different types of informality and the product choice of firms. First, firm-level regulations that are costly to comply with deter firm registration and accordingly drive informal sector participation at the firm-level, as found in the seminal work of de Soto (1989).[2] Second, in a similar fashion, product market regulations drive firm-level informality (Loayza et al., 2005). Third, in the case of Ghana where registered, and hence formal, firms did not register some of their products, product-level regulation matters: Ghana's TMPs state "the cumbersome nature of the registration procedures and the fees charged being too high" (Essegbey et al. (2014, p. 2)) as the reasons for the evasion of regulation and hence informality at the product-level. Fourth, product-level costs and regulation have been found to be determinants of firms' product choices in the recent literature on multi-product firms (Bernard et al., 2010; Goldberg et al., 2010), and consequently informal firms seem to be less diversified than formal firms (CIEM, 2012).[3] This indicates two things. First, multiple policies impact firms at different levels and accordingly their product choice and informality decision. Second, there exists not just an informal sector, but also informality within the formal sector in the form of unregistered products

[1]This literature considers, for instance, the effect of informality on unemployment (Fields, 1975), size dualism (Rauch, 1991), quality dualism (Banerji and Jain, 2007), contractual dualism (Basu et al., 2011) and trade liberalization (Becker, 2014).

[2]Several subsequent papers confirm firm registration cost and complex registration procedures as main drivers of informality (Djankov et al., 2002; Auriol and Warlters, 2005; Antunes and de V. Cavalcanti, 2007; Ulyssea, 2010).

[3]Other key findings from this literature are the connection of firm productivity with product scope and diversification (Schoar, 2002; Iacovone and Javorcik, 2008; Bernard et al., 2009, 2010), adjustment of product scope and scale of firms upon trade liberalization (Feenstra and Ma, 2007; Baldwin and Gu, 2009; Arkolakis and Muendler, 2010; Eckel and Neary, 2010; Bernard et al., 2011; Mayer et al., 2011; Nocke and Yeaple, 2013) and the importance of within-firm distribution of resources among products as drivers of aggregate output of an economy (Bernard et al., 2010; Goldberg et al., 2010; Navarro, 2012; Söderbom and Weng, 2012).

produced by registered firms. Given the focus of the extant theoretical informality literature on only firm-level regulation and as no model to date captures all the aforementioned empirical facts, the effect of firm- and product-level regulations jointly require further scrutinity. Therefore, I develop a model of heterogeneous multi-product firms that endogenizes product choice and informality at both the product- and firm-level to answer two questions: What are the impacts of both firm- and product-level regulations on informality? How does product-level regulation impact firms' product choice and diversification in the presence of informality?

This model, along the lines of Bernard et al. (2010), provides a tractable setup to answer these questions and consists of four core components. First, firms are heterogeneous in productivity and, conditional on firm productivity, are characterized by heterogeneous skills in the production of a continuum of goods. Second, firms are exposed to both firm- and product-level regulation and firms' decision on whether to comply with or evade regulation occurs at two levels. Informality at the firm-level is defined to be firm registration non-compliance. Similarly, informality at the product-level is defined to be product registration non-compliance. This allows for registered, and hence formal, firms that may produce goods informally by evading product-level regulation corresponding to Essegbey et al. (2014).[4] Third, formality decisions at both levels are made under the consideration of the costs and benefits of formality and informality. Firms operate rationally and informality at both levels is therefore an entrepreneurial choice (e.g. de Mel et al., 2013; La Porta and Shleifer, 2014). Fourth, in line with the empirical diversification literature, the model features three indicators of product diversification to examine firms' product choice: product scope, a Product-Gini and a product-level Herfindahl index.[5]

The model results suggest that in addition to firm-level regulation, product-level regulation is also an important driver of informality. While firm-level regulation has a direct effect on firms' decision to register their business, an increase in the cost to comply with the product-level regulation incentivizes firms to informalize along two dimensions. First, higher product regulation cost has a direct effect in making the production of goods in compliance with registration requirements more costly and leads to fewer registered goods. Second, an increase in this cost has an indirect effect by decreasing the profitability of registering a business, and therein having the option to produce registered goods.

Furthermore, the model highlights the effect of product-level regulation on firms' product choice and diversification. In terms of product scope, i.e. the number of distinct products produced by

[4]The complete evasion of product registration of informal firms seems particularly fitting for Ghana's TMP, given that the sector, as described in Essegbey et al. (2014), was originally entirely informal.

[5]The empirical literature on multi-product firms and international trade commonly measures diversification by product scope, that is the number of distinct products per firm (Iacovone and Javorcik, 2008; Bernard et al., 2009; Arkolakis and Muendler, 2010; Bernard et al., 2010; Goldberg et al., 2010; Bernard et al., 2011). In the literature on sectoral and export diversification two alternative measures are commonly employed: the Herfindahl index as absolute (Imbs and Wacziarg, 2003; Klinger and Lederman, 2004; Koren and Tenreyro, 2007; Cadot et al., 2011) and the Product-Gini as relative measure of diversification (Imbs and Wacziarg, 2003; Cadot et al., 2011). For a detailed discussion on the different indices the interested reader is referred to Cadot et al. (2013). I do not derive a product-level Theil index. Both the Gini-coefficient and the Theil index are scale independent measures of inequality (see e.g. Allison, 1978; Bourguignon, 1979; Thon, 1982). As the Pareto distribution used in this model is scale invariant, both the Gini and Theil index only depend on the shape parameter of the Pareto distribution and lead to similar results.

a firm, diversification is increasing in firm productivity. Formal firms are more diversified than informal firms as found in CIEM (2012) and product diversification is independent of the product-level regulation. However, the Product-Gini, which considers the product-level revenue distribution of a firm, predicts the opposite. Due to the product-level regulation, registered, and hence formal, firms produce their highest-revenue products in compliance with regulation, while all other products are produced in evasion thereof. Therefore, the majority of firm revenue is concentrated among registered products and formal firms are less diversified than informal firms which do not register any product. Lastly, the product-level Herfindahl index captures both the number of products and the skewed revenue distribution. Accordingly, diversification is increasing in productivity, but controlling for productivity formal firms are less diversified than informal firms. The pattern of diversification along the firm productivity spectrum is particularly interesting in light of the recent work on sectoral and export diversification that also finds a pattern of diversification, albeit along the economic development path (e.g. Imbs and Wacziarg, 2003; Koren and Tenreyro, 2007; Cadot et al., 2011). Moreover, this result shows the importance of considering product diversification not just in terms of number of distinct products, but also through product revenue-based measures, which can lead to different results.

This paper makes therefore four important contributions. First, this is the first model to examine the drivers of informality in a multi-product setting. Second, the model is the first to consider specifically product-level regulation and its impact on informality as well as product diversification in the light of the recent empirical evidence. Third, the model captures informality within the formal sector, that is evasion of product registration by registered firms, that has been empirically observed, yet theoretically neglected in the previous literature. Fourth, in addition to measuring product diversification simply by product scope, the model proposes the use of two revenue-based measures, the Gini coefficient and Herfindahl index that commonly have been employed in empirical studies at the industry-level, at the firm-level. As this paper shows, the two indicators can readily be implemented in this type of multi-product firm model, facilitate bringing theoretical models to the data and provide for richer predictions. Therefore, this parsimonious setup is the first model to jointly consider firm heterogeneity, multiple products and informality in one framework and provides empirically testable predictions that shed light on the sector and diversification choices of firms in the presence of both product- and firm-level regulation.

Through these contributions, the paper has clear policy implications. First, while the complexity of firm-level regulations have taken the primary focus in the discourse on curbing informality (see e.g. World Bank (2013)), the results of this paper show that product-level regulations should also be part of the discussion. Moreover, as formal firms are predicted to be less diversified than informal firms, the often targeted formalization of economies could entail a loss in aggregate product variety and hence welfare. Therefore, the implications of informality and diversification for aggregate welfare deserve further scrutiny in future work.

I proceed with a description of the model in section 2, and section 3 concludes the paper.

2 The model

Consider an economy with an aggregate consumer that supplies L units of labor (section 2.1). A continuum of firms, heterogeneous in productivity, produce a range of goods and decide to comply with or evade regulations at the firm- and product-level (section 2.2).

2.1 Demand: The aggregate consumer

The aggregate consumer has a utility function described by the following CES-preferences over a continuum of identical products a in the interval $[0, 1]$:

$$U = \left[\int_0^1 C_a^{\iota} \mathrm{d}a\right]^{\frac{1}{\iota}}, 0 < \iota < 1, \tag{1}$$

where $\kappa \equiv \frac{1}{1-\iota}$ is the elasticity of substitution between products. The consumer derives utility from the consumption of differentiated varieties of each of the products. Therefore, the consumption of product a consists of a consumption index C_a of individual varieties v. The consumption index C_a is described by

$$C_a = \left[\int_{v \in V} q_a(v)^{\rho} \mathrm{d}v\right]^{\frac{1}{\rho}}, 0 < \rho < 1, \tag{2}$$

where $q_a(v)$ is the consumption of variety v of product a and $\rho \equiv \frac{\sigma-1}{\sigma}$. σ is the elasticity of substitution between product varieties. I assume $\sigma > \kappa$, that is a higher substitutability between product varieties than between products themselves. The price index for each product a depends on the prices of the individual varieties $p_a(v)$ and follows as

$$P_a = \left[\int_{v \in V} p_a(v)^{1-\sigma} \mathrm{d}v\right]^{\frac{1}{1-\sigma}}. \tag{3}$$

The wage rate w serves as the numeraire. Given the focus on firm informality, labor market impacts of regulation and informality lie outside of the scope of this paper and therefore are not explicitly considered.[6] The aggregate consumer maximizes utility subject to the constraint that the aggregate expenditure R_a over the continuum of products is equal to aggregate labor income described by L with w normalized to 1:

$$L = \int_0^1 C_a P_a \mathrm{d}a = \int_0^1 R_a \mathrm{d}a. \tag{4}$$

As common in a monopolistic competition setup, the utility-maximizing $q_a(v)$, i.e. the demand for variety v of product a, depends on the aggregate expenditure for the product R_a, the product's price index P_a, the price of the variety $p_a(v)$ and the elasticity of substitution σ:

[6] The interested reader is referred to the model in Becker (2014) that jointly considers heterogeneous firms, informality and labor markets.

$$q_a(v) = R_a P_a^{\sigma-1} p_a(v)^{-\sigma}. \tag{5}$$

2.2 Supply: Formal and informal firms

In a manner well known from Melitz (2003), firms are initially identical. Upon entering the market, however, firms do not just draw a firm-specific productivity φ, but also a product-specific skill β_a for each product of the continuum of products. This allows each firm to produce a unique variety v of each of the products. However, firms with the same firm productivity φ behave in the same manner and are accordingly henceforth only indexed by it. A firm's marginal cost of producing a product is decreasing in both its firm-specific productivity φ and the product-specific skill β_a. Firm output $q_a(\varphi, \beta_a)$ of good a is linear in labor input l_a for the product, that is $q_a(\varphi, \beta_a) = \varphi\beta_a l_a$.

Firms decide on their compliance with regulations, and hence their formality, on two levels.[7] First, firms are legally required to register, which incurs firm-level fixed cost F_f. Alternatively, firms can choose to evade registration by only paying F_i and become informal at the firm-level. I assume $F_i < F_f$ to capture the findings of the empirical literature that firm registration, particularly in the developing country context, is an arduous procedure and costly (Schneider and Enste, 2000; Djankov et al., 2002; Auriol and Warlters, 2005; Antunes and de V. Cavalcanti, 2007). Henceforth, all informal sector variables feature subscript i and formal sector variables subscript f.

Second, firms are legally required to register each of their products at product-level fixed cost f_f. As a benefit of compliance with product-level regulation, and hence formality at the product-level, firms experience a productivity bonus $\lambda \in [0, 1)$ for the registered product and firm productivity for the product becomes $\frac{\varphi}{1-\lambda}$. The productivity bonus can be seen as the result of protection by the rule of law for that product (de Soto, 1989; La Porta and Shleifer, 2008; Dabla-Norris et al., 2008). Alternatively, firms can evade product-level regulation by paying f_i. In the context of Ghana's TMPs, "[the firms] who did not register any of their products gave reasons such as the cumbersome nature of the registration procedures and the fees charged being too high." (Essegbey et al. (2014, p. 2)) To capture this observation, I assume $f_i < f_f$. However, evading product-level registration entails the probability of government enforcement $\delta \in (0, 1)$ and loss of product revenue upon detection of that specific product,[8] a reflection of the institutional quality of the economy (Loayza, 1996; Dabla-Norris et al., 2008). All firm- and product-level fixed costs are measured in labor units.

Formal and informal firms maximize their profit by charging a constant markup $(1/\rho)$ over the product-specific marginal cost for each of their products. Considering the aforementioned costs and benefits, the prices for products produced in evasion of and in compliance with product regulation are described by:[9]

[7] Appendix E develops an extension with export regulations and provides for additional interesting results.

[8] Intuitively, if a government agent discovers an unregistered product on the market, the specific product and the associated revenue will be confiscated.

[9] The following product-level prices and revenues are the same for formal and informal firms and only differ between registered and unregistered products.

$$p_{ai}(\varphi, \beta_a) = ((1-\delta)^{\sigma} \rho\varphi\beta_a)^{-1} \text{ and } p_{af}(\varphi, \beta_a) = \left(\rho \frac{\varphi}{1-\lambda}\beta_a\right)^{-1}. \tag{6}$$

Given consumer demand for the products of each firm described by (5), the revenue for product a produced in evasion of or compliance with product regulation by a firm with productivity φ and product skill β_a is

$$r_{ai}(\varphi, \beta_a) = (1-\delta)^{\sigma} R_a \left(\rho P_a \varphi \beta_a\right)^{\sigma-1} \text{ and } r_{af}(\varphi, \beta_a) = R_a \left(\rho P_a \frac{\varphi}{1-\lambda}\beta_a\right)^{\sigma-1}. \tag{7}$$

Lastly, taking the product-level fixed costs of into account, the product-level profits for product a differ according to firm productivity φ, product-specific skill β_a and whether a product is registered or not. They are described by

$$\pi_{ai}(\varphi, \beta_a) = (1-\delta)^{\sigma} \frac{R_a}{\sigma} \left(\rho P_a \varphi \beta_a\right)^{\sigma-1} - f_i \text{ and } \pi_{af}(\varphi, \beta_a) = \frac{R_a}{\sigma} \left(\rho P_a \frac{\varphi}{1-\lambda}\beta_a\right)^{\sigma-1} - f_f. \tag{8}$$

The firm-specific productivity φ is assumed to be distributed Pareto with $G(\varphi) = 1 - \varphi^{-\alpha}$ and $g(\varphi) = \alpha\varphi^{-\alpha-1}$, due to the fit of the Pareto distribution to the empirically observed firm productivity distribution (Axtell, 2001; Helpman et al., 2004). Additionally, I follow Bernard et al. (2011) in assuming that the product-specific skill is also distributed Pareto with $Z(\beta) = 1 - \beta^{-k}$ and $z(\beta) = k\beta^{-k-1}$. The lower bound of both distributions is normalized to 1. Lastly, I assume $\alpha > k > \sigma - 1$ to ensure a finite mean firm size.

I proceed by solving the model in two steps. First, I solve the firms' choice of products, taking their decision on joining the formal or informal sector as given (section 2.2.1). Second, I solve for the firms' choice of sector given their firm productivity φ (section 2.2.2). Lastly, I examine the impact of product-regulation on diversification as measured by product scope (section 2.2.3), Product-Gini (section 2.2.4) and Herfindahl index (section 2.2.5), and conclude by comparing the three indicators (sections 2.2.6).

2.2.1 Firms' product choice

In the model, corresponding to the Ghanaian TMP sector where some registered firms evade product-level regulation (Essegbey et al., 2014), registered firms may evade regulation for their products, and hence produce them informally. Unregistered, and hence informal, firms evade regulations for all their products. With the same production technology and demand structure for all identical products, the pricing of a product and accordingly the product-level profit depends solely on firm productivity φ, product-specific skill β and whether a product is registered or not. Accordingly, I henceforth omit the product-specific subscript a.

For informal firms there exists a product skill threshold level $\beta_i^*(\varphi)$ at which a firm with productivity φ just breaks even in producing an unregistered product. The reason for this is the product-level fixed cost f_i that limits the profitability of an unregistered product $\pi_i(\varphi, \beta)$ described in (8). $\beta_i^*(\varphi)$ is defined by

$$\pi_i(\varphi, \beta_i^*(\varphi)) = 0. \tag{9}$$

Any product for which the firm draws a higher product skill than $\beta_i^*(\varphi)$ can be profitably produced by the firm. Any product with a lower product skill is unprofitable and hence will not be produced. Notably, $\beta_i^*(\varphi)$ depends on the firm-specific productivity φ. A high firm productivity compensates for a low product-specific skill and accordingly high-productivity firms are able to produce even low product skill products profitably, that is $\beta_i^*(\varphi)$ is decreasing in φ.[10]

Formal firms, on the other hand, may produce goods in compliance with or evasion of product-level regulations. Intuitively, as product registration at f_f is more costly than informal production at f_i, firms only register products if the registration provides them for a higher profitability than informal production. Accordingly, products in which a firm draws a low product-specific skill will not be registered and high skill products will be registered. Therefore, the lower product-skill threshold $\beta_i^*(\varphi)$ is defined as for informal firms by (9). In addition, there exists a threshold level $\beta_f^*(\varphi)$ that determines the minimum product skill threshold for registered products defined by

$$\pi_i(\varphi, \beta_f^*(\varphi)) = \pi_f(\varphi, \beta_f^*(\varphi)) \tag{10}$$

above which registration of a product is more profitable than unregistered production. In sum, if a formal firm draws a product-specific skill above $\beta_f^*(\varphi)$, it will produce the product in compliance with the product regulation. A skill draw below $\beta_f^*(\varphi)$, but above $\beta_i^*(\varphi)$ leads to the evasion of product regulation for that product, and a draw below $\beta_i^*(\varphi)$ means that the firm will not produce the product since the production would incur negative profits. The sorting $\beta_i^*(\varphi) < \beta_f^*(\varphi)$ is ensured if $\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} < \frac{f_f}{f_i}$, which I henceforth assume holds.[11] The condition follows economic intuition. Only if the benefit of formal production relative to the potential government enforcement of informality is smaller than the cost of registering a good relative to producing an unregistered good, a higher product skill is required to benefit from registration of a good compared to unregistered production.[12]

[10]Mathematically $\frac{\partial \beta_i^*(\varphi)}{\partial \varphi} = -\left[\frac{f_i \sigma}{(1-\delta)^{\sigma} R}\right]^{\frac{1}{\sigma-1}} (\rho P)^{-1} \varphi^{-2} < 0$.

[11]The condition follows from the assumption that, given the same firm productivity, producing goods in evasion of product regulation is profitable at a lower product skill level than producing goods in compliance of that regulation. This requires that $\beta_i^* < \beta_f^*$ for β_i^* from $\pi_i(\beta_i^*, \varphi) = 0$ & β_f^* from $\pi_f(\beta_f^*, \varphi) = 0$. $\beta_i^* = \left[\frac{f_i \sigma}{(1-\delta)^{\sigma} R}\right]^{\frac{1}{\sigma-1}} (\rho P \varphi)^{-1}$ and $\beta_f^* = \left[\frac{F_f \sigma}{R}\right]^{\frac{1}{\sigma-1}} \left(\rho P \frac{\varphi}{1-\lambda}\right)^{-1}$. Hence, $\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} < \frac{f_f}{f_i}$. As both $\pi_i(\varphi, \beta)$ and $\pi_f(\varphi, \beta)$ are monotonically increasing in β, single-crossing of the two functions is ensured and $\beta_i^*(\varphi) < \beta_f^*(\varphi)$ holds.

[12]In this model, firms produce goods of the same quality at different prices. It is conceivable that unregistered, and hence informal products, are of lower quality. An alternative interpretation of Melitz (2003)-type models is that firms produce products of different quality at the same cost. Given the alternative interpretation, this model implicitly captures a quality and hence demand difference between registered and unregistered products. Heterogeneous firms and heterogeneous product quality are explicitly modeled in Verhoogen (2008), albeit in a setup of heterogeneous workers.

Since the product skill draws of firms over the continuum of identical products are i.i.d. and given the law of large numbers, the expected profit of a firm over the continuum of products is equal to the expected profit for an individual product, which equals the probability of drawing a product-specific skill above the threshold levels. Accordingly, total firm profit of an informal firm that does not comply with firm-level regulation depends on firm productivity φ and is described by

$$\pi_i(\varphi) = \int_{\beta_i^*(\varphi)}^{\infty} \left[(1-\delta)^{\sigma} \frac{R}{\sigma} (\rho P \varphi \beta)^{\sigma-1} - f_i\right] z(\beta) \mathrm{d}\beta - F_i. \tag{11}$$

Formal firms, on the other hand, can produce products both in evasion of and/or compliance with product-level regulation, and pay a firm-level registration fee F_f:

$$\begin{aligned} \pi_f(\varphi) = & \int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} \left[(1-\delta)^{\sigma} \frac{R}{\sigma} (\rho P \varphi \beta)^{\sigma-1} - f_i\right] z(\beta) \mathrm{d}\beta \\ & + \int_{\beta_f^*(\varphi)}^{\infty} \left[\frac{R}{\sigma} \left(\rho P \frac{\varphi}{1-\lambda} \beta\right)^{\sigma-1} - f_f\right] z(\beta) \mathrm{d}\beta - F_f. \end{aligned} \tag{12}$$

The share of registered products that a formal firm produces is equal to the probability of drawing a product skill above the threshold $\beta_f^*(\varphi)$, i.e. $1 - Z(\beta_f^*(\varphi)) = \beta_f^*(\varphi)^{-k}$. The share of unregistered products is described by the probability of drawing a product-specific skill above the informal product-skill threshold $\beta_i^*(\varphi)$, but below the formal threshold $\beta_f^*(\varphi)$, i.e. $Z(\beta_f^*(\varphi)) - Z(\beta_i^*(\varphi)) = \beta_i^*(\varphi)^{-k} - \beta_f^*(\varphi)^{-k}$. Using (9) and (10), I can write the share of unregistered products relative to the share of registered products as

$$\frac{\beta_i^*(\varphi)^{-k} - \beta_f^*(\varphi)^{-k}}{\beta_f^*(\varphi)^{-k}} = \left[\frac{f_i}{f_f - f_i}\right]^{\frac{-k}{\sigma-1}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-k}{\sigma-1}} - 1. \tag{13}$$

The right hand side of (13) is independent of φ. Therefore, while the share of the whole continuum of products that is produced by a firm depends on its firm productivity φ, the relative share of unregistered to registered products is the same for all formal firms and independent of φ. A comparative statics exercise on (13) provides for intuitive results. Parameters that increase the relative profitability of production in compliance with product regulation (decrease in f_f or increase in f_i, λ, δ) lead to a higher relative share of registered products.

Proposition 1. *The share of unregistered relative to registered products of a formal firm is increasing in f_f and decreasing in f_i, λ and δ.*

Proof. See appendix A. □

2.2.2 Firms' sector choice

Having solved the product choice of firms, as second step I solve firms' sector choice. Given their firm-level productivity φ, firms choose to either register their firm and become formal, evade firm-level registration and become informal or not produce at all according to the profitability of the activity (e.g. de Mel et al., 2013; La Porta and Shleifer, 2014), that is $\max\{\pi_i(\varphi), \pi_f(\varphi), 0\}$. This poses the question how informal and formal firms are distributed over the productivity spectrum. The empirical literature provides an answer to this question: informal firms are lower-productivity firms and formal firms are higher-productivity firms (La Porta and Shleifer, 2008; de Paula and Scheinkman, 2011). Analogously to the product skill threshold levels, the firm-level fixed costs lead to the firm-level productivity threshold levels φ_i^* and φ_f^*, which are the productivity levels above which firms select into informal and formal production, respectively. The thresholds are determined with the help of (11) and (12) by

$$\pi_i(\varphi_i^*) = 0 \tag{14}$$

and

$$\pi_i(\varphi_f^*) = \pi_f(\varphi_f^*). \tag{15}$$

If a firm draws a firm-level productivity below φ_i^*, it will not produce at all. For a draw above φ_i^*, but below φ_f^*, the firm becomes informal, and for a draw above φ_f^* the firm becomes formal. The sorting $\varphi_i^* < \varphi_f^*$ is ensured if $1 < \left[\frac{(1-\lambda)^{1-\sigma}-(1-\delta)^{\sigma}}{(1-\delta)^{\sigma}}\right]^{\frac{k}{\sigma-1}} < \frac{F_f}{F_i}\left[\frac{f_f-f_i}{f_i}\right]^{\frac{k+1-\sigma}{\sigma-1}}$, which I henceforth assume holds.[13] Intuitively, this condition compares the benefit of formality relative to informality to the costs of formality relative to informality. There are two parts to the comparison. First, the benefit of formality relative to informality $\left[\frac{(1-\lambda)^{1-\sigma}-(1-\delta)^{\sigma}}{(1-\delta)^{\sigma}}\right]^{\frac{k}{\sigma-1}}$ has to be greater than 1 for formality to be attractive for any firm. Second, if the formality relative to informality costs ratio $\frac{F_f}{F_i}\left[\frac{f_f-f_i}{f_i}\right]^{\frac{k+1-\sigma}{\sigma-1}}$ is greater than the benefit ratio, then only high-productivity firms will find it beneficial to become formal.

Similar to the share of registered and unregistered products, due to an i.i.d. firm productivity

[13] The condition results from two assumptions. First, I assume that informal sector firms break even at a lower productivity level than formal firms, which requires that $\varphi_i^* < \varphi_f^*$ for φ_i^* from $\pi_i(\varphi_i^*) = 0$ & φ_f^* from $\pi_f(\varphi_f^*) = 0$. $\varphi_i^* = F_i^{\frac{1}{k}}\left[\frac{R(1-\delta)^{\sigma}}{\sigma}\right]^{\frac{1}{\sigma-1}}(\rho P)^{-1} f_i^{\frac{k+1-\sigma}{(\sigma-1)k}}\left[\frac{\sigma-1}{k+1-\sigma}\right]^{\frac{-1}{k}}$ and $\varphi_f^* = F_f^{\frac{1}{k}}\left[\frac{R(1-\delta)^{\sigma}}{\sigma}\right]^{\frac{1}{\sigma-1}}(\rho P)^{-1} f_i^{\frac{k+1-\sigma}{(\sigma-1)k}}\left[\frac{\sigma-1}{k+1-\sigma}\right]^{\frac{-1}{k}}\left[\frac{f_i}{f_f-f_i}\right]^{\frac{k+1-\sigma}{(1-\sigma)k}}\left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}}-1\right]^{\frac{k}{\sigma-1}}$. Hence, $\left[\frac{(1-\lambda)^{1-\sigma}-(1-\delta)^{\sigma}}{(1-\delta)^{\sigma}}\right]^{\frac{k}{\sigma-1}} < \frac{F_f}{F_i}\left[\frac{f_i}{f_f-f_i}\right]^{\frac{\sigma-1-k}{\sigma-1}}$. Second, I assume that the marginal profitability with respect to productivity is higher in the formal sector, which requires $\frac{\partial\pi_i}{\partial\varphi} < \frac{\partial\pi_f}{\partial\varphi}$. $\frac{\partial\pi_i}{\partial\varphi} = k\varphi^{k-1}\left[\frac{R(1-\delta)^{\sigma}}{\sigma}\right]^{\frac{-k}{1-\sigma}}(\rho P)\, f_i^{\frac{\sigma-1-k}{\sigma-1}}\left[\frac{\sigma-1}{k+1-\sigma}\right]$ and $\frac{\partial\pi_f}{\partial\varphi} = k\varphi^{k-1}\left[\frac{R(1-\delta)^{\sigma}}{\sigma}\right]^{\frac{-k}{1-\sigma}}(\rho P)\, f_i^{\frac{\sigma-1-k}{\sigma-1}}\left[\frac{\sigma-1}{k+1-\sigma}\right]\left[\frac{f_i}{f_f-f_i}\right]^{\frac{k+1-\sigma}{\sigma-1}}\left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}}-1\right]^{\frac{k}{\sigma-1}}$.
Hence, $1 < \left[\frac{(1-\lambda)^{1-\sigma}-(1-\delta)^{\sigma}}{(1-\delta)^{\sigma}}\right]^{\frac{k}{\sigma-1}}$. The two conditions jointly are $1 < \left[\frac{(1-\lambda)^{1-\sigma}-(1-\delta)^{\sigma}}{(1-\delta)^{\sigma}}\right]^{\frac{k}{\sigma-1}} < \frac{F_f}{F_i}\left[\frac{f_i}{f_f-f_i}\right]^{\frac{\sigma-1-k}{\sigma-1}}$. As both $\pi_i(\varphi)$ and $\pi_f(\varphi)$ are monotonically increasing in φ, single-crossing of the two functions is ensured and $\varphi_i^* < \varphi_f^*$ holds.

distribution and the law of large numbers, the share of formal firms is equal to the probability of a productivity draw above the formality threshold φ_f^*, i.e. $1-G(\varphi_f^*) = \varphi_f^{*-\alpha}$, and the share of informal firms equals the probability of drawing a productivity between φ_i^* and φ_f^*, i.e. $Z(\varphi_f^*) - Z(\varphi_i^*) = \varphi_i^{*-\alpha} - \varphi_f^{*-\alpha}$. Using (14) and (15) that determine the productivity cutoff levels, I derive the share of informal firms relative to formal firms:

$$\frac{\varphi_i^{*-\alpha} - \varphi_f^{*-\alpha}}{\varphi_f^{*-\alpha}} = \left[\frac{F_i}{F_f - F_i}\right]^{-\frac{\alpha}{k}} \left[\frac{f_i}{f_f - f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}} - 1. \tag{16}$$

Intuitively, factors that make formal sector participation relatively more costly or decrease its benefits (increase in F_f, f_f, decrease in F_i, f_i, λ, δ) lead to a relatively larger informal sector. Importantly, when considering both types of regulation, not just firm-level, but also product-level regulation f_f steers informal sector participation.[14] This provides the rational for the findings of Loayza et al. (2005), who show product market regulations to be drivers of informality in a cross-country context.

Proposition 2. *The share of informal relative to formal firms is increasing in F_f, f_f and decreasing in F_i, f_i, λ and δ.*

Proof. See appendix B. □

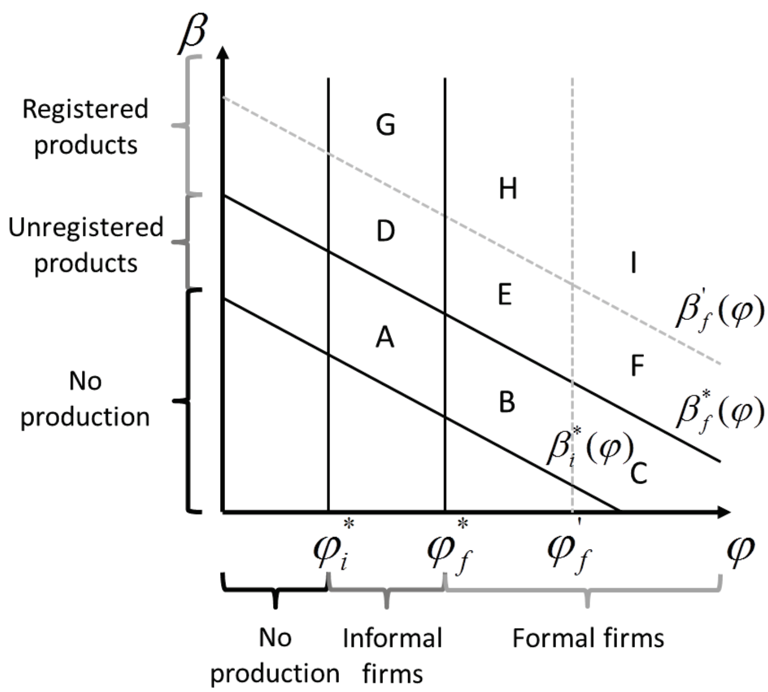


Figure 1: Sorting of firms and products along the productivity and product skill dimensions.

[14]This is particularly striking in contrast to a setup with only firm-level regulations in place as derived in appendix F or in Becker (2014).

The results of Propositions 1 and 2 can be graphically illustrated to highlight the effect of product regulation on both the share of informal firms and unregistered products in the market. Figure 1 pictures all products of an economy within the firm productivity and product skill dimensions. The curves $\beta_i^*(\varphi)$ and $\beta_f^*(\varphi)$ are the product skill threshold levels following from (9) and (10):

$$\beta_i^*(\varphi) = \left[\frac{f_i\sigma}{(1-\delta)^\sigma R}\right]^{\frac{1}{\sigma-1}} (\rho P\varphi)^{-1} \tag{17}$$

and

$$\beta_f^*(\varphi) = \left[\frac{(f_f - f_i)\,\sigma}{R}\right]^{\frac{1}{\sigma-1}} (\rho P\varphi)^{-1} \left[(1-\lambda)^{1-\sigma} - (1-\delta)^\sigma\right]^{\frac{1}{1-\sigma}}. \tag{18}$$

As higher productivity firms are able to produce products at a lower product-specific skill, the product skill thresholds $\beta_i^*(\varphi)$ and $\beta_f^*(\varphi)$ are decreasing in firm productivity.[15] The firm productivity thresholds φ_i^* and φ_f^*, however, are independent of product skill. They follow from the combination of their definitions (14) and (15) with the product skill cutoff thresholds (17) and (18):

$$\varphi_i^* = F_i^{\frac{1}{k}} \left[\frac{k+1-\sigma}{\sigma-1}\right]^{\frac{1}{k}} f_i^{\frac{\sigma-k-1}{(1-\sigma)k}} \left[\frac{R(1-\delta)^\sigma}{\sigma}\right]^{\frac{1}{1-\sigma}} (\rho P)^{-1} \tag{19}$$

and

$$\varphi_f^* = [F_f - F_i]^{\frac{1}{k}} \left[\frac{k+1-\sigma}{\sigma-1}\right]^{\frac{1}{k}} [f_f - f_i]^{\frac{\sigma-k-1}{(1-\sigma)k}} \left[\frac{R(1-\delta)^\sigma}{\sigma}\right]^{\frac{1}{1-\sigma}} (\rho P)^{-1} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^\sigma} - 1\right]^{\frac{1}{\sigma-1}}. \tag{20}$$

Firms with firm productivity above φ_i^* and below φ_f^* evade firm registration. Firms with productivity level above φ_f^* register their firm, thus being formal firms. Similarly, products for which a formal firm draws a product skill above $\beta_i^*(\varphi)$ but below $\beta_f^*(\varphi)$ are produced in evasion of product regulation, and products with product skill above $\beta_f^*(\varphi)$ are registered products. Unregistered products accordingly take up space ABCDG. However, the effect of a change in firm-level regulation F_f differs to that of a change in product-level regulation f_f. More costly firm-level regulation, i.e. an increase in F_f, ceteris paribus affects firms' decision to register their business and shifts the productivity threshold level from φ_f^* to $\varphi_f^{'}$ ($\frac{\partial\varphi_f^*}{\partial F_f} > 0$ from (20)). Yet, the decision on the compliance with product regulation is not affected. When deciding on product registration, the firm level regulation cost F_f is already sunk and firms optimize only at the product-level, which mathematically means that (18) does not depend on F_f.[16] The result is an increase in the share of informal firms, and accordingly unregistered products. Unregistered products take up space ABCDEGH and informalization at the product-level has increased. In contrast to this, an increase in product-level regulation cost f_f makes not just product registration more costly and thus shifts

[15] Mathematically, $\frac{\partial\beta_i^*(\varphi)}{\partial\varphi} = -\left[\frac{f_i\sigma}{(1-\delta)^\sigma R}\right]^{\frac{1}{\sigma-1}} (\rho P)^{-1}\varphi^{-2} < 0$ and
$\frac{\partial\beta_f^*(\varphi)}{\partial\varphi} = -\left[\frac{(f_f-f_i)\sigma}{R}\right]^{\frac{1}{\sigma-1}} \left[(1-\lambda)^{1-\sigma} - (1-\delta)^\sigma\right]^{\frac{1}{1-\sigma}} (\rho P)^{-1}\varphi^{-2} < 0.$

[16] $\beta_i^*(\varphi)$ and φ_i^* do not depend on F_f and are unaffected.

the product registration threshold from $\beta_f^*(\varphi)$ to $\beta_f^{'}(\varphi)$ ($\frac{\partial \beta_i^*(\varphi)}{\partial f_f} > 0$ from (17)), but also makes formal sector participation less attractive ($\frac{\partial \varphi_i^*}{\partial f_f} > 0$ from (19)) and shifts the productivity threshold from φ_f^* to $\varphi_f^{'}$.[17] As a result, both the share of informal firms and unregistered products in the economy rise to take up space ABCDEFGH, while the number of registered products decreases to only take up space I. By having an effect on both the share of informal firms and unregistered products, product-level regulation has a stronger effect than firm-level regulation on the extent of informality at the product-level.

2.2.3 Product Scope

A firm's product scope $K(\varphi)$ is defined as the number of distinct products a firm produces. Because product skill draws are i.i.d. over the continuum of products, $K(\varphi)$ is equal to the probability of drawing a product skill above the product skill threshold $\beta_i^*(\varphi)$, i.e. $1 - Z(\beta_i^*(\varphi))$. Both informal and formal firms do not register their lowest product skill goods and the relevant product skill threshold is in both cases $\beta_i^*(\varphi)$. Hence, the product-level regulation does not affect the product scope of firms. The product scope of both type of firms are

$$K_i(\varphi) = 1 - Z(\beta_i^*(\varphi)) = \beta_i^*(\varphi)^{-k} \text{ and } K_f(\varphi) = 1 - Z(\beta_i^*(\varphi)) = \beta_i^*(\varphi)^{-k}. \tag{21}$$

As the threshold is decreasing in firm productivity, product scope is increasing in firm productivity.[18] Therefore, higher productivity firms are more diversified measured by the product scope, which corresponds to the theoretical and empirical findings of Bernard et al. (2010) and Bernard et al. (2011).

Proposition 3. *Product scope is increasing in firm productivity.*

2.2.4 Product-Gini

Originally a measure of inequality, the Gini coefficient is commonly employed as a measure of diversification at the industry-level, for instance in the context of sectoral and export diversification (e.g. Imbs and Wacziarg, 2003; Cadot et al., 2011). I employ the Gini coefficient at the product-level to examine the product revenue distribution and hence the product diversification of a firm. The derivation of the Product-Gini for informal firms requires two steps. First, I derive γ, i.e. the share of products that are produced with a product skill below $\bar{\beta}$ of the total number of products a firm produces:

$$\gamma = \frac{\int_{\beta_i^*(\varphi)}^{\bar{\beta}} \beta z(\beta) \mathrm{d}\beta}{\int_{\beta_i^*(\varphi)}^{\infty} \beta z(\beta) \mathrm{d}\beta} = 1 - \left(\frac{\bar{\beta}}{\beta_i^*(\varphi)} \right)^{-k}. \tag{22}$$

[17] $\beta_i^*(\varphi)$ and φ_i^* do not depend on f_f and are unaffected.

[18] Substituting for $\beta_i^*(\varphi)$ leads to $K_i(\varphi) = \left[\frac{f_i \sigma}{(1-\delta)^\sigma R}\right]^{\frac{k}{1-\sigma}} (\rho P \varphi)^k$ and accordingly $\frac{\partial K_i(\varphi)}{\partial \varphi} > 0$.

Next, I derive the share of product revenue of products with product skill below $\bar{\beta}$ relative to total firm revenue Q_i as

$$Q_i = \frac{\int_{\beta_i^*(\varphi)}^{\bar{\beta}} r_i(\varphi,\beta)z(\beta)\mathrm{d}\beta}{\int_{\beta_i^*(\varphi)}^{\infty} r_i(\varphi,\beta)z(\beta)\mathrm{d}\beta} = 1 - \left(\frac{\bar{\beta}}{\beta_i^*(\varphi)}\right)^{\sigma-k-1}. \tag{23}$$

Lastly, substituting (22) into (23) leads to the Lorenz curve $Q_i(\gamma)$:

$$Q_i(\gamma) = 1 - [1-\gamma]^{\frac{\sigma-k-1}{-k}}, \tag{24}$$

which has the desired properties $Q_i(0) = 0$, $Q_i(1) = 1$ and $\frac{\partial Q_i(\gamma)}{\partial \gamma} > 0$. The Product-Gini of informal firms G_i derives from

$$G_i = 1 - 2\int_0^1 Q_i(\gamma)\mathrm{d}\gamma = \frac{\sigma-1}{2k+1-\sigma}. \tag{25}$$

Given its relative nature, G_i is independent of firm productivity φ, and hence the total number of products a firm produces. The product-level Gini only captures the distribution of the product revenues relative to the whole firm revenue. Firms that are neither registered on a firm- nor on a product-level are subject to government enforcement δ for all products and thus enforcement does not affect the relative distribution. Further, the distribution of product revenue follows the product skill distribution and accordingly G_i only depends on the product skill distribution parameter k and the elasticity of substitution between the products of different firms σ.[19]

The derivation of the Product-Gini of formal firms G_f is more complicated as formal firms register only some of their products. The individual steps necessary to derive G_f can be found in appendix C.

$$\begin{aligned} G_f = G_i \Bigg\{ \frac{2}{\Omega(\sigma-1)} \Bigg[\frac{2k+1-\sigma}{2} \left[\Omega - 2\phi^k \left[(1-\delta)^{\sigma-1} + \Omega \right] \right] \\ -\phi^{1+2k-\sigma} k \left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma-1} \right] - (1-\delta)^{\sigma-1}(\sigma-1-k) \Bigg] \Bigg\}, \end{aligned} \tag{26}$$

where $\phi \equiv (\beta_i^*(\varphi)/\beta_f^*(\varphi))$ and $\Omega \equiv \left[(1-\delta)^\sigma + \phi^{k+1-\sigma} \left[(1-\lambda)^{1-\sigma} - (1-\delta)^\sigma \right] \right]$. Besides the Pareto distribution parameter k and the elasticity of substitution between products σ, G_f depends on ϕ and Ω that are proxies for the relative share of unregistered to registered products produced by a formal firm. Firms that are formal at the firm-level do not register a share of their goods under the probability of government enforcement δ and produce their registered products with the productivity bonus λ. Given this, unregistered products generate a relatively lower revenue per

[19] Informal firms only differ in their threshold product skill level $\beta_i^*(\varphi)$ and hence the lower bound of their product revenue distribution. A truncated Pareto is still Pareto, albeit with a lower minimum bound and mean described by $\beta_i^*(\varphi)\,[k/(k-1)]$. As the Gini-coefficient is scale independent (see e.g. Allison, 1978; Thon, 1982), only the Pareto distribution parameter matters and all firms that are informal at the firm-level are characterized by the same product-level Gini. Notably, the same G_i would also hold if there were no product-level regulations in place and a firm would produce all products. See appendix F.2 for a derivation of this result.

product and registered products generate a relatively higher revenue. Thus, the Lorenz curve has two segments.

In a model where firms register all their products, as for instance in Bernard et al. (2010), firms' product diversification would be described by $G_i(\varphi)$. However, as formal firms may comply with or evade product regulation, their Lorenz curve is kinked and product diversification is described by $G_f(\varphi)$. The product-level regulation f_f has two effects. First, it influences firms' choice of firm-level informality or formality, and thus whether diversification of the firm is measured by G_i or G_f. Second, f_f enters ϕ, determines the relative share of unregistered products and hence the absolute value of G_f.

Proposition 4. *The Product-Gini is independent of firm productivity conditional on a firms' sector choice.*

2.2.5 Herfindahl Index

The Herfindahl index is an absolute measure of diversification that captures both the distribution of a firm's revenue among its products and the number of distinct products. To measure the product diversification of a firm I derive it in two steps. For informal firms I first derive the revenue share $s_i(\beta)$ of an unregistered product that is produced with product skill β of total firm revenue:

$$s_i(\beta) = \frac{r_i(\varphi, \beta)}{\int_{\beta_i^*(\varphi)}^{\infty} r_i(\varphi, \beta) z(\beta) \mathrm{d}\beta} = \frac{k+1-\sigma}{k} \beta_i^*(\varphi)^{1+k-\sigma} \beta^{\sigma-1}. \tag{27}$$

Second, the product-level Herfindahl index for informal firms $H_i(\varphi)$ is computed as the integral over the squared revenue shares for all produced products:

$$H_i(\varphi) = \int_{\beta_i^*(\varphi)}^{\infty} s_i(\beta)^2 z(\beta) \mathrm{d}\beta = H \beta_i^*(\varphi)^k, \tag{28}$$

where $H = (k+1-\sigma)^2 / [k(k+2-2\sigma)] > 1$.[20] Similar to the Product-Gini, the Herfindahl index depends on the product skill distribution parameter k and the elasticity of substitution between products σ, both captured by H.[21] Further, the product-level Herfindahl index, similar to the product scope, depends on firm productivity φ. Higher productivity firms are characterized by a lower $\beta_i^*(\varphi)$ and produce more distinct product. Hence, firm revenue is distributed over a wider range of products and the Herfindahl index indicates these firms as more diversified.[22]

The product-level Herfindahl index of formal firms $H_f(\varphi)$ consists of two pieces. First, the product revenue share of unregistered products and second the product revenue share of registered

[20]Rewriting $H > 1$ leads to $(\sigma^2 - 1)(\sigma - 1) > 0$, which holds because $\sigma > 1$.

[21]Without any product-level regulations in place, $\beta_i^*(\varphi) = 1$ and the product-level Herfindahl reduces to H. This result is derived in appendix F.3.

[22]Substituting for $\beta_i^*(\varphi)$ leads to $H_i(\varphi) = H \left[\frac{f_i \sigma}{(1-\delta)^{\sigma} R}\right]^{\frac{k}{\sigma-2}} (\rho P \varphi)^{-k}$ and accordingly $\frac{\partial H_i(\varphi)}{\partial \varphi} < 0$.

products. The derivation is more arduous than for informal firms and the interested reader is therefore referred to appendix D.

$$H_f(\varphi) = H\beta_i^*(\varphi)^k\Omega^{-2}\omega, \tag{29}$$

where $\omega \equiv \left[(1-\delta)^{2\sigma} + \phi^{k+2-2\sigma}\left[(1-\lambda)^{2-2\sigma} - (1-\delta)^{2\sigma}\right]\right]$ and $\Omega^{-2}\omega > 1$.[23] Notably, $H_f(\varphi)$ consists of three parts. First, H reflects the product revenue distribution as a result of the product skill distribution within a firm and the elasticity of substitution between product varieties. Second, $\beta_i^*(\varphi)^k$ captures the increasing number of products a higher productivity firm can produce due to a lower product skill threshold.[24] A larger number of products reduces the Herfindahl index.[25] Lastly, $\Omega^{-2}\omega$ reflects the skew of the revenue distribution, because the share of registered products generates relatively higher product revenue than the share of unregistered products. The less even distribution of revenue among all products entails less diversification and increases the Herfindahl index.

Because there exists informality at the product-level, some firms diversification is described by $G_f(\varphi)$. Without the possibility of the evasion of product-level regulation, as for instance in Bernard et al. (2010), diversification of all firms would be described by $H_i(\varphi)$. Therefore, similar to the Product-Gini, also the Herfindahl index is affected by the product-level regulation f_f in two ways. First, the product-level regulation affects firms' choice of sector and accordingly whether diversification of the firm is measured by $H_i(\varphi)$ or $H_f(\varphi)$. Second, f_f enters $\Omega^{-2}\omega$, which captures the extent to which the revenue distribution is skewed through informal and formal production of goods, and thus the absolute value of $H_f(\varphi)$.

Proposition 5. *The product-level Herfindahl index is decreasing in firm productivity conditional on a firms' sector choice.*

2.2.6 Comparison of the indicators

The diversification of an informal and a formal firm, measured by the relative product scope, $K_i(\varphi_i)/K_f(\varphi_f)$, can be compared using (21):

$$\frac{K_i(\varphi_i)}{K_f(\varphi_f)} = \left(\frac{\varphi_i}{\varphi_f}\right)^k < 1. \tag{30}$$

The difference in product scope only depends on the relative productivity difference between the two firms. Because informal sector firms are characterized by a lower productivity than formal sector firms (cf. (16)), they produce a smaller number of distinct products and are less diversified as measured by product scope. This corresponds to the empirical finding in CIEM (2012).

[23] $\Omega^{-2}\omega > 1$ holds if $\Omega^2 - \omega < 0$. Rewriting leads to $\phi^{1-\sigma}\left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right]\left[\phi^k - 1\right] + 2(1-\delta)^{\sigma}\left[1-\phi^{1-\sigma}\right] < 0$, which holds because $\phi \in (0,1)$, $1-\sigma < 0$ and $(1-\lambda)^{1-\sigma} > (1-\delta)^{\sigma}$.

[24] Mathematically $\frac{\partial \beta_i^*(\varphi)}{\partial \varphi} < 0$.

[25] Substituting for $\beta_i^*(\varphi)$ leads to $H_f(\varphi) = H\left[\frac{f_i\sigma}{(1-\delta)^{\sigma}R}\right]^{\frac{k}{\sigma-2}}(\rho P\varphi)^{-k}\Omega^{-2}\omega$ and accordingly $\frac{\partial H_f(\varphi)}{\partial \varphi} < 0$.

Contrary to that, (25) and (26) in combination show that unregistered firms are more diversified than registered firms measured by the Product-Gini:

$$\frac{G_i}{G_f} = \left[\frac{2}{\Omega(\sigma-1)}\right]^{-1} \left[\frac{2k+1-\sigma}{2}\left[\Omega - 2(1-\delta)^{\sigma}\left[1-\phi^k\right] - 2\phi^k\Omega\right]\right.$$
$$\left. + k\phi^{2k+1-\sigma}\left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right] + k(1-\delta)^{\sigma}\right]^{-1} < 1. \tag{31}$$

The product-level Gini captures the concentration of revenue of formal firms among registered products through $\phi \in (0,1)$. Informal firms produce only unregistered products. Therefore, formal firms are less diversified than informal firms and $G_i/G_f < 1$.[26] Notably, this result is independent of firm productivity and holds when comparing any informal to any formal firm. Despite that the product-level regulation f_f affects the absolute value of G_f through ϕ, $G_f > G_i$ holds in any case.

Lastly, using (28) and (29) leads to the relative difference in diversification of firms in both sectors measured by the Herfindahl index, $H_i(\varphi_i)/H_f(\varphi_f)$, which depends on firms' productivity difference, formality status and economic parameters:

$$\frac{H_i(\varphi_i)}{H_f(\varphi_f)} = \Omega^2\omega^{-1}\left(\frac{\varphi_f}{\varphi_i}\right)^k. \tag{32}$$

The difference in product diversification between informal and formal firms is driven by two components. First, higher-productivity firms are able to produce a wider range of distinct products given product-level regulation and hence the relative difference in diversification of two firms depends on their relative productivities. Second, formal firms produce both registered and unregistered products. As most of formal firms' profit is generated by registered goods, the product revenue distribution of formal firms is skewed. This is captured by $\Omega^2\omega^{-1} < 1$ that decreases $H_i(\varphi_i)/H_f(\varphi_f)$ and indicates that, controlling for firm productivity φ, formal firms are less diversified than informal firms. Similar to the case of the Product-Gini, while the product-level regulation affects $H_f(\varphi)$ through $\Omega^2\omega^{-1}$, formal firms are always less diversified than informal firms given the same firm productivity φ.[27]

Proposition 6. *As measured by product scope, formal firms are more diversified than informal firms. By contrast, formal firms are less diversified than informal firms measured by the Product-Gini and, when controlling for productivity, by the product-level Herfindahl index.*

Figure 2 is a stylized visualization of the three findings summarized in Proposition 6 of product diversification over the productivity spectrum. Product scope, as measured by the number of

[26] $\left[\frac{2}{\Omega(\sigma-1)}\right]^{-1}\left[\frac{2k+1-\sigma}{2}\left[\Omega - 2(1-\delta)^{\sigma}\left[1-\phi^k\right] - 2\phi^k\Omega\right] + k\phi^{2k+1-\sigma}\left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right] + k(1-\delta)^{\sigma}\right]^{-1} < 1$ can be rewritten as $(k+1-\sigma)\left[\left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right]\left[\phi^{k+1-\sigma} - \phi^{2k+1-\sigma}\right]\right] > 0$. Because $(1-\lambda)^{1-\sigma} > (1-\delta)^{\sigma}$, $k+1-\sigma > 0$ and $\phi \in (0,1)$, it must be that $G_i/G_f < 1$.

[27] These diversification results are particularly interesting in comparison to the variant of the model developed in appendix F in which firms do not face a product-level regulation and as a result all firms are equally diversified as measured by any of the three indicators.

distinct products, is increasing in firm productivity over the whole spectrum and formal firms are more diversified than informal firms. Measured by the Product-Gini, formal firms are less diversified than informal firms. However, the diversification only depends on formality status and not firm productivity. Lastly, diversification measured by the Herfindahl index combines features of the two other indicators. On the one hand, diversification is increasing in firm productivity. On the other hand, there is a diversification gap between informal and formal firms, where formal firms are indicated as less diversified given their skewed product revenue distribution.

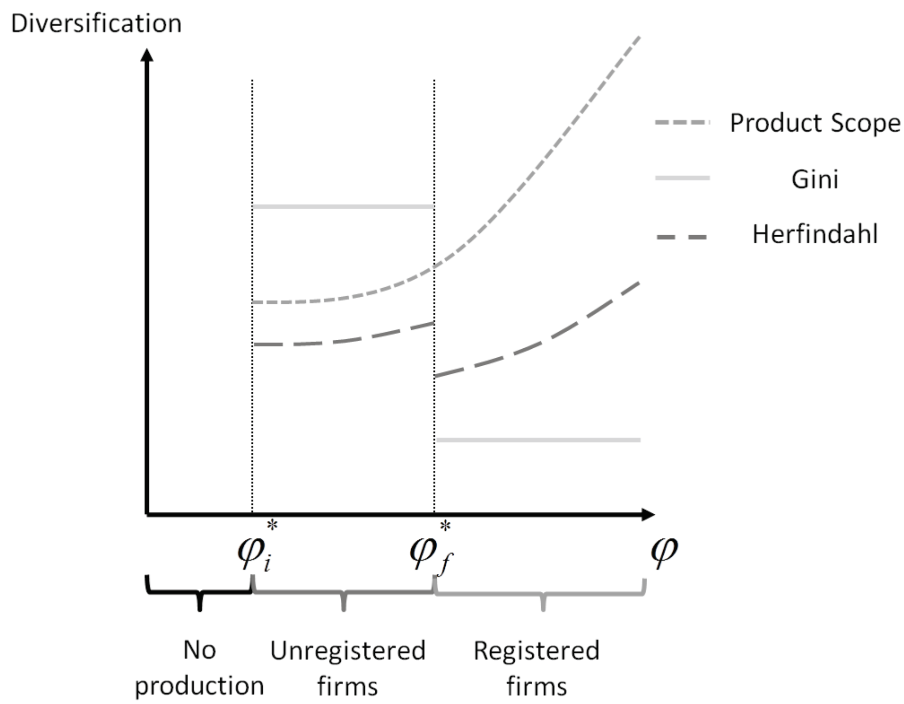


Figure 2: Product diversification along the productivity spectrum measured by the three indicators.

3 Conclusion

In this paper I build a model of heterogeneous multi-product firms in the presence of both product- and firm-level regulations. I capture a multi-dimensional concept of informality and explain firms' sector choice and diversification decision in the presence of these two type of regulations. The results indicate that the product-level regulation is an important driver of informality at both firm- and product-level. Further, the model highlights that in addition to product scope, also product revenue-based measures are important to understand diversification. More specifically, the model shows that utilizing product scope leads to the opposite prediction than using a Product-Gini or Herfindahl index, that is informal firms are indicated to be less or more diversified than formal firms.

The present work offers two avenues for fruitful future research. First, considering both product-and firm-level regulations in a general equilibrium setup would provide a richer set of predictions. Increasing product-level regulation might not only lead to an adjustment of the number of products by a given a firm, but also have general equilibrium impacts that further affect the demand for all products. This extension would be particularly interesting with regard to the impacts of trade on informality in a multi-product setting. If foreign market access is reserved for formal firms and registered products, resource reallocation is likely. Similarly, extending the model to explicitly feature labor market frictions along the lines of Becker (2014) would provide interesting insights into the impact of product-level regulation on employment, particularly in general equilibrium. Resource reallocation jointly with labor market frictions could potentially lead to greater inequality between workers employed in heterogeneous firms. Therefore, future research should carefully consider the far-reaching effects of product-level regulation, which my model shows to be a new and important driver of informality and determinants of product diversification.

Appendix

A Share of unregistered relative to registered products

$$\left(\frac{\beta_i^*(\varphi)}{\beta_f^*(\varphi)}\right)^{-k} - 1 = \left[\frac{f_i}{f_f - f_i}\right]^{\frac{-k}{\sigma-1}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-k}{\sigma-1}} - 1.$$

$$\frac{\partial\left(\frac{\beta_i^*(\varphi)}{\beta_f^*(\varphi)}\right)^{-k}}{\partial f_i} = \left[\frac{-k}{\sigma-1}\right] \left[\frac{f_f}{(f_f-f_i)^2}\right] \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-k}{\sigma-1}-1} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-k}{\sigma-1}} < 0.$$

$$\frac{\partial\left(\frac{\beta_i^*(\varphi)}{\beta_f^*(\varphi)}\right)^{-k}}{\partial f_f} = \left[\frac{-k}{\sigma-1}\right] \left[\frac{-f_i}{(f_f-f_i)^2}\right] \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-k}{\sigma-1}-1} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-k}{\sigma-1}} > 0.$$

$$\frac{\partial\left(\frac{\beta_i^*(\varphi)}{\beta_f^*(\varphi)}\right)^{-k}}{\partial \lambda} = -k \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-k}{\sigma-1}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-k}{\sigma-1}-1} (1-\lambda)^{-\sigma}(1-\delta)^{-\sigma} < 0.$$

$$\frac{\partial\left(\frac{\beta_i^*(\varphi)}{\beta_f^*(\varphi)}\right)^{-k}}{\partial \delta} = \left[\frac{-k\sigma}{\sigma-1}\right] \left[\frac{f_i}{f_f-f_i}\right]^{\frac{1}{\sigma-1}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{1}{\sigma-1}-1} (1-\lambda)^{1-\sigma}(1-\delta)^{-\sigma-1} < 0.$$

B Share of informal relative to formal firms

$$\left(\frac{\varphi_i^*}{\varphi_f^*}\right)^{-\alpha} - 1 = \left[\frac{F_i}{F_f-F_i}\right]^{-\frac{\alpha}{k}} \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}} - 1.$$

$$\frac{\partial\left(\frac{\varphi_i^*}{\varphi_f^*}\right)^{-\alpha}}{\partial F_i} = -\frac{\alpha}{k} \left[\frac{F_f}{(F_f-F_i)^2}\right] \left[\frac{F_i}{F_f-F_i}\right]^{-\frac{\alpha}{k}-1} \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}} < 0.$$

$$\frac{\partial\left(\frac{\varphi_i^*}{\varphi_f^*}\right)^{-\alpha}}{\partial F_f} = \frac{\alpha}{k} \left[\frac{F_i}{(F_f-F_i)^2}\right] \left[\frac{F_i}{F_f-F_i}\right]^{-\frac{\alpha}{k}-1} \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}} > 0.$$

$$\frac{\partial\left(\frac{\varphi_i^*}{\varphi_f^*}\right)^{-\alpha}}{\partial f_i} = \left[\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}\right] \left[\frac{f_f}{(f_f-f_i)^2}\right] \left[\frac{F_i}{F_f-F_i}\right]^{-\frac{\alpha}{k}} \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}-1} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}} < 0.$$

$$\frac{\partial\left(\frac{\varphi_i^*}{\varphi_f^*}\right)^{-\alpha}}{\partial f_f} = \left[\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}\right] \left[\frac{f_i}{(f_f-f_i)^2}\right] \left[\frac{F_i}{F_f-F_i}\right]^{-\frac{\alpha}{k}} \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}-1} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}} > 0.$$

$$\frac{\partial\left(\frac{\varphi_i^*}{\varphi_f^*}\right)^{-\alpha}}{\partial \lambda} = -\alpha \left[\frac{F_i}{F_f-F_i}\right]^{-\frac{\alpha}{k}} \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}-1} (1-\lambda)^{-\sigma}(1-\delta)^{-\sigma} < 0.$$

$$\frac{\partial\left(\frac{\varphi_i^*}{\varphi_f^*}\right)^{-\alpha}}{\partial \delta} = \left[\frac{-\alpha}{\sigma-1}\right] \left[\frac{F_i}{F_f-F_i}\right]^{-\frac{\alpha}{k}} \left[\frac{f_i}{f_f-f_i}\right]^{\frac{-\alpha(k+1-\sigma)}{(\sigma-1)k}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}-1} (1-\lambda)^{1-\sigma}(1-\delta)^{-\sigma-1} < 0.$$

C Product-Gini of formal firms

The first segment of the Lorenz curve consists of unregistered products. Therefore, I first derive the number of products produced with a product skill below $\bar{\beta}$:

$$\gamma = \frac{\int_{\beta_i^*(\varphi)}^{\bar{\beta}} \beta z(\beta) \mathrm{d}\beta}{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} \beta z(\beta) \mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\infty} \beta z(\beta) \mathrm{d}\beta} = 1 - \left(\frac{\bar{\beta}}{\beta_i^*(\varphi)} \right)^{-k}. \tag{33}$$

The next step is calculating the share in revenue of unregistered products that are produced with a product skill below $\bar{\beta}$:

$$Q_1 = \frac{\int_{\beta_i^*(\varphi)}^{\bar{\beta}} r_i(\varphi, \beta) z(\beta) \mathrm{d}\beta}{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} r_i(\varphi, \beta) z(\beta) \mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\infty} r_f(\varphi, \beta) z(\beta) \mathrm{d}\beta} = \frac{(1-\delta)^{\sigma-1}}{\Omega} \left[\left(\frac{\bar{\beta}}{\beta_i^*(\varphi)} \right)^{\sigma-k-1} - 1 \right], \tag{34}$$

where $\phi \equiv (\beta_i^*(\varphi)/\beta_f^*(\varphi))$ and $\Omega \equiv \left[(1-\delta)^{\sigma} + \phi^{k+1-\sigma}\left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right]\right]$. Combining (33) and (34) leads to the first segment of the Lorenz curve $Q_1(\gamma)$:

$$Q_1(\gamma) = \frac{(1-\delta)^{\sigma-1}}{\Omega} \left[[1-\gamma]^{\frac{k+1-\sigma}{k}} - 1 \right]. \tag{35}$$

The second segment of the Lorenz curve consists of registered products. The number of products produced with product skill below $\bar{\beta}$ are

$$\gamma = \frac{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} \beta z(\beta) \mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\bar{\beta}} \beta z(\beta) \mathrm{d}\beta}{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} \beta z(\beta) \mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\infty} \beta z(\beta) \mathrm{d}\beta} = 1 - \left(\frac{\bar{\beta}}{\beta_i^*(\varphi)} \right)^{-k}. \tag{36}$$

Subsequently, I calculate the share in revenue of products produced with product skill below $\bar{\beta}$:

$$Q_2 = \frac{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} r_i(\varphi, \beta) z(\beta) \mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\bar{\beta}} r_f(\varphi, \beta) z(\beta) \mathrm{d}\beta}{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} r_i(\varphi, \beta) z(\beta) \mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\infty} r_f(\varphi, \beta) z(\beta) \mathrm{d}\beta} = 1 + \frac{(1-\lambda)^{1-\sigma}}{\Omega} \left(\frac{\bar{\beta}}{\beta_i^*(\varphi)} \right)^{\sigma-k-1}. \tag{37}$$

Combining (36) and (37), leads to the second segment of the Lorenz curve $Q_2(\gamma)$:

$$Q_2(\gamma) = 1 + \frac{(1-\lambda)^{1-\sigma}}{\Omega} [1-\gamma]^{\frac{k+1\sigma}{k}}. \tag{38}$$

Jointly, the Lorenz curve is described by

$$Q(\gamma) = \begin{cases} Q_1(\gamma) \text{ if } \gamma \in [0, b) \\ Q_2(\gamma) \text{ if } \gamma \in [b, 1] \end{cases}, \tag{39}$$

where $b = 1 - \phi^k$ is the share of unregistered products produced by a firm. The Lorenz curve $Q(\gamma)$

has the desired properties $Q(0) = 0$, $Q(1) = 1$, $Q_1(b) = Q_2(b)$ and $\frac{\partial Q(\gamma)}{\partial \gamma} > 0$. The Gini-coefficient follows from $G_f = 1 - 2\int_0^1 Q(\gamma)\mathrm{d}\gamma$:

$$G_f = G_i \left\{ \frac{2}{\Omega(\sigma-1)} \left[\frac{2k+1-\sigma}{2} \left[\Omega - 2\phi^k \left[(1-\delta)^{\sigma-1} + \Omega \right] \right] \right. \right.$$
$$\left. \left. -\phi^{1+2k-\sigma} k \left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma-1} \right] - (1-\delta)^{\sigma-1}(\sigma-1-k) \right] \right\}.$$

D Herfindahl index of formal firms

The revenue share of an unregistered product that is produced with product skill β of total firm revenue is

$$s_1(\beta) = \frac{r_i(\varphi,\beta)}{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} r_i(\varphi,\beta)z(\beta)\mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\infty} r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta} =$$
$$(1-\delta)^{\sigma} \frac{k+1-\sigma}{k} \beta_i^*(\varphi)^{1+k-\sigma}\Omega^{-1}\beta^{\sigma-1}. \tag{40}$$

Moreover, the revenue share of a registered product that is produced with product skill β of total firm revenue is

$$s_2(\beta) = \frac{r_f(\varphi,\beta)}{\int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} r_i(\varphi,\beta)z(\beta)\mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\infty} r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta} =$$
$$(1-\lambda)^{1-\sigma} \frac{k+1-\sigma}{k} \beta_i^*(\varphi)^{1+k-\sigma}\Omega^{-1}\beta^{\sigma-1}. \tag{41}$$

The Herfindahl index results from the combination of the two:

$$H_f = \int_{\beta_i^*(\varphi)}^{\beta_f^*(\varphi)} s_1(\beta)^2 z(\beta)\mathrm{d}\beta + \int_{\beta_f^*(\varphi)}^{\infty} s_2(\beta)^2 z(\beta)\mathrm{d}\beta = H\beta_i^*(\varphi)^k\Omega^{-2}\omega,$$

where $\omega \equiv (1-\delta)^{2\sigma} + \phi^{k+2-2\sigma}\left[(1-\lambda)^{2-2\sigma} - (1-\delta)^{2\sigma}\right]$.

E International trade

In this section I consider an extension of the model featuring international trade. Consider a world of $n+1$ symmetric countries. Besides informality and formality, firms can choose to export a share of their products to n countries. Export activities require the firm-level fixed costs $F_x > F_f$ to reflect the findings of Roberts and Tybout (1997) that fixed costs are critical determinants of export participation. Intuitively, firms need not just register their firm, but face additional costs and time requirements to comply with firm-level export regulations. Given the mounting empirical evidence on the exclusion of informal firms from international trade (Batra et al., 2003; Bigsten et al., 2004; La Porta and Shleifer, 2008), I proceed with the assumption that the registration for exporting entails an intense government screening that leads to the probability of government enforcement $\delta = 1$ and accordingly renders informal production economically infeasible. To model the product registration requirement and learning about export markets, I assume that firms incur a fixed cost $f_x > f_f$ for each exported product at the product-level. The transportation of products to foreign markets is costly and modeled as variable iceberg cost $\tau > 1$, that is for one good to arrive at the destination country, τ goods have to be shipped. The product-level profits for exported goods are therefore described by

$$\pi_x(\varphi, \beta) = n\tau^{1-\sigma}\frac{R}{\sigma}\left(\rho P\frac{\varphi}{1-\lambda}\beta\right)^{\sigma-1} - f_x. \tag{42}$$

As firms have different product skills for the individual products they produce, this poses the question which products will be exported. The empirical literature provides an answer to this question (e.g. Iacovone and Javorcik, 2008; Bernard et al., 2011): core products, i.e. the products the firm has the highest product skill in, are sold both at home and abroad. Peripheral products, i.e. products that the firm has a low product skill in, are sold only at home. Mathematically, if $f_f < \frac{f_x\tau^{\sigma-1}}{n}$, then firms export high-product skill products and sell low-product skill products only domestically.[28] The condition corresponds to the firm-level condition for export participation in Melitz (2003) and Bernard et al. (2011).[29]

The firm-level profits and product skill threshold levels for informal and formal firms and important assumptions are as described in section 2.2. Exporting firms, however, cannot produce goods informally. Therefore, the product skill threshold levels for domestic and exported products $\beta_t^*(\varphi)$ and $\beta_x^*(\varphi)$, respectively, of exporting firms are determined by

$$\pi_f(\varphi, \beta_t^*(\varphi)) = 0 \tag{43}$$

[28] I derive $\beta_t^*(\varphi)$ from $\pi_f(\varphi, \beta_t^*(\varphi) = 0$ and $\beta_x^*(\varphi)$ from $\pi_x(\varphi, \beta_x^*(\varphi) = 0$. Accordingly, $\beta_t^*(\varphi) = \left[\frac{f_f\sigma}{R}\right]^{\frac{1}{\sigma-1}}(\rho P\frac{\varphi}{1-\lambda})^{-1}$ and $\beta_x^*(\varphi) = \left[\frac{f_x\sigma}{Rn\tau^{1-\sigma}}\right]^{\frac{1}{\sigma-1}}(\rho P\frac{\varphi}{1-\lambda})^{-1}$. To ensure $\beta_t^*(\varphi) < \beta_x^*(\varphi)$, $f_f < \frac{f_x\tau^{\sigma-1}}{n}$ has to hold.

[29] Notably, in this setup f_x is a one-time payment independent of the number of export markets n to reflect stringent product-regulation that occurs once for a product independent of the number of destination markets.

and

$$\pi_x(\varphi, \beta_x^*(\varphi)) = 0. \tag{44}$$

Notably, $\beta_t^*(\varphi)$ differs from $\beta_f^*(\varphi)$ in that $\beta_t^*(\varphi)$ is the threshold at which formal production is economically feasible, whereas $\beta_f^*(\varphi)$ defines the threshold at which formal production is more profitable than informal production of a product. Total profit of exporting firms therefore is

$$\begin{aligned}\pi_x(\varphi) = &\int_{\beta_t^*(\varphi)}^{\infty} \left[\frac{R}{\sigma}\left(\rho P \frac{\varphi}{1-\lambda}\beta\right)^{\sigma-1} - f_f\right] z(\beta)\mathrm{d}\beta - F_f \\ &+ \int_{\beta_x^*(\varphi)}^{\infty} \left[\frac{n\tau^{1-\sigma}R}{\sigma}\left(\rho P \frac{\varphi}{1-\lambda}\beta\right)^{\sigma-1} - f_x\right] z(\beta)\mathrm{d}\beta - F_x. \end{aligned} \tag{45}$$

Firms now have four choices and maximize profit according to $\max\{0, \pi_i(\varphi), \pi_f(\varphi), \pi_x(\varphi)\}$. That means they can either seize production, remain fully informal, register their firm, but evade product regulation for a share of their products, or register their firm for domestic and export activities and register their entire catalog of products. The empirical literature on trade and firm productivity finds that high-productivity firms export and low-productivity firms only serve the domestic market (e.g. Bernard and Jensen, 1995; Roberts and Tybout, 1997; Delgado et al., 2002; Wagner, 2007). To capture this, the threshold level φ_x^* above which firms decide to become exporters is defined by

$$\pi_f(\varphi_x^*) = \pi_x(\varphi_x^*). \tag{46}$$

In sum, if a firm draws a firm-level productivity below φ_i^*, it will not produce at all. For a draw above φ_i^*, but below φ_f^*, the firm becomes informal. For a draw above φ_f^*, but below φ_x^* the firm becomes formal. Lastly, a draw above φ_x^* leads the firm to become an exporter. The sorting $\varphi_f^* < \varphi_x^*$ is ensured if

$$\begin{aligned}&\left[\frac{F_f - F_i}{F_x}\right]^{\frac{1}{k}} \left[\frac{(f_f - f_i)^{\frac{\sigma-k-1}{1-\sigma}}}{(1-\lambda)^k}\left[f_f^{\frac{\sigma-k-1}{\sigma-1}} + (n\tau^{1-\sigma})^{\frac{k}{\sigma-1}} f_x^{\frac{\sigma-k-1}{\sigma-1}})\right] \Xi^{\frac{k}{1-\sigma}}\right. \\ &\left. + \frac{k}{(1-\sigma)\Xi}\left[1 + (1-\delta)^{\sigma}\left[\phi^{\sigma-k-1} - 1\right]\right] + \frac{1-\sigma}{(f_f - f_i)(\sigma-k-1)}\left[f_f + f_i\left[\phi^{-k} - 1\right]\right]\right]^{\frac{1}{k}} < 1,\end{aligned}$$

with $\Xi \equiv \left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right]$. The condition follows from the combination of the formality cutoff condition described by (15) and (46).

The share of the exported products of the continuum of products is described by $1 - Z(\beta_x^*(\varphi)) = \beta_x^*(\varphi)^{-k}$. Similarly, the share of domestically sold products is determined by $1 - Z(\beta_t^*(\varphi)) = \beta_t^*(\varphi)^{-k}$. Therefore, I can express the share of exported relative to domestically sold products as

$$\left(\frac{\beta_x^*}{\beta_t^*}\right)^{-k} = \left[\frac{f_x}{f_f}\right]^{\frac{-k}{\sigma-1}} n^{\frac{k}{\sigma-1}} \tau^{-k}. \tag{47}$$

The results of a comparative statics analysis on (47) are intuitive. Factors that increase the profitability of exporting (decrease in f_x, τ or increase in n) or increase the cost of domestic sales (increase in f_f) increase the relative share of exported products.

E.1 Product Scope

The number of distinct products produced by an exporting firm is determined by the probability of drawing a product skill above the firm-specific product skill threshold of production. This follows from

$$K_x(\varphi) = 1 - Z(\beta_t^*(\varphi)) = \beta_t^*(\varphi)^{-k}. \tag{48}$$

E.2 Product-Gini

The Lorenz curve that captures the product revenue distribution of exporters has two distinct segments. The first segment considers only domestically sold products and the second steeper segment captures the revenue generated from products that are sold both at home and abroad. I first derive the share of products that are produced with a product skill below $\bar{\beta}$ of the total number of products a firm produces

$$\gamma = \frac{\int_{\beta_t^*(\varphi)}^{\bar{\beta}} \beta z(\beta) \mathrm{d}\beta}{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} \beta z(\beta) \mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\infty} \beta z(\beta) \mathrm{d}\beta} = 1 - \left(\frac{\bar{\beta}}{\beta_t^*(\varphi)}\right)^{-k} \tag{49}$$

The next step is calculating the revenue share of domestically sold products that are produced with product skill below $\bar{\beta}$ of total firm revenue:

$$Q_1 = \frac{\int_{\beta_t^*(\varphi)}^{\bar{\beta}} r_f(\varphi, \beta) z(\beta) \mathrm{d}\beta}{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} r_f(\varphi, \beta) z(\beta) \mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\infty} (1 + n\tau^{1-\sigma}) r_f(\varphi, \beta) z(\beta) \mathrm{d}\beta} = \frac{\left[\left(\frac{\bar{\beta}}{\beta_f^*(\varphi)}\right)^{\sigma-k-1} - 1\right]}{\Psi}, \tag{50}$$

where $\xi = (\beta_t^*(\varphi)/\beta_x^*(\varphi))$ and $\Psi \equiv 1 + n\tau^{1-\sigma}\xi^{k+1-\sigma}$. The first segment of the Lorenz curve is

$$Q_1(\gamma) = \frac{\left[[1-\gamma]^{\frac{k+1-\sigma}{k}} - 1\right]}{\Psi}. \tag{51}$$

The second segment consists of the goods that are also exported. The share of products that are produced with a product skill below $\bar{\beta}$ of all products is

$$\gamma = \frac{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} \beta z(\beta) \mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\bar{\beta}} \beta z(\beta) \mathrm{d}\beta}{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} \beta z(\beta) \mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\infty} \beta z(\beta) \mathrm{d}\beta} = 1 - \left(\frac{\bar{\beta}}{\beta_t^*(\varphi)}\right)^{-k}. \tag{52}$$

The share in revenue of the products that are also exported and produced with a product skill below $\bar{\beta}$ is

$$Q_2 = \frac{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} r_i(\varphi,\beta)z(\beta)\mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\bar{\beta}} r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta}{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} r_i(\varphi,\beta)z(\beta)\mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\infty}(1+n\tau^{1-\sigma})r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta} = 1 + \frac{\left(\frac{\bar{\beta}}{\beta_t^*(\varphi)}\right)^{\sigma-k-1}}{\Psi}. \tag{53}$$

In combination they lead to the second segment of the Lorenz curve:

$$Q_2(\gamma) = 1 + \frac{[1-\gamma]^{\frac{k+1\sigma}{k}}}{\Psi}. \tag{54}$$

Jointly, the Lorenz curve is described by

$$Q(\gamma) = \begin{cases} Q_1(\gamma) \text{ if } \gamma \in [0,c) \\ Q_2(\gamma) \text{ if } \gamma \in [c,1] \end{cases}, \tag{55}$$

where $c = 1-\xi^k$ is the share of domestically sold goods relative to the share products that are also exported. The Lorenz curve $Q(\gamma)$ has the desired properties $Q(0)=0$, $Q(1)=1$, $Q_1(c)=Q_2(c)$ and $\frac{\partial Q(\gamma)}{\partial \gamma} > 0$. The Gini-coefficient follows from $G_x = 1 - 2\int_0^1 Q(\gamma)\mathrm{d}\gamma$:

$$\begin{aligned} G_x = G\Bigg\{\frac{2}{\Psi(\sigma-1)}\Bigg[&\frac{2k+1-\sigma}{2}\left[\Psi - 2\xi^k\left[\Psi+1\right]\right] \\ &-\xi^{1+2k-\sigma}kn\tau^{1-\sigma} - k\Bigg]\Bigg\}. \end{aligned} \tag{56}$$

G_x depends on ξ, a proxy for the share of exported relative to only domestically sold products, and Ψ that captures the skew of the product revenue distribution.

E.3 Herfindahl index

The Herfindahl index of exporting firms considers the revenue of two types of products: the products only sold domestically and the products that are sold at home and abroad. The second group of products, the ones also exported, generate a higher revenue due to the multiple markets they're sold in.

The revenue share $s_t(\beta)$ of a domestically sold good product with product skill β of total firm revenue is

$$\begin{aligned} s_t(\beta) = &\frac{r_f(\varphi,\beta)}{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\infty}(1+n\tau^{1-\sigma})r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta} = \\ &\frac{k+1-\sigma}{k}\beta_t^*(\varphi)^{1+k-\sigma}\Psi^{-1}\beta^{\sigma-1}. \end{aligned} \tag{57}$$

Moreover, the revenue share $s_x(\beta)$ of a product that is also exported and produced with product skill β of total firm revenue is

$$s_x(\beta) = \frac{(1+n\tau^{1-\sigma})r_f(\varphi,\beta)}{\int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\infty}(1+n\tau^{1-\sigma})r_f(\varphi,\beta)z(\beta)\mathrm{d}\beta} = (1+n\tau^{1-\sigma})\frac{k+1-\sigma}{k}\beta_f^*(\varphi)^{1+k-\sigma}\Psi^{-1}\beta^{\sigma-1}. \tag{58}$$

The Herfindahl index results from the combination of the two:

$$H_x = \int_{\beta_t^*(\varphi)}^{\beta_x^*(\varphi)} s_t(\beta)^2 z(\beta)\mathrm{d}\beta + \int_{\beta_x^*(\varphi)}^{\infty} s_x(\beta)^2 z(\beta)\mathrm{d}\beta = H\beta_t^*(\varphi)^k\Psi^{-2}\psi, \tag{59}$$

where $\psi \equiv \left[1+\xi^{k+2-2\sigma}\left[2n\tau^{1-\sigma}+n^2\tau^{2-2\sigma}\right]\right]$ and $\Psi^{-2}\psi > 1$.[30]

As in the case of firms that produce both unregistered and registered products, the Herfindahl index for firms that export a share of their goods and sell some of their products only domestically consists of three major components. First, H captures the product skill distribution and elasticity of substitution between product varieties. $\beta_t^*(\varphi)^k$ reflects the number of products a firm produces depending on its firm productivity φ. A higher firm productivity leads to a lower threshold, larger number of products produced by a firm and a lower Herfindahl index.[31] Lastly, $\Psi^{-2}\psi$ captures the concentration of firm revenue among a few core products that are sold both at home and abroad and therefore increases the Herfindahl index.

E.4 Comparisons

First, I compare diversification of exporting firms relative to informal and formal firms by comparing product scope of the firms.

$$\frac{K_i(\varphi_i)}{K_x(\varphi_x)} = \left[\frac{f_f}{f_i}\right]^{\frac{k}{\sigma-1}}(1-\delta)^{\frac{\sigma k}{\sigma-1}}(1-\lambda)^k\left(\frac{\varphi_i}{\varphi_x}\right)^k. \tag{60}$$

Exporting firms produce all their goods in compliance with product-level regulations. Therefore, the difference in product scope between informal and exporting firms depends besides their relative productivity difference also on the costs and benefits of informal and formal production. The two components work in opposite directions. First, exporting firms are higher-productivity firms and are therefore able to produce goods at a lower product skill level. Therefore, the relative productivity difference decreases the informal relative to the exporting product scope. Second, evading product-level regulation allows firms to produce goods at a lower product skill and hence

[30] $\Psi^{-2}\psi > 1$ holds if $\Psi^2 - \psi < 0$. Rewriting leads to
$n^2\tau^{2-2\sigma}\xi^{k+2-2\sigma}\left[\xi^k - 1\right] + 2n\tau^{1-\sigma}\xi^{k+1-\sigma}\left[1-\xi^{1-\sigma}\right] < 0$, which holds because $\xi \in (0,1)$ and $1-\sigma < 0$.

[31] Substituting for $\beta_t^*(\varphi)$ leads to $H_x(\varphi) = H\left[\frac{f_f\sigma}{R}\right]^{\frac{k}{\sigma-2}}(\rho P\frac{\varphi}{1-\lambda})^{-k}\Psi^{-2}\psi$ and accordingly $\frac{\partial H_x(\varphi)}{\partial\varphi} < 0$.

increases the informal relative to the formal product scope.[32]

$$\frac{K_f(\varphi_f)}{K_x(\varphi_x)} = \left[\frac{f_f}{f_i}\right]^{\frac{k}{\sigma-1}} (1-\delta)^{\frac{\sigma k}{\sigma-1}} (1-\lambda)^k \left(\frac{\varphi_f}{\varphi_x}\right)^k. \tag{61}$$

As formal firms can evade product registration, the product scope of formal relative to exporting firms is the same as the previous relationship between informal and exporting firms and consists of the same two components. Comparing the lowest-productivity exporting firm and the highest-productivity formal firm, i.e. $(\varphi_f/\varphi_x) = 1$, shows that export participation entails less diversification, as measured by product scope, for the marginal exporter.

Second, I compare diversification between the different firms using the Product-level Gini.

$$\frac{G_i}{G_x} = \left[\frac{2}{\Psi(\sigma-1)}\right]^{-1} \left[\frac{2k+1-\sigma}{2}\left[\Psi - 2\left[1-\xi^k\right] - 2\xi^k\Psi\right] + kn\tau^{1-\sigma}\xi^{2k+1-\sigma} + k\right]^{-1} < 1. \tag{62}$$

Exporting firms produce all their goods in compliance with product regulation, however, the majority of their revenue stems from core goods that are sold domestically and abroad. ξ captures this as a proxy for the share of products sold domestically relative to the ones that are also exported. The intuition is as follows. When a select group of core products generate the majority of firm revenue, in this case the goods that are also exported, firms are less diversified than when all goods are sold only domestically. Therefore, $G_i/G_x < 1$ and informal firms are more diversified than exporting firms.[33]

$$\frac{G_f}{G_x} = \left[\frac{G_i}{G_f}\right]^{-1} \left[\frac{G_i}{G_x}\right]. \tag{63}$$

Measured by the Product-Gini, formal non-exporting firms may or may not be more diversified than exporting firms. As the Product-Gini is a relative measure, the absolute number of products produced by each firm does not matter and hence firm productivity does not influence this measure of diversification. The two components that matter are the extent to which non-exporting firms' revenue is concentrated among registered products versus the extent to which exporting firms' revenue is concentrated among the core products that are sold domestically and exported. Depending on which of the two components dominates, non-exporting firms are more or less diversified than exporting firms. To demonstrate that G_f/G_x can be smaller or greater than 1, I compute the numerical value of the key variables for two scenarios as shown in table 1. The first scenario considers a world consisting of only home and abroad. The second scenario assumes three foreign countries. All parameters are within the constraints of the model.

Lastly, I compare product diversification of firms using the Herfindahl index.

[32] Given an earlier assumption, $\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} < \frac{f_f}{f_i}$ holds. Rewriting leads to $(f_f/f_i)^{\frac{k}{\sigma-1}}(1-\delta)^{\frac{\sigma k}{\sigma-1}}(1-\lambda)^k > 1$.

[33] $\frac{G_i}{G_x} < 1$ if $\frac{2}{\Psi(\sigma-1)}\left[\frac{2k+1-\sigma}{2}\left[\Psi - 2\left[1-\xi^k\right] - 2\xi^k\Psi\right] + kn\tau^{1-\sigma}\xi^{2k+1-\sigma} + k\right] > 1$.
The condition can be rewritten as $(k+1-\sigma)n\tau^{1-\sigma}\left[\xi^{k+1-\sigma} - \xi^{2k+1-\sigma}\right] > 0$. Because $\xi \in (0,1)$, $k+1-\sigma > 0$ and $k > 1$ the condition holds and the Product-Gini of informal firms is lower than the one of exporting formal firms.

Table 1: Economy parameters for two scenarios.

		Scenario 1	Scenario 2
Total workforce	L	1.0	1.0
Elasticity of substitution	σ	3.4	3.4
Product skill distribution parameter	k	5.0	5.0
Formal fixed cost	F_f	4.0	4.0
Informal fixed cost	F_i	1.0	1.0
Exporting fixed cost	F_x	10.0	10.0
Formal fixed cost	f_f	4.0	4.0
Informal fixed cost	f_i	1.0	1.0
Exporting fixed cost	f_x	10.0	10.0
Productivity bonus	λ	0.3	0.3
Probability of detection	δ	0.1	0.1
Iceberg transportation cost	τ	1.4	1.4
Number of foreign countries	n	**1**	**3**
$\frac{G_f}{G_x}$		**1.36**	**0.94**

$$\frac{H_i(\varphi_i)}{H_x(\varphi_x)} = \left[\frac{f_i}{f_f}\right]^{\frac{k}{\sigma-1}} (1-\delta)^{\frac{-k\sigma}{\sigma-1}}(1-\lambda)^{-k}\Psi^2\psi^{-1}\left(\frac{\varphi_x}{\varphi_i}\right)^k. \tag{64}$$

The diversification of informal relative to exporting firms measured by the Herfindahl index consists of three components. As before, the first component is the relative productivity difference φ_x/φ_i. Second, as informal firms evade product regulation for all and exporting firms for none of their products, the diversification ratio depends also on the relative costs and benefits of evasion of versus compliance with product regulation $[f_i/f_f]^{\frac{k}{\sigma-1}} (1-\delta)^{\frac{-k\sigma}{\sigma-1}}(1-\lambda)^{-k}$. The evasion of product-level regulation allows firms to produce goods at a lower product skill and therefore more overall products. As a result, this component indicates a higher relative diversification of informal firms.[34] Lastly, $\Psi^2\psi^{-1}$ captures the fact that exporting firms sell their core products domestically and abroad. The concentration of firm revenue on a few core products entails a less equal product revenue distribution and accordingly a higher Herfindahl index. In sum, given the same firm productivity, informal firms are more diversified than exporting firms.

$$\frac{H_f(\varphi_f)}{H_x(\varphi_x)} = \left[\frac{f_i}{f_f}\right]^{\frac{k}{\sigma-1}} (1-\delta)^{\frac{-k\sigma}{\sigma-1}}(1-\lambda)^{-k}\Psi^2\psi^{-1}\Omega^{-2}\omega\left(\frac{\varphi_x}{\varphi_f}\right)^k. \tag{65}$$

Lastly, comparing formal non-exporting and exporting firms brings all the previous components together. On the one hand, the ability of non-exporting firms to produce informally, captured by $[f_i/f_f]^{\frac{k}{\sigma-1}} (1-\delta)^{\frac{-k\sigma}{\sigma-1}}(1-\lambda)^{-k}$, and the revenue concentration of exporting firms on core products, $\Psi^2\psi^{-1}$, indicate a higher diversification of non-exporting relative to exporting firms. On the other hand, the revenue concentration on formal products of non-exporting firms, $\Omega^{-2}\omega$, and the productivity difference between exporting and non-exporting firms, $\varphi_x/\varphi_f > 1$, indicates a lower diversification of non-exporting relative to exporting firms. The result is ambiguous.

[34] Given an earlier assumption, $\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} < \frac{f_f}{f_i}$ holds. Rewriting leads to $(f_i/f_f)^{\frac{k}{\sigma-1}}(1-\delta)^{\frac{-\sigma k}{\sigma-1}}(1-\lambda)^{-k} < 1$.

F Firm-level regulation without product-level regulation

This section considers the model setup without product-level regulation, along the lines of Becker (2014). Firms only face firm-level fixed cost F_f, but can evade the registration fee and become informal by paying fixed cost $F_i < F_f$ just to set up shop. Firms aiming to export have to pay $F_x > F_f$, which reflects the additional costs of learning about export markets as well as export regulations. This captures the finding of Roberts and Tybout (1997) that sunk costs are critical determinants of export participation. As there are no boundaries to the profitability of the production of each product, every firm produces all products of the continuum. Notably, variables with subscript x describe export activities and $\pi_x(\varphi)$ is the firm-level profit of exporting that firms receive in addition to their domestic profit. The informal, formal and exporting profit of a firm with productivity φ respectively are

$$\pi_i(\varphi) = \int_1^{\infty} \left[(1-\delta)^{\sigma} \frac{R}{\sigma} (\rho P \varphi \beta)^{\sigma-1}\right] z(\beta) \mathrm{d}\beta - F_i, \tag{66}$$

$$\pi_f(\varphi) = \int_1^{\infty} \left[\frac{R}{\sigma} \left(\rho P \frac{\varphi}{1-\lambda} \beta\right)^{\sigma-1}\right] z(\beta) \mathrm{d}\beta - F_f \tag{67}$$

and

$$\pi_x(\varphi) = \int_1^{\infty} \left[\frac{n\tau^{1-\sigma} R}{\sigma} \left(\rho P \frac{\varphi}{1-\lambda} \beta\right)^{\sigma-1}\right] z(\beta) \mathrm{d}\beta - F_x. \tag{68}$$

The threshold levels φ_i^* and φ_f^* are determined as in section 2.2 by (14) and (15) and informal sector firms are lower-productivity firms than formal sector firms. The sorting $\varphi_i^* < \varphi_f^*$ is ensured if $\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} < \frac{F_f}{F_i}$.[35] Notably, the condition is equivalent to the corresponding one in Becker (2014).

In addition to the cutoff productivity levels defined by (14) and (15), firms are able to export and do so if exporting leads to positive profits. The threshold level φ_x^* above which firms decide to become exporters is defined by

$$\pi_x(\varphi_x^*) = 0. \tag{69}$$

Export activities are conducted by firms in addition to their domestic sales. That means firms maximize their profit by deciding on $\max\{0, \pi_i(\varphi), \pi_f(\varphi), \pi_f(\varphi) + \pi_x(\varphi)\}$. The empirical literature on firm-level productivity and exporting finds that higher-productivity firms export and lower-productivity firms supply only the domestic market (Bernard and Jensen, 1995, 1999; Delgado et al., 2002; Wagner, 2007). This sorting of $\varphi_f^* < \varphi_x^*$ is ensured if $\frac{\left[F_f - F_i\right]^{\frac{1}{\sigma-1}}}{\left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right]^{\frac{1}{\sigma-1}}} < \frac{F_x^{\frac{1}{\sigma-1}} \tau}{n^{\frac{1}{\sigma-1}} (1-\lambda)^{-1}}$.[36]

[35] The condition results from the assumption that informal firms break even at a lower productivity level than formal firms. Breaking even at a lower productivity requires that $\varphi_i^* < \varphi_f^*$ for φ_i^* from $\pi_i(\varphi_i^*) = 0$ & φ_f^* from $\pi_f(\varphi_f^*) = 0$. $\varphi_i^* = \left[\frac{F_i \sigma}{(1-\delta)^{\sigma} R}\right]^{\frac{1}{\sigma-1}} (\rho P) \left[\frac{k+1-\sigma}{k}\right]^{\frac{1}{\sigma-1}}$ and $\varphi_f^* = \left[\frac{F_f \sigma}{R}\right]^{\frac{1}{\sigma-1}} (\rho P)^{\frac{1}{1-\lambda}} \left[\frac{k+1-\sigma}{k}\right]^{\frac{1}{\sigma-1}}$. Hence, $\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} < \frac{F_f}{F_i}$. Since both profit functions are monotonically increasing in productivity, a single crossing of the profit functions is ensured and the productivity sorting is achieved.

[36] Ensuring $\varphi_x^* > \varphi_f^*$ for φ_f^* from $\pi_i(\varphi_f^*) = \pi_f(\varphi_f^*)$ and φ_x^* from $\pi_x(\varphi_x^*) = 0$ is sufficient to sort domestic productivity

Intuitively, only if the cost benefit ratio of formality is lower than the one of exporting, only the highest-productivity firms will export. This condition also corresponds to Becker (2014).

The share of informal and formal firms in an economy is equal to the probability of drawing a productivity above the respective threshold level. Specifically, the share of formal firms is $1 - G(\varphi_f^*) = \varphi_f^{*-k}$ and the share of informal firms follows from $G(\varphi_f^*) - G(\varphi_i^*) = \varphi_i^{*-k} - \varphi_f^{*-k}$. Using (14) and (15), I derive the number of informal relative to formal firms depending only on the policy parameters:

$$\frac{\varphi_i^{*-\alpha} - \varphi_f^{*-\alpha}}{\varphi_f^{*-\alpha}} = \left[\frac{F_i}{F_f - F_i}\right]^{\frac{-\alpha}{\sigma-1}} \left[\frac{(1-\lambda)^{1-\sigma}}{(1-\delta)^{\sigma}} - 1\right]^{\frac{-\alpha}{\sigma-1}} - 1. \tag{70}$$

The effect of changing these parameters is intuitive. For an increase in factors that decrease formal sector profitability relative to the informal sector (increase in F_f or decrease in λ, δ, F_i), the relative share of informal firms‘ increases.

F.1 Product Scope

Since there are is no product-level regulation in place, the minimum product skill threshold equals the lower limit of product skill draws, i.e. 1. Thus, all firms produce all products.[37] Moreover, as all firms produce all products, product scope is equal for informal, formal and exporting firms.

F.2 Product-Gini

First, I derive the number of products with product skill below $\bar{\beta}$ relative to the total amount of products:

$$\gamma = \frac{\int_1^{\bar{\beta}} \beta z(\beta) \mathrm{d}\beta}{\int_1^{\infty} \beta z(\beta) \mathrm{d}\beta} = 1 - \bar{\beta}^{-k}. \tag{71}$$

Next, the revenue share of products with product skill below $\bar{\beta}$ of total firm revenue is

$$Q = \frac{\int_1^{\bar{\beta}} r(\varphi, \beta) z(\beta) \mathrm{d}\beta}{\int_1^{\infty} r(\varphi, \beta) z(\beta) \mathrm{d}\beta} = 1 - \bar{\beta}^{\sigma-k-1}. \tag{72}$$

Lastly, both equations can be combined to calculate the Lorenz curve

$$Q(\gamma) = 1 - [1-\gamma]^{\frac{\sigma-k-1}{-k}} \tag{73}$$

and the Product-Gini

levels below export productivity levels. This results in
$\varphi_f^* = (F_f - F_i)^{\frac{1}{\sigma-1}} \left[\frac{R}{\sigma}\right]^{\frac{1}{1-\sigma}} \left[\frac{(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}}{(1-\delta)^{\sigma}}\right]^{\frac{1}{1-\sigma}} \left[\frac{k+1-\sigma}{k}\right]^{\frac{1}{\sigma-1}} (\rho P)^{-1}$ and
$\varphi_x^* = F_x^{\frac{1}{\sigma-1}} \left[\frac{R}{\sigma}\right]^{\frac{1}{1-\sigma}} n^{\frac{1}{1-\sigma}} \tau (1-\lambda) \left[\frac{k+1-\sigma}{k}\right]^{\frac{1}{\sigma-1}} (\rho P)^{-1}$. In combination, $\frac{\left[F_f - F_i\right]^{\frac{1}{\sigma-1}}}{\left[(1-\lambda)^{1-\sigma} - (1-\delta)^{\sigma}\right]^{\frac{1}{\sigma-1}}} < \frac{F_x^{\frac{1}{\sigma-1}} \tau}{n^{\frac{1}{\sigma-1}} (1-\lambda)^{-1}}$.

[37] Mathematically: $K(\varphi) = 1 - Z(1) = 1$.

$$G = 1 - 2\int_0^1 Q(\gamma)\mathrm{d}\gamma = \frac{\sigma - 1}{2k + 1 - \sigma}. \tag{74}$$

Because the relative revenue shares are not affected by the sector choice, the derivation of the Product-Gini is the same for all firms and the product revenue distribution follows the Pareto product skill distribution. Given its relative nature, G is independent of firm productivity φ. Therefore, all firms are equally diversified as measured by the Product-Gini. Moreover, product diversification measured by the Product-Gini is increasing in the elasticity of substitution σ.[38]

F.3 Herfindahl Index

The revenue share $s(\beta)$ of a product produced with product skill β of total firm revenue is

$$s(\beta) = \frac{r(\varphi, \beta)}{\int_1^\infty r(\varphi, \beta)z(\beta)\mathrm{d}\beta} = \frac{k + 1 - \sigma}{k}\beta^{\sigma - 1} \tag{75}$$

and the Herfindahl index follows from

$$H = \int_1^\infty s(\beta)^2 z(\beta)\mathrm{d}\beta = \frac{(k + 1 - \sigma)^2}{k(k + 2 - 2\sigma)}. \tag{76}$$

This derivation follows the same procedure for informal, formal and exporting firms, as sector specific differences do not influence the relative revenue shares. Moreover, all firms produce the whole continuum of products. Therefore, all firms are equally diversified, as measured by the Herfindahl index, and diversification is increasing in the elasticity of substitution σ.[39]

[38] $\frac{\partial G}{\partial \sigma} = \frac{2k}{(\sigma - 2k - 1)^2} > 0.$

[39] $\frac{\partial H}{\partial \sigma} = \frac{-2k(k+1-\sigma)(k+2-2\sigma)+2k(2k+2-2\sigma)}{k^2(k+2-2\sigma)^2} > 0$, because $\sigma > 1$.

References

Allison, P. D. 1978. Measures of inequality. *American Sociological Review*, 43(6):865–880.

Antunes, A. R. and de V. Cavalcanti, T. V. 2007. Start up costs, limited enforcement, and the hidden economy. *European Economic Review*, 51(1):203–224.

Arkolakis, C. and Muendler, M.-A. 2010. The extensive margin of exporting products: A firm-level analysis. Working Paper 16641, National Bureau of Economic Research.

Auriol, E. and Warlters, M. 2005. Taxation base in developing countries. *Journal of Public Economics*, 89(4):625–646.

Axtell, R. L. 2001. Zipf distribution of U.S. firm sizes. *Science*, 293(5536):1818–1820.

Baldwin, J. and Gu, W. 2009. The impact of trade on plant scale, production-run length and diversification. In Dunne, T., Jensen, J. B., and Roberts, M. J., editors, *Producer Dynamics: New Evidence from Micro Data*, NBER Chapters, pages 557–592. National Bureau of Economic Research, Inc, Cambridge, MA.

Banerji, A. and Jain, S. 2007. Quality dualism. *Journal of Development Economics*, 84(1):234–250.

Basu, A. K., Chau, N. H., and Kanbur, R. 2011. Contractual dualism, market power and informality. IZA Discussion Papers 5845, Institute for the Study of Labor (IZA).

Batra, G., Kaufmann, D., and Stone, A. H. W. 2003. *Investment climate around the world: Voices of the firms from the world business environment survey*. World Bank, Washington, D.C.

Becker, D. 2014. Heterogeneous firms and informality: The effects of trade liberalization on labor markets. Working Paper 2014-13, Charles H. Dyson School of Applied Economics and Management, Cornell University.

Bernard, A. B. and Jensen, J. B. 1995. Exporters, jobs, and wages in U.S. manufacturing: 1976-1987. *Brookings Papers on Economic Activity. Microeconomics*, 1995:67–112.

Bernard, A. B. and Jensen, J. B. 1999. Exceptional exporter performance: Cause, effect, or both? *Journal of International Economics*, 47(1):1–25.

Bernard, A. B., Jensen, J. B., and Schott, P. K. 2009. Importers, exporters and multinationals: A portrait of firms in the U.S. that trade goods. In Dunne, T., Jensen, J. B., and Roberts, M. J., editors, *Producer Dynamics: New Evidence from Micro Data*, NBER Chapters, pages 513–552. National Bureau of Economic Research, Inc, Cambridge, MA.

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

Bernard, A. B., Redding, S. J., and Schott, P. K. 2011. Multiproduct firms and trade liberalization. *Quarterly Journal of Economics*, 126(3):1271–1318.

Bigsten, A., Kimuyu, P., and Lundvall, K. 2004. What to do with the informal sector? *Development Policy Review*, 22(6):701–715.

Bourguignon, F. 1979. Decomposable income inequality measures. *Econometrica*, 47(4):901–920.

Cadot, O., Carrére, C., and Strauss-Kahn, V. 2011. Export diversification: What's behind the hump? *Review of Economics and Statistics*, 93(2):590–605.

Cadot, O., Carrére, C., and Strauss-Kahn, V. 2013. Trade diversification, income, and growth: What do we know? *Journal of Economic Surveys*, 27(4):790–812.

CIEM. 2012. Characteristics of the Vietnamese business environment: Evidence from a SME survey in 2011. Technical report, Central Institute for Economic Management (CIEM), Hanoi.

Dabla-Norris, E., Gradstein, M., and Inchauste, G. 2008. What causes firms to hide output? The determinants of informality. *Journal of Development Economics*, 85(12):1–27.

de Beer, J., Fu, K., and Wunsch-Vincent, S. 2013. The informal economy, innovation and intellectual property - Concepts, metrics and policy considerations. Economic Research Working Paper 10, World Intellectual Property Organization.

de Mel, S., McKenzie, D., and Woodruff, C. 2013. The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5 (2):122–150.

de Paula, A. and Scheinkman, J. A. 2011. The informal sector: An equilibrium model and some empirical evidence from Brazil. *Review of Income and Wealth*, 57:8–26.

de Soto, H. 1989. *The other path*. Harper and Row, New York, NY.

Delgado, M. A., Fariñas, J. C., and Ruano, S. 2002. Firm productivity and export markets: A non-parametric approach. *Journal of International Economics*, 57(2):397–422.

Djankov, S., La Porta, R., Lopez-De-Silanes, F., and Shleifer, A. 2002. The regulation of entry. *Quarterly Journal of Economics*, 117(1):1–37.

Eckel, C. and Neary, J. P. 2010. Multi-product firms and flexible manufacturing in the global economy. *Review of Economic Studies*, 77(1):188–217.

Essegbey, G. O., Awuni, S., Essegbey, I. T., Akuffobea, M., and Micah, B. 2014. Country study on innovation, intellectual property and the informal economy: Traditional herbal medicine in Ghana. Committee on Development and Intellectual Property (CDIP), Thirteenth Session CDIP/13/INF/2, World Intellectual Property Organization.

Feenstra, R. and Ma, H. 2007. Optimal choice of product scope for multiproduct firms under monopolistic competition. Working Paper 13703, National Bureau of Economic Research.

Fields, G. S. 1975. Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs. *Journal of Development Economics*, 2(2):165–187.

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

Helpman, E., Melitz, M. J., and Yeaple, S. R. 2004. Export versus FDI with heterogeneous firms. *American Economic Review*, 94(1):300–316.

Iacovone, L. and Javorcik, B. S. 2008. Multi-product exporters: Diversification and micro-level dynamics. Policy Research Working Paper Series 4723, World Bank.

Imbs, J. and Wacziarg, R. 2003. Stages of diversification. *American Economic Review*, 93(1):63–86.

Klinger, B. and Lederman, D. 2004. Discovery and development: An empirical exploration of "new" products. Policy Research Working Paper Series 3450, World Bank.

Koren, M. and Tenreyro, S. 2007. Volatility and development. *Quarterly Journal of Economics*, 122(1):243–287.

La Porta, R. and Shleifer, A. 2008. The unofficial economy and economic development. *Brookings Papers on Economic Activity*, 2008:275–352.

La Porta, R. and Shleifer, A. 2014. Informality and development. *Journal of Economic Perspectives*, 28(3):109–126.

Loayza, N. V. 1996. The economics of the informal sector: A simple model and some empirical evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45:129–162.

Loayza, N. V., Oviedo, A. M., and Serven, L. 2005. The impact of regulation on growth and informality - cross-country evidence. Policy Research Working Paper Series 3623, World Bank.

Mayer, T., Melitz, M. J., and Ottaviano, G. I. 2011. Market size, competition, and the product mix of exporters. Working Paper 16959, National Bureau of Economic Research.

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

Navarro, L. 2012. Plant level evidence on product mix changes in Chilean manufacturing. *Journal of International Trade & Economic Development*, 21(2):165–195.

Nocke, V. and Yeaple, S. 2013. Globalization and multiproduct firms. Working Paper 19409, National Bureau of Economic Research.

Rauch, J. E. 1991. Modelling the informal sector formally. *Journal of Development Economics*, 35 (1):33–47.

Roberts, M. J. and Tybout, J. R. 1997. The decision to export in Colombia: An empirical model of entry with sunk costs. *American Economic Review*, 87(4):545–564.

Schneider, F. and Enste, D. H. 2000. Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38(1):77–114.

Schneider, F., Buehn, A., and Montenegro, C. E. 2010. New estimates for the shadow economies all over the world. *International Economic Journal*, 24(4):443–461.

Schoar, A. 2002. Effects of corporate diversification on productivity. *Journal of Finance*, 57(6): 2379–2403.

Söderbom, M. and Weng, Q. 2012. Multi-product firms, product mix changes and upgrading: Evidence from China's state-owned forest areas. *China Economic Review*, 23(4):801–818.

Thon, D. 1982. An axiomatization of the Gini coefficient. *Mathematical Social Sciences*, 2(2): 131–143.

Ulyssea, G. 2010. Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, 91(1):87–99.

Verhoogen, E. A. 2008. Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector. *Quarterly Journal of Economics*, 123(2):489–530.

Wagner, J. 2007. Exports and productivity: A survey of the evidence from firm-level data. *World Economy*, 30(1):60–82.

World Bank. 2013. *Doing Business 2014: Understanding Regulations for Small and Medium-Size Enterprises*. World Bank Group, Washington, D.C.

OTHER A.E.M. WORKING PAPERS

WP No	Title	Fee (if applicable)	Author(s)
2014-20	Social Protection, Vulnerability and Poverty		Kanbur, R.
2014-19	What is a "Meal"? Comparative Methods of Auditing Carbon Offset Compliance for Fuel Efficient Cookstoves		Harrell, S., Beltramo, T., Levine, D., Blalock, G. and A. Simons
2014-18	Informality: Causes, Consequences and Policy Responses		Kanbur, R.
2014-17	How Useful is Inequality of Opportunity as a Policy-Construct?		Kanbur, R. and A. Wagstaff
2014-16	The Salience of Excise vs. Sales Taxes on Healthy Eating: An Experimental Study		Chen, X., Kaiser, H. and B. Rickard
2014-15	'Local' producer's production functions and their importance in estimating economic impacts		Jablonski, B.B.R. and T.M. Schmit
2014-14	Rising Inequality in Asia and Policy Implications		Zhuang, J., Kanbur, R. and C. Rhee
2014-13	Heterogeneous firms and informality: the effects of trade liberalization on labor markets		Becker, D.
2014-12	Minimum Wage Effects at Different Enforcement Levels: Evidence from Employment Surveys in India		Soundararajan.V.
2014-11	Thresholds, Informality, and Partitions of Compliance		Kanbur, R. and M. Keen
2014-10	On the Political Economy of Guest Workers Programs in Agriculture		Richard, B.
2014-09	Groupings and the Gains from Tagging		Kanbur, R. and M. Tuomala
2014-08	Economic Policy in South Africa: Past, Present, and Future		Bhorat, H., Hirsch, A., Kanbur, R. and M. Ncube
2014-07	The Economics of China: Successes and Challenges		Fan, S., Kanbur, R., Wei, S. and X. Zhang
2014-06	Mindsets, Trends and the Informal Economy		Kanbur, R.

Paper copies are being replaced by electronic Portable Document Files (PDFs). To request PDFs of AEM publications, write to (be sure to include your e-mail address): Publications, Department of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853-7801. If a fee is indicated, please include a check or money order made payable to Cornell University for the amount of your purchase. Visit our Web site (*http://dyson.cornell.edu/research/wp.php)* for a more complete list of recent bulletins.