Intra- and inter-industry misallocation and comparative advantage*

José Pulido†

May 2019
[Latest version here]

Abstract

What are the implications of factor misallocation in an open economy? This paper shows how firm-level resource misallocation can affect the relative unit cost of producing a good across sectors, distorting the "natural" comparative advantage of a country. First, sectors with a larger extent of within-industry factor misallocation face larger productivity losses, which reduce their relative export capability. Second, misallocation of factors across industries can alter sectors' sizes and distort their average productivity through firms' selection effects, affecting their comparative advantage too. After presenting evidence on how metrics of intra- and inter-industry factor misallocation are related to the observed patterns of comparative advantage, this paper explores the general equilibrium effects of both types of misallocation in an open economy and their role in shaping industry export capabilities. For this, I use a model of international trade with endogenous selection of heterogeneous firms in which the allocation of multiple factors within and across industries is inefficient. I compute a counterfactual equilibrium in which misallocation in capital, skilled labor and unskilled labor is removed in Colombia, a country whose firm-level data allows me to obtain precise measures of misallocation. The reallocation of factors allows Colombia to specialize in industries with "natural" comparative advantage and generates a substantial change in its industrial composition, which leads to a rise in the ratio of exports to GDP by 18 p.p. This industrial composition effect is absent in the workhorse models of factor misallocation under closed economies.

Keywords: Comparative advantage, firm-level misallocation, TFP, general equilibrium.
J.E.L. Classification: F12, D24, D61

---

*I am especially grateful for the advice and guidance of my PhD supervisor at the University of British Columbia, Matilde Bombardini, and my committee Tomasz Święcki, Amartya Lahiri and Vanessa Alviarez. I thank Yaniv Yedid-Levi, Jesse Perla, Victoria Hnatkovska, Giovanni Gallipoli and Nouri Najjar for their helpful comments and collaboration. I also thank Rodrigo Adão, Andrés Rodríguez-Claire, Svetlana Demidova, Chang Tai Hsieh and Loukas Karabarbounis for their feedback, and the seminar participants at the 2017 CEA, Central Bank of Colombia, Universidad del Rosario, Universidad EAFIT and the Sauder-VSE trade group, for their comments. I am grateful to the Colombian Statistics Institute (DANE) for providing access to the manufacturing data. All remaining errors are mine.

†Banco de la República, Colombia. E-mail: jpulidpe@banrep.gov.co.

1 Introduction

What are the implications of firm-level factor misallocation in open economies? In recent years, a growing body of research has strived to understand how factor misallocation across heterogeneous firms can account for differences in aggregate productivity across countries.[1] The main insight from this literature is that, given a fixed endowment of production factors in the economy and a certain distribution of physical productivity across firms, the inefficient allocation of inputs across production units generates sizable losses in aggregate TFP. Under standard assumptions on the demand and production structure, and regardless the underlying cause of the inefficient use of resources – regulations, financial constraints, information asymmetries, crony capitalism, etc. – the amount of misallocation can be measured by the extent to which the marginal returns to factors varies within countries. Some evidence suggests a broader dispersion of those returns in developing economies (Banerjee and Duflo (2005), Hsieh and Klenow (2009), Bartelsman, Haltiwanger and Scarpetta (2013)), implying larger productivity losses for those countries.

However, most of the literature on the effects of resource misallocation on the aggregate economic performance has focused on closed economies.[2] In open economies, if the extent of factor misallocation varies not only across countries but also across industries, it could also shape the relative unit costs of production across sectors, distorting the comparative advantage of country.[3] For example, consider the broad range of industrial policies that several East Asian countries introduced during the post-war period, intended to promote some strategic industries. Such policies could have generated not only reallocation of factors towards targeted industries but also an increase in resource misallocation across firms within those sectors given the distortionary nature of some instruments used: selective investment tax credits, public enterprises, depreciation allowances, etc.[4] Thus, the likely improvement in the export capability of targeted sectors due to the reduction in the average cost of the factors, compared to untargeted industries, could have been countered by decreases in their sectoral TFP, due to their larger extent of within-industry factor misallocation. A relevant

[1]For an extensive review, see Restuccia and Rogerson (2013) or Hopenhayn (2014*a*).

[2]In the trade literature, most of the analysis has been addressed from a different angle: the effect of trade on a metric of firm-level misallocation, such as mark-ups dispersion (Epifani and Gancia (2011), Edmond, Midrigan and Daniel Yi (2015)) or how much plant survival depend on productivity (Eslava et al. (2013)). Others have studied the effects of trade liberalization for welfare in economies with factor misallocation, papers that are mentioned below.

[3]As usual, comparative advantage describes the differences of the average unit cost of a good across industries relative to the same differences in a reference country. Hence, the sources of comparative advantage comprise all primitive variables that affect the three determinants of the unit costs in an industry: sectoral average productivities, factors prices and the number of varieties produced. Those sources include not only "natural" differences in technology distributions or factor endowments, but also, in a world with economies to scale, differences in the primitive determinants of industries' scale (i.e. entry barriers) and, as I show in this paper, the extent of resource misallocation both within and across industries.

[4]For details of East Asian industry policies, see for example Rodrik (1995), Chang (2006) or Lane (2017).

question here is then how to assess the role of those policies in shaping comparative advantage through their effect on the allocation of resources. Did those policies accentuate or distort the "frictionless" patterns of industrial specialization?

This paper explores how firm-level factor misallocation can influence the core determinants of industries' export capabilities in an open economy, and hence, the patterns of industrial specialization. I do this by addressing the following two questions. First, does resource misallocation explain observed industries' export capabilities once we control for the "frictionless" sources of comparative advantage? Second, if so, what are the implications of removing such misallocation for the comparative advantage of a country and its industrial composition taking into account all general equilibrium effects?

To verify the role of firm-level factor misallocation as a determinant of comparative advantage, I first present empirical evidence on how standard metrics of firm-level misallocation are related to the observed patterns of export capability of Colombian industries, once we control for the "natural" determinants of comparative advantage. The choice of Colombia is due to the fact that its manufacturing firm-level data, considered one of the richest in the world (De Loecker and Goldberg (2014)), offers a better understanding of the role of firms' efficiency in aggregate productivity. A unique feature of the data is the possibility to obtain direct measures of firms' physical productivity (TFPQ) using plant-level deflators for firms' inputs and outputs. Those measures of TFPQ allow me to decompose the contribution of efficiency, demand shocks and factor distortions in the sectoral TFP. As my metric of export capability, I use the estimates of the exporter-industry fixed effect derived from a gravity equation, an approach that has gained popularity as a measure of "revealed" comparative advantage, RCA hereafter (Costinot, Donaldson and Komunjer (2012), Levchenko and Zhang (2016), Hanson, Lind and Muendler (2015), French (2017)). I regress the Colombian RCA measure relative to the United States on indicators of both intra- and inter-industry misallocation, exploiting their variation over time. I control for the "natural" sources of comparative advantage using total endowments interacted with factor intensities and efficient sectoral productivities, which capture Heckscher-Ohlin and Ricardian forces respectively. I find that firm-level misallocation have a quantitative relevance for shaping Colombian RCA with a magnitude similar to the one observed for the "natural" determinants.

Next, I examine the general equilibrium channels with which firm-level factor misallocation can shape relative industries' unit costs and hence comparative advantage. This exploration, which is the main contribution of this paper, takes into account several adjustments that are absent when removing factor misallocation under a closed economy. For example, consider first the impact of firm-level misallocation within industries only. As it is well known, this type of misallocation generates losses in sectoral TFP. In a closed economy setting with a fixed mass of firms, as in Hsieh and Klenow (2009), HK hereafter, the gains in sectoral efficiency from removing intra-industry misallocation do not generate reallocation of factors across sectors under the standard two-tier (Cobb Douglas-CES) demand system.[5] Instead, in an open economy, even with the same demand structure and a fixed mass of firms, sectoral revenue shares are endogenously determined and depend not only on how substitutable goods are across sectors, but also on the gains from industrial specialization due to comparative advantage. Removing intra-industry misallocation in a country leads to two types of adjustments on factor prices, absent in a closed economy. First, it produces a change in the relative factor prices across countries to restore trade balance equilibrium, a result analogous to the introduction of a set of sectoral-specific productivity shocks in standard Ricardian models. And second, it changes the relative real factor returns depending on the adjustment of relative prices of goods, as in the standard Heckscher-Ohlin model.

Furthermore, when allowing for endogenous entry and selection across firms, as in the closed economy models of Bartelsman, Haltiwanger and Scarpetta (2013), Yang (2017) or Adamopoulos et al. (2017), TFP gains and their general equilibrium effects on factor prices are magnified by the adjustment in the extensive margin (the number of operating firms) after removing misallocation. This effect is sizable since it involves a drastic recomposition of incumbent firms: a withdrawal of low-efficiency firms that survived because of factor misallocation plus the addition of potential high-efficiency firms that were not able to operate under allocative inefficiency. In monopolistically competitive industries this recomposition of firms affect the scale of the sectors and the average productivities through firms' selection effects, impacting industries' relative unit costs. Finally, the marginal returns of the factors might differ on average across sectors, suggesting the presence of inter-industry misallocation as well. Simultaneously removing this type of misallocation affects the direction of sectoral factor reallocations and the magnitude of the adjustments on relative factor prices, which produces further adjustments on average productivities under self-selecting heterogenous firms.

To consider all these general equilibrium channels, I use a tractable multi-country, multi-factor and multi-sector model of international trade à la Melitz (2003) in which the allocation of multiple factors across heterogeneous firms is inefficient. I employ wedge analysis to characterize the observed dispersion in the marginal returns of the factors abstracting from the underlying cause of misallocation, an approach introduced by Restuccia and Rogerson (2008) and HK in this context and inspired by the business cycle literature.[6] Under this approach, each firm is represented by a draw of "true" efficiency – physical productivity or TFPQ – and a

[5]Constant revenue shares across sectors imply that the efficiency gained by each industry, translated into a lower aggregate price index, is automatically followed by an increase in demand, so there are not inter-industry factor reallocations and their relative prices do not adjust. Under a more general demand (two-tier CES) there is reallocation of factors across sectors, but abstracting from inter-industry misallocation, the effect on factor prices is marginal (see HK and Appendix C).

[6]Wedge analysis was first developed as accounting methodology in the business-cycle literature by Cole and Ohanian (2002), Mulligan (2005), Chari, Kehoe and McGrattan (2007) and Lahiri and Yi (2009) among others. For recent uses in the literature on factor misallocation, see for example Adamopoulos et al. (2017), Brandt, Tombe and Zhu (2013), Bartelsman, Haltiwanger and Scarpetta (2013), Gopinath et al. (2017), Hopenhayn (2014*b*), Oberfield (2013), Święcki (2017), Tombe (2015) and Yang (2017) among others.

vector of wedges, whose elements represent the differences between the returns of each primary factor for the firm and the average returns in the economy. I derive a theoretically consistent gravity equation along the lines of Chaney (2008), Arkolakis, Costinot and Rodríguez-Clare (2012) and Melitz and Redding (2014) that incorporates the impact of wedges on the determinants of bilateral exports, in particular on the exporter industry fixed effect, my measure of RCA. I then investigate the effect of removing firm-level misallocation of a country on its bilateral exports and hence on its RCA.

To this end, I obtain counterfactual equilibria solving the model in relative changes, using the "exact hat algebra" method proposed by Dekle, Eaton and Kortum (2008). Each counterfactual incorporates the whole set of general equilibrium effects of reallocating factors to their efficient allocation and is not demanding in terms of data requirements. I perform the exercises using a world composed of 47 countries and an aggregate rest of the world, three production factors and 25 tradable sectors, to evaluate the effect of Colombian firm-level factor misallocation on its comparative advantage schedule. I use Bils, Klenow and Ruane's (2018) method to estimate the dispersion in marginal products in the presence of additive measurement error in revenue and inputs. This methodology exploits the fact that in the absence of measurement error the elasticity of revenues with respect to inputs should not vary for plants with different average products. Hence, panel data can be used to back out the "true" marginal product dispersion by estimating how such elasticity changes for plants with different average products. Moreover, since overhead factors (necessary to account for endogenous selection) are analogous to an unobservable additive term in measured inputs, this methodology allows me to overcome the problem of measuring the variance of the marginal products of the factors directly from the dispersion of their average products in the presence of fixed costs; a key issue of models with self-selection of heterogenous firms (Bartelsman, Haltiwanger and Scarpetta (2013)).

The results of the counterfactual exercise suggest that in Colombia resource misallocation plays a major role in shaping comparative advantage. In the case of an extreme reform in which factor misallocation is entirely removed within and across industries, the ratio of exports to manufacturing GDP rises by 18 p.p. and welfare, measured as real expenditure, grows 75%.[7] The large boost in exports is due to the increase in the dispersion of the schedule of comparative advantage, which leads to higher degrees of industrial specialization in the frictionless equilibrium. For instance, the whole chemical sector (both industrial chemicals and other chemicals such as paints, medicines, soaps or cosmetics) climbs to the top of the national export capability ranking, and ends up in the first percentile of the counterfactual RCA world distribution. The opposite case occurs in industries whose comparative advantage in the actual data seems to be due only to factor misallocation, particularly computer, electronic and optical products, transportation equipment, petroleum and machinery and equipment.

[7]The growth in real expenditure is equivalent to the TFP gains in a closed economy model.

These four industries shrink and practically disappear, indicating a non-interior solution in the counterfactual equilibrium.[8]

The model also delivers a decomposition of the change in the RCA measure after removing factor misallocation into three terms, each of which corresponds to a single component of the relative unit cost across industries: the average TFP, factors prices, and the number of produced varieties. I find that the adjustment in the relative number of produced varieties (i.e., in the extensive margin), which is generated by the reallocation of factors across industries, contributes the most to the change in the RCA. This is because in the intensive margin the gains in average TFP relative to the rest of the world are offset in large part by the rise in the relative factor prices, and the remaining effect does not vary much across industries.

Related literature

This paper belongs to a recent strand of research that evaluates the implications of factor misallocation in open economies, as in Ho (2012), Tombe (2015), Święcki (2017), Caliendo, Parro and Tsyvinski (2017), Costa-Scottini (2018), Berthou et al. (2018) and Chung (2018). My approach is different with respect to most of those papers and is in line with Caliendo, Parro and Tsyvinski's (2017) one: instead of analyzing the effect of a trade reform in an economy with factor misallocation, the objective is to evaluate the consequences of removing the observed misallocation on the structure of the economy, particularly on the patterns of industrial specialization due to comparative advantage. Regarding my theoretical framework, the papers with the closest models to the one used here are Ho (2012), Costa-Scottini (2018), Berthou et al. (2018) and Chung (2018); which use different variations of open-economy models with firm-level factor distortions and selection effects of heterogenous firms. My multi-country, multi-sector and multi-factor model shares some features with those papers, but it differs in several aspects.[9] My empirical implementation is also different, since I obtain counterfactuals without relying on the combination of estimating and calibrating large sets of structural parameters.[10] Further, different to Tombe (2015) and Święcki (2017), who use a Eaton and Kortum's (2002) type of model to study welfare and the gains from trade under

[8] The feasibility of non-interior solutions in multi-sector Pareto-Melitz type of models is established by Kucheryavyy, Lyn and Rodríguez-Clare (2019). Under a similar setup to the one used in this paper, it is guaranteed that the general equilibrium is unique, but not necessarily an interior solution.

[9] Unlike the mentioned papers and because my main focus is on comparative advantage, I let misallocation arise in any factor market. This can distort industries' advantages in unit costs based on the relative size of the countries' factor endowments (Heckscher-Ohlin forces). My framework also accounts simultaneously for both intra- and inter-industry misallocation. Finally, unlike Ho (2012) and Costa-Scottini (2018), I do not constrain factor distortions to be size-dependent. With size-dependent distortions, the model behaves exactly as a Melitz model with a unique physical productivity cut-off. Thus, the selection effects of distortions do not generate rank-reversals, which are necessary to obtain the large TFP gaps attributed to factor misallocation (Hopenhayn (2014*a*),Hopenhayn (2014*b*)).

[10] Instead, I use the "exact hat algebra" method proposed by Dekle, Eaton and Kortum (2008) that is not demanding in terms of data requirements.

the presence of sectoral distortions, and thus only inter-industry misallocation, my model is able to generate ex-post misallocation across industries as result of differences in the first and second moments of the underlying distributions of factor distortions across sectors, which allows me to have rich interactions between the extent of intra- and inter-industry factor misallocation.

My model has the same interactions between country, industry, and firm characteristics in general equilibrium as the multi-factor models that exhibit factor reallocations, both within and across industries, in response to trade shocks, particularly Bernard, Redding and Schott (2007) and Balistreri, Hillberry and Rutherford (2011). In my case, the introduction of resource misallocation generates a new source of comparative advantage that distorts the frictionless trade equilibrium. Instead of a full characterization of the inefficient equilibrium properties, my focus is mostly on the implications of allocative inefficiency for the industrial specialization patterns. Therefore, my primary interest relies on the counterfactual exercise of removing the misallocation. Finally, this paper is also related to the trade literature concentrated on gravity equations to derive indirect measures of relative export capability, as in Costinot, Donaldson and Komunjer (2012), Hanson, Lind and Muendler (2015), Levchenko and Zhang (2016), and French (2017). I use the same approach to obtain revealed comparative advantage measures, which are the main metric of interest in my counterfactual exercises.

The organization of this paper is as follows. Section 2 presents the empirical motivation. I first introduce the empirical measure of RCA derived from a standard gravity equation, and next I propose a strategy to evaluate the impact of different metrics of Colombian factor misallocation on its comparative advantage. Section 3 introduces the theoretical model and derives the effect of firms' wedges on the gravity equation, particularly on exporter-industry fixed effects, the measure of RCA. I also offer an overview of the general equilibrium channels that each type of misallocation can trigger using model simulations under a simple parametrization. Section 4 presents the counterfactual exercise of removing firm-level misallocation in Colombia, to compute the effect of the two types of misallocation on its industries' comparative advantage. I also evaluate some departures from the baseline model. Section 5 concludes.

2 Empirical motivation

In this section I present empirical evidence on how factor misallocation is related to the comparative advantage of a country. For this, I first introduce the empirical measure of RCA derived from a standard gravity equation and I explain how this measure is linked to the relative producer price index. Next, I decompose the price index in terms of the "natural" sources of comparative advantage and metrics of factor misallocation. Finally, I propose a strategy to evaluate the relation between the metrics of factor misallocation and the measures of RCA, controlling for the "natural" sources of comparative advantage.

2.1 A measure of RCA

A wide range of the new trade models deliver a gravity equation, in which comparative advantage has an important role as a predictor of bilateral trade flows. In the generic formulation of the gravity equation, bilateral exports of country i to country j, denoted by X_{ij}, can be expressed as the combination of three forces: i) a factor that represents "capabilities" of exporter i as a supplier to all destinations; ii) a factor that characterizes the demand for foreign goods of importer j; iii) a factor that captures bilateral accessibility of destination j to exporter i, which combines trade costs and other bilateral frictions. The gravity equation can be estimated at the industry level, in order to reduce aggregation bias.[11] With cross-sectional data the standard procedure involves taking logs and estimating a regression with fixed effects:

$$\ln x_{ijs} = \delta_{is} + \delta_{js} + \delta_{ij} + \varepsilon_{ijs} \tag{1}$$

where δ_{is}, the exporter-industry fixed effect, characterizes factor i), "capabilities" of exporter i in industry s; δ_{js}, the importer-industry fixed effect, captures factor ii), the demand for foreign goods of importer j in industry s; and $\delta_{ij} + \varepsilon_{ijs}$ represent factor iii), bilateral accessibility of j to i, a component that involves characteristics of the bilateral relation independent of the sector (distance, common language, etc.), absorbed by the exporter-importer fixed effect δ_{ij}, plus sector-specific bilateral frictions and measurement error, represented by the term ε_{ijs}.

In this way, the estimate of the industry-exporter fixed effect characterizes the relative country's productive potential in an industry and, given the structure of the gravity equation, it is "clean" from other determinants that affect bilateral trade flows. Since it is only identified up to a double normalization, that is, it has meaning only when it is compared to a reference country and industry, it can be interpreted as a measure of "revealed" comparative advantage (RCA), an approach that has increasingly gained relevance in the trade literature (Costinot, Donaldson and Komunjer (2012), Hanson, Lind and Muendler (2015), and Levchenko and Zhang (2016)). In contrast to traditional measures of RCA, as Balassa's (1965) index, the fixed effect estimate is a valid measure of countries' fundamental patterns of comparative advantage (French (2017)). Moreover, it has better statistical properties than Balassa's index, especially lower ordinal ranking bias and higher time stationarity (Leromain and Orefice (2014)).

Figure 1 displays for Colombia the RCA measures of the 25 manufacturing industries listed in Table A.2 of the Appendix A.1. I rely on the CEPII trade and production database, developed for de Sousa, Mayer and Zignago (2012). I use bilateral trade flows among 47 countries plus a rest of the world aggregate for 1995. The set of countries is listed in Table A.1 of Appendix A.1. Similar to Hanson, Lind and Muendler (2015), I use as a reference country and industry the mean over all countries and industries, so the RCA can be interpreted as a measure of Colombian industries' capabilities relative to a "typical" country and a "typical"

[11] For a detailed explanation about the necessary conditions for a trade model to yield a structural gravity equation, see Head and Mayer (2014). On the aggregation bias see Anderson and Yotov (2010, 2016).

sector.[12] The logarithmic transformation in equation (1) poses two well-known econometric issues for an estimation by OLS. First, zeros in bilateral exports are not likely random in the data, and since OLS drops those observations, it introduces sample-selection bias. Second, the coefficients of log-linearized models estimated by OLS are biased in the presence of heteroskedasticity (Silva and Tenreyro, 2006). In Monte Carlo simulations, Head and Mayer (2014) find that the Tobit model proposed in Eaton and Kortum (2001) (EK-Tobit hereafter) and the Poisson pseudo-maximum-likelihood estimator (PPML hereafter) proposed in Silva and Tenreyro (2006) are the two estimating methods which, depending on the structure of the error of the underlying data generating process, produce unbiased coefficients for exogenous variables in a gravity formulation.[13] Thus, Figure 1 compares the estimates obtained by EK-Tobit (vertical axis) and PPML (horizontal axis). Noticeably, the ranking across sectors in the cross section is not strongly affected by the estimation method.

The determinants of the exporter-industry fixed effect vary according to the sources of comparative advantage in the considered theoretical model. However, a common feature across all standard models is that such determinants are collapsed in the reduced-form of the relative producer price index at the industry level compared to a reference country ($\frac{P_{is}P_{i's'}}{P_{is'}P_{i's}}$), as a measure of the relative unit cost of producing across industries (French (2017)).[14] For example, in Ricardian models, as in Eaton and Kortum (2002), such ratio depends only on sectoral fundamental efficiencies, the source of comparative advantage at the heart of the Ricardian theory.[15] In a Heckscher-Ohlin model, as in Deardorff (1998), the ratio depends on the factor prices weighted by sectoral factor intensities, reflecting the balance between the relative sizes of factor endowments and the technology requirements. In the Krugman (1980) model, it depends only on the relative number of varieties produced, quantifying the effect of differences in the increasing returns to scale across industries on the aggregate prices. In the Pareto version of the Melitz (2003) model, the ratio is analogous to that in Krugman (1980), adjusted by the lower bound of the Pareto's productivity distribution, so the support of the firms' physical productivities also plays a role. Multi-factor models with heterogenous firms, as in Bernard, Redding and Schott (2007) or in this paper, combine all mentioned sources of

[12]Therefore, letting $\hat{\delta}_{is}$ be an estimate of δ_{is} in regression (1), RCA of country i in sector s is defined as:

$$RCA_{is} = \left[exp(\hat{\delta}_{is})/exp(\sum_{s}^{S} \frac{1}{S}\hat{\delta}_{is}) \right] / \left[exp(\sum_{i}^{N} \frac{1}{N}\hat{\delta}_{is})/exp(\sum_{s}^{S}\sum_{i}^{N} \frac{1}{S*N}\hat{\delta}_{is}) \right]$$

[13]Under heteroskedasticity in the form of a constant variance to mean ratio PPML performs better, whereas under homoskedastic log-normal errors the Tobit proposed by Eaton and Kortum (2001) is preferred.

[14]Strictly, French (2017) shows that country i has comparative advantage in sector s, compared to country $i\prime$ and industry $s\prime$, if the relative price of country i in sector s in autarky is smaller than the same price in country i': $\frac{\bar{P}_{is}\bar{P}_{i's'}}{\bar{P}_{is'}\bar{P}_{i's}} < 1$ where $\bar{P}_{is}$ is the counterfactual price index in industry s of country i in autarky.

[15]The implicit assumption is that sectors share the same intra-industry heterogeneity in the distribution of varieties' productivities. If the heterogeneity varies across sectors, the productivity dispersion can be an additional source of comparative advantage (Bombardini, Gallipoli and Pupato (2012)).

comparative advantage in the reduced form of the relative price index.

The model with resource misallocation in an open-economy in the next section delivers an analytical expression of the exporter-industry fixed effect taking into account endogenous entry and selection of firms, features that will provide a rich theoretical grounding to the RCA measure. However, at this point we can use the insights from the most well-known misallocation framework, Hsieh and Klenow (2009) (HK hereafter), to decompose the producer price index in its different determinants and empirically test whether the components due to firm-level misallocation are related to the metrics of RCA, once we control for the remaining sources of export capability.

2.2 Decomposing the price index under factor misallocation

The starting point in the HK framework to evaluate the implications of firm-level factor misallocation relies on the distinction between physical productivity (TFPQ, defined as the ratio of physical output to inputs) and revenue productivity (TFPR, defined as the ratio of revenues to inputs), first proposed by Foster, Haltiwanger and Syverson (2008). Assume a standard monopolistic competition framework in which firms differ in terms of efficiency –i.e. in the TFPQ or Hicks-neutral productivity–, but use the same constant returns to scale technology in each industry. Moreover, assume firms face a CES demand, with the same elasticity of substitution in all industries. In this simple economy if factor markets are frictionless the following two implications emerge: i) TFPR is equalized across firms within industries;[16] and ii) the sectoral TFP can be computed as a power mean of firms' TFPQ. Any dispersion in firms' TFPR within a sector is a signal of within-industry factor misallocation, and leads to a loss in sectoral TFP.

Of course, the reliability of the dispersion of TFPR as a measure of intra-industry factor misallocation depends on the plausibility of the considered assumptions. Some recent papers have tried to quantify the contribution of other possible sources of variation in TFPR, that do not imply factors are misallocated. Those sources can be classified in two categories: i) model misspecification (by incorrectly not assuming heterogeneity in inputs, variable markups or adjustments costs, for example) and ii) pure measurement error. In Table 1 I present a brief survey of some calculations in the recent literature regarding the contributions of those possible sources of variation in TFPR, each one derived from an extended structural model that takes into account the corresponding cause. The main conclusion is that the sources related to model misspecification individually have a relative small contribution. For the Colombian case, this is in line with Eslava and Haltiwanger (2018), who show, using an extension of the generic model to account for demand shocks (idiosyncratic at the firm and the firm-product

[16] This is simply because TFPR is the product of firm's price and TFPQ. With constant mark-ups, prices vary across firms only due to marginal costs. In turn, with all firms facing the same factor prices and the described technologies, the only source of variation in marginal costs is TFPQ. Hence, differences in TFPQ are perfectly translated into (the inverse of) prices, leaving TFPR invariant.

levels) and heterogeneity in factor prices, that under the assumption of imposing the right demand elasticities and returns to scale, the TFPR is a good proxy for actual distortions, especially for those correlated with fundamentals.[17]

However, Table 1 also suggests that measurement error seems to be a relevant contributor to the dispersion of the TFPR.[18] To take into account measurement error, Bils, Klenow and Ruane (2018) propose a method to compute the true dispersion in TFPR in the presence of additive and orthogonal measurement error in revenues and inputs, using panel data. The methodology exploits the fact that in the absence of measurement error the elasticity of revenues with respect to inputs should not vary for plants with different average products; see section 4.2 for a detailed explanation. In what follows I use Bils, Klenow and Ruane's (2018) methodology to obtain measures of intra-industry misallocation that correct for measurement error, but, for tractability – and given the evidence cited above – I abstract from the sources related to model misspecification.[19]

More formally, assume that the production technology is Cobb-Douglas (CD) such that q units of variety m in a manufacturing industry s in country i are produced using a set of L homogenous factors z_l, physical productivity (TFPQ) a_m and factor intensities α_{ls}: $q_m = a_m \prod_l^L z_{lm}^{\alpha_{ls}}$ (I omit industry and country subscripts for firm-specific variables). Denote firms' revenue by r_m and the inverse of the constant mark-up by ρ. The sectoral production function is then: $Q_{is} = A_{is} \prod_l^L Z_{ils}^{\alpha_{ls}}$ (capital letters denote aggregates) where the sectoral TFP A_{is} depends on the distribution of physical productivities and the extent of intra-industry factor misallocation. In frictionless factor markets the (efficient) sectoral TFP is the power mean of firms' TFPQ, $(A_{is}^e)^{\sigma-1} = \sum_m a_m^{\sigma-1}$, and all firms face the same price for their homogenous inputs, say w_l for factor z_{lm}, leading to TFPR equalization across firms within industries, with values equal to $\frac{1}{\rho} \prod_l^L w_{il}^{\alpha_{ls}}$. Since the sectoral price index can be expressed as the ratio between the sectoral TFPR and the industry TFP, it can be in turn decomposed in terms of "natural" sources of comparative advantage and measures of factor misallocation as:

$$\ln P_{is} = \ln TFPR_{is} - \ln A_{is} = \sum_l^L \alpha_{ls} \left[\ln \left(1 + \bar{\theta}_{ils} \right) + \ln w_{il} \right] - \ln A_{is}^e - \ln AEM_{is} \qquad (2)$$

where $\left(1 + \bar{\theta}_{ils}\right)$ is defined as the ratio of the observed marginal revenue product (MRP) of factor l at the sector level, $\frac{\alpha_{ls} R_{is}}{Z_{ils}}$, to its return in the efficient allocation, $\frac{w_l}{\rho}$, that is:

[17]However, using the TFPR alone yields a substantial overstatement of the contribution of factor misallocation in accounting for firms growth volatility, a dimension that is not studied here.

[18]Further, Gollin and Udry (2019) recently show that measurement error can also importantly account for dispersion in TFPQ.

[19]Moreover, since in this paper I always exploit the variation in the misallocation measures across sectors, unless the causes related to misspecification have an heterogeneous impact across sectors, the obtained results are robust to this omission.

$\left(1+\bar{\theta}_{ils}\right) \equiv \frac{\rho\alpha_{ls}R_{is}}{w_l Z_{ils}}$; and AEM_{is} corresponds to the ratio sectoral TFP to the efficient one, $AEM_{is} \equiv A_{is}/A^e_{is}$. Those two ratios quantify the extent of resource misallocation. In the first case, the sectoral wedge $\left(1+\bar{\theta}_{ils}\right)$ characterizes the magnitude of inter-industry misallocation in factor l, and thus $\prod_l^L (1+\bar{\theta}_{ils})^{\alpha_{ls}}$ is a factor-intensity weighted measure of inter-industry misallocation.[20] In the second case, AEM_{is} characterizes the amount of within-industry factor misallocation, with $0 \leq AEM_{is} \leq 1$ and values closer to 1 reflecting less misallocation. According to the implications of the model, this measure is inversely related to the within-industry variance of the TFPR.[21] In Appendix C I offer more details about the relative importance of both types of factor misallocation in a closed economy.

Therefore, the decomposition in equation (2) reveals the theoretical determinants of the RCA measure under resource misallocation: i) the efficient TFP, A^e_{is}, which depends exclusively on the distribution of physical productivities across firms; ii) the geometric average of factor prices, $\prod_l^L w_{il}^{\alpha_{ls}}$, which in equilibrium can be recovered as the interaction between factor endowments and intensities;[22] iii) the geometric average of inter-industry wedges, $\prod_l^L (1+\bar{\theta}_{ils})^{\alpha_{ls}}$, a measure of inter-industry misallocation; and iv) the measure of intra-industry misallocation, AEM_{is}. Notice that, since the first component is related to technical efficiency and the second component to relative factor abundance, they represent the "Ricardian" and "Heckscher-Ohlin" sources of comparative advantage, respectively, whereas the two latter terms summarize both inter- and intra-industry resource misallocation. I use these four components (in logs) as explanatory variables in a regression of the RCA measure derived from the fixed effects, to test our hypothesis.

2.3 Relation between RCA and misallocation measures

Ideally, the suggested regression would require measures of the four variables in a large set of countries and industries, and thus comparable firm-level data for several countries. Given the infeasibility of this approach, I propose a two-stage strategy that exploits the time variation in the measures of RCA for Colombia relative to the United States (US) using panel-data. In the first stage, I estimate the panel data-version of equation (1), allowing the fixed effects in each cross section vary over time. That is, with data for the same set of countries in the

[20]The sectoral wedge $\left(1+\bar{\theta}_{ils}\right)$ can be also computed as the harmonic weighted average of analogue wedges at the firm-level, with weights given by firms' shares in sectoral revenue.

[21]In the case of a log-normal distribution of factor distortions across firms, the correlation is perfect. See Chen and Irarrazabal (2015) for the proof.

[22]Particularly if we set$w_l = \rho R / \sum_s \frac{Z_{ls}}{\alpha_{ls}}$ (where R is total revenue,$\sum_s R_s$), relative factor prices satisfy the equilibrium values for an allocative efficient closed economy, given by $\frac{w_l}{w_k} = \bar{Z}_k \sum_s \alpha_{ls}\beta_s / \bar{Z}_l \sum_s \alpha_{ks}\beta_s$ where $\bar{Z}_l$ is the total endowment of factor l (for a more details see Appendix C).

period 1991-1998, I run the regression:

$$\ln X_{ijst} = \delta_{ist} + \delta_{ijt} + \delta_{jst} + \varepsilon_{ijst} \tag{3}$$

where the exporter-industry-year fixed effect δ_{ist} identifies the triple difference of bilateral flows across exporters i and i', sectors s and s' and years t and t'; that is, the variation of RCA_{is} between time t and t', denoted by $dRCA_{ist}$. To compute $dRCA_{ist}$, instead of global means, I take as the reference country i' the US, the reference year t' the first year in the panel (1991), and the reference industry s' the sector with the median number of zeros bilateral flows in the data (footwear).[23] In the second stage, I regress the estimates of $dRCA_{ist}$ for Colombian industries on the four theoretical determinants of comparative advantage, constructed using micro-level data. Each variable is transformed to be expressed as the double difference first with respect to the reference industry and second with respect to the reference year, and then is normalized by the corresponding difference in the producer price index in the US (obtained from the NBER-CES manufacturing database), using the same industry and year of reference.[24]

The introduction of the time-dimension poses an additional challenge for the fixed effects estimators. Particularly, we must appraise the incidental parameter problem (Neyman and Scott (1948)), which generates an asymptotic bias for the fixed effects estimators when the number of time periods is small. Fernández-Val and Weidner (2016) prove that under exogenous regressors, in a Poisson model this bias is zero, which make PPML preferable over EK-Tobit as estimating method in the first stage. Thus, Table 2 displays the results for the standardized coefficients of the regression in the second stage, using PPML to obtain the exporter-industry-year fixed effects in the first stage. The estimation of the second stage is by weighted OLS, using the reciprocal of the error variance in the first stage as weighting matrix.[25] In the first column I present the results for the measure of intra-industry misallocation AEM_{is}, based on the direct measures of firms' TFPQ using plant-level deflators for firms'

[23]Therefore, letting $\hat{\delta}_{ist}$ be an estimate of δ_{ist} in the regression (3), $dRCA_{ist}$ of country i in sector s at time t is defined as:

$$dRCA_{ist} = \left[\frac{exp(\hat{\delta}_{ist})}{exp(\hat{\delta}_{is't})} / \frac{exp(\hat{\delta}_{i'st})}{exp(\hat{\delta}_{i's't})}\right] / \left[\frac{exp(\hat{\delta}_{is91})}{exp(\hat{\delta}_{is'91})} / \frac{exp(\hat{\delta}_{i's91})}{exp(\hat{\delta}_{i's'91})}\right]$$

where i' =US and s' = Footwear (7). As I show below, the results are not very sensitive to the choice of s'.

[24]This transformation intends to reflect the fact that the variation in RCA should be related to the change in the relative producer price indices compared to the same change in the country of reference: $dRCA_{ist} = F((\frac{P_{ist}}{P_{is't}} / \frac{P_{is0}}{P_{is'0}}) / (\frac{P_{i'st}}{P_{i's't}} / \frac{P_{i's0}}{P_{i's'0}}))$. Notice that in this approach we compare the growth on the relative prices (with respect to the reference year) across countries, so any difference in the measurement of relative prices across countries is absorbed by the difference over time.

[25]The use of weighted OLS seeks to alleviate the impossibility to bootstrap standard errors to account for the uncertainty in the estimation of the first stage. Given the high-dimensionality of the set of fixed effects involved in the non-linear regression by PPML in the first stage, the estimation is infeasible in standard econometric software as ®STATA, so I take advantage of the sparsity pattern of the problem and use a specialized solver that deals efficiently with sparse problems (SNOPT). However, the estimation is still highly time consuming, which makes infeasible to bootstrap standard errors.

inputs and outputs, that allows me to isolate the influence of demand shocks. In the second column, I use instead for the measure of intra-industry misallocation the within-industry variance of firms' TFPR, corrected by measurement error following Bils, Klenow and Ruane's (2018) methodology.

In both specifications, the measures of intra and inter-industry misallocation, once we control for the "natural" sources of export capability, are significantly correlated with our RCA measure and display the expected signs: positive for the intra-industry misallocation measure AEM_{is} (negative in the case of the within-industry variance of TFPR) and negative for the inter-industry misallocation measure $\prod_{l}^{L}\left(1+\bar{\theta}_{ils}\right)^{\alpha_{ls}}$. Moreover, the magnitude of the standardized coefficients suggests that both types of misallocation have a similar impact for shaping Colombian RCA, and they are not less important relative to the "Ricardian" and "Heckscher-Ohlin" determinants. These correlations are robust to the choice of the reference industry and the aggregation of countries. For instance, in column 3 I replicate the first specification using the sector with the lowest number of zeros as reference industry (machinery exc. electrical) whereas in column 4 I aggregate the 48 countries into 20 regions. The results are qualitatively similar. Therefore, the empirical evidence suggests that resource misallocation can play a role shaping the schedule of comparative advantage in Colombia. The model in the next section offers theoretical grounding to this insight.

3 A model of firm-level misallocation in an open economy

In this section, I introduce a model of international trade à la Melitz (2003) in which the allocation of factors within and across industries is inefficient. Next, I derive a theoretically consistent gravity equation following the lead of Arkolakis, Costinot and Rodríguez-Clare (2012) and Melitz and Redding (2014), assuming certain restrictions on the ex-ante joint distribution of TFPQ and factor distortions. Finally, I study the effects of both intra- and inter-industry factor misallocation on the reduced-form expression of the exporter-industry fixed effect derived from the gravity equation, my measure of RCA, using model simulations under a simple parametrization.

3.1 Model setup

Denote by m a single variety, i the exporting country, j the importing country, s an industry and l a homogenous production factor. Assume there are N possibly asymmetric countries, S industries and L homogenous primary factors. Hereafter capital letters denote aggregates, lower case letters firm-specific variables and for simplicity, I omit again sector subscripts for firm-specific variables. Each country i consumes according a two-tier utility function, with an upper-level CD with expenditure shares β_{is} across sectors and a lower-level CES with elasticity of substitution σ across varieties; let $\rho = \frac{\sigma-1}{\sigma}$. Each firm produces a variety m using

L homogenous primary factors (each one denoted by z_{ilm}) and a CD production technology with factor intensities α_{ls} (different factor intensities across industries, but equal for the same industry across countries). Firms are characterized by a Hicks-neutral physical productivity (TFPQ) a_{im} and a vector of L factor-distortions: $\vec{\theta}_{im} = \{\theta_{i1m}, \theta_{i2m}, ...\theta_{iLm}\}$, which are drawn from a joint ex-ante distribution $G_{is}\left(a, \vec{\theta}\right)$. There is a fixed cost of production f_{is} in terms of the composite input bundle, and each industry faces an exogenous probability of exit δ_{is}.

There is a fixed cost f^x_{ijs} to access market j from country i in sector s, defined in terms of the composite input bundle, and a transportation iceberg-type cost $\tau_{ijs} \geq 1$, with $\tau_{iis} = 1$. Let w_{il} denote the price of factor l in country i in absence of distortions, unobservable and common for all firms. Firms in country i face an idiosyncratic distortion θ_{ilm} (given by the l-th element of $\vec{\theta}_{im}$) in the market of primary factor l, such that the input price perceived by the firm is $(1 + \theta_{ilm})\, w_{il}$. Define $f_{ijs} = f^x_{ijs}$ if $j \neq i$; $f_{ijs} = f^x_{iis} + f_{is}$ otherwise (so domestic market fixed costs incorporates both "market access" and fixed production costs, whereas the export cost includes only the market access cost). The minimum "operational" cost to sell a variety m of country i in country j is:

$$c_{ijm}\left(q_{ijm}\right) = \omega_{is}\Theta_{im}\left(\frac{\tau_{ijs}q_{ijm}}{a_{im}} + f_{ijs}\right) \tag{4}$$

where $\Theta_{im} = \prod\limits_{l}^{L} (1 + \theta_{ilm})^{\alpha_{ls}}$ is a factor-intensity weighted geometric average of firm wedges and $\omega_{is} = \prod\limits_{l}^{L} (w_{il}/\alpha_{ls})^{\alpha_{ls}}$ is the prevalent factor price of the composite input bundle for the firms with zero draws of $\vec{\theta}_{im}$. Hereafter I refer to this cost as the total "operational" cost, which includes the variable cost of production and the fixed costs of production and delivery. Notice that this is a standard cost function in a multi-factor Melitz-type setting, the only difference here is that the composite input bundle's price perceived by the firm is a combination of both distortions and the underlying factor prices. Moreover, this cost function could be derived from a primal problem considering the following technology to produce and deliver one unit of variety m of country i in country j:

$$q_{ijm} = \frac{a_{im}}{\tau_{ijs}}\left(\prod_{l}^{L} z^{\alpha_{ls}}_{ijlm} - f_{ijs}\right) = \frac{a_{im}}{\tau_{ijs}}\left(z_{ijm} - f_{ijs}\right) \tag{5}$$

Here z_{ijlm} represents the total amount of primary factor l "embedded" in the production and delivery of variety m from country i in country j, and z_{ijm} the corresponding composite input bundle. Notice that z_{ijlm} includes the demand of primary factor l to pay both variable and fixed costs.

Profit maximization implies a firm charges a price p_{ijm} in each destination j equal to a fixed mark-up (ρ^{-1}) over its marginal cost: $p_{ijm} = \tau_{ijs}\Theta_{im}\omega_{is}/\rho a_{im}$. Quantities, revenues and profits of variety m from country i sold in country j are (respectively):

$$q_{ijm} = p_{ijm}^{-\sigma} E_{js} P_{js}^{d\sigma-1} \; ; \; r_{ijm} = p_{ijm}^{1-\sigma} E_{js} P_{js}^{d\sigma-1}; \; \breve{\pi}_{ijm} = \frac{1}{\sigma} r_{ijm} - \omega_{is}\Theta_{im} f_{ijs} \tag{6}$$

where E_{js} is the total expenditure of country j in varieties of industry s and P_{js}^{d} the corresponding consumer price index, variables that are defined below. It is straightforward to show the following relation between revenues from destination j and the corresponding total "operational" cost: $c_{ijm} = \rho r_{ijm} + \omega_{is}\Theta_{im} f_{ijs}$. Revenue productivity (TFPR) of selling variety m in destination j, denoted by ψ_{ijm}, is the ratio between revenue and the input used in production: $\psi_{ijm} \equiv r_{ijm}/\left(z_{ijm} - f_{ijs}\right) = p_{ijm} a_{im}/\tau_{ijs} = \Theta_{im}\omega_{is}/\rho$. Notice that although this destination-specific TFPR is not directly observable, since the allocation of factors to production for a given destination is unobservable, profit maximization implies that firms equate this value across all destinations, as the natural consequence of the absence of destination-specific frictions at the firm level. Hence, total TFPR must coincide with this value. In the absence of frictions in factor markets, there is TFPR equalization across firms within an industry (factor intensities make TFPR vary across sectors) for all destinations. Thus, in an efficient allocation, a firm's performance with respect to its competitors depends uniquely on relative TFPQ. In contrast, in the presence of factor misallocation, firms with higher TFPQ or lower TFPR (due to a low geometric average of firm wedges, Θ_{im}), holding the rest constant, set lower prices and hence sell higher quantities, obtaining higher revenues and profits in all markets.

Denote by ξ_{ijlm} the marginal revenue product (MRP) of factor l "embedded" in the production of variety m from country i to country j. Once again this MRP is not directly observable, but it is a useful concept to illustrate the consequences of factor misallocation. After some manipulation, it is possible to obtain the following relation between ξ_{ijlm} and the total "operational" cost: $\xi_{ijlm} = \alpha_{ls} c_{ijm}/\rho z_{ijlm}$. Notice that because of the presence of fixed costs, the MRP is no longer directly proportional to the average revenue product, a result emphasized in Bartelsman, Haltiwanger and Scarpetta (2013). From the FOC of the minimization cost problem of the firm, we know that $\left(1+\theta_{ilm}\right) w_{il} z_{ijlm} = \alpha_{ls} c_{ijm}$, which derives into $\xi_{ijlm} = \left(1+\theta_{ilm}\right)\omega_{is}/\rho$. That is, an efficient allocation of factors in an open economy requires MRP equalization across firms over all industries for all destinations, TFPR equalization within industries for all destinations,[26] but because of fixed costs, there is not average revenue products equalization.

Firms produce for a given destination only if they can make non-negative profits. Since profits in each market depend on both TFPQ and TFPR, this condition defines a cutoff frontier $a_{ijs}^{*}\left(\Theta\right)$ for each destination j, such that $\breve{\pi}_{ijs}\left(a_{ijs}^{*}\left(\Theta\right), \Theta\right) = 0 \; \forall \; i, j, s$. For a given combination of factor wedges Θ of firms in country i industry s, i.e., a given value of TFPR, $a_{ijs}^{*}\left(\Theta\right)$ indicates the minimum TFPQ required to earn non-negative profits in destination j

[26] Notice also that TFPR of variety m sold in destination j can be expressed as a factor-intensity weighted geometric average of the MRP: $\psi_{ijm} = \prod_{l}^{L} \left(\xi_{ijlm}/\alpha_{ls}\right)^{\alpha_{ls}}$.

. Define a^*_{ijs} as the TFPQ cutoff value for firms with TFPR equal to $\frac{\omega_{is}}{\rho}$ in destination j, i.e. firms with draws of distortions equal to zero: $a^*_{ijs} \equiv a^*_{ijs}(1)$. It is straightforward to derive the specific functional form of the cutoff functions in terms of a^*_{ijs} and Θ:

$$a^*_{ijs}(\Theta) = a^*_{ijs}\Theta^{\frac{1}{\rho}} \text{ with } a^*_{ijs} \equiv a^*_{ijs}(1) = \frac{\tau_{ijs}}{\rho}\left(\frac{E_{js}P_{js}^{\sigma-1}}{\sigma f_{ijs}}\right)^{\frac{1}{1-\sigma}} \omega_{is}^{\frac{1}{\rho}} \,\forall\, i, j, s. \tag{7}$$

The function $a^*_{ijs}(\Theta)$ is increasing in Θ (and thus in TFPR) reflecting the fact that larger wedges reflect higher marginal cost of the inputs, becoming more difficult to sell to the corresponding market. The existence of these cutoff functions, instead of unique threshold values for physical productivity, implies that the introduction of factor misallocation triggers selection effects that are absent in the efficient allocation. For example, some firms productive enough to operate in an undistorted counterfactual can no longer keep producing either because their distortions draws turn their profits negative or because even with a small "good" draw, the possible strengthening of competition due to the presence of highly positive distorted firms does not make it profitable for them to stay in the respective market. And the opposite could occur with some low productive firms, which will be able to survive in each market leading to misallocation of resources.[27]

To analyze the selection effects of resource misallocation, notice first that all cutoff functions across destinations share the same functional forms. Particularly, cutoff values for exporting to destination j are $\Lambda_{ijs} = \tau_{ijs}\left(E_{js}P_{js}^{d\sigma-1}f_{iis}/E_{is}P_{is}^{d\sigma-1}f_{ijs}\right)^{\frac{1}{1-\sigma}}$ times larger than domestic cutoff values. Thus, a simple representation of the firms in an open economy can be done in the space $a \times \Theta$, illustrated in Figure 2. In this space, each firm in sector s, characterized by a pair of draws (a, Θ), is represented by a single point. Profits are an increasing function of TFPQ and a decreasing function of TFPR, so firms with draws closer to the upper-left corner are more profitable. For simplicity, consider the destination j different to i with the lowest ratio Λ_{ijs} for country i in sector s in Panel A. Only firms with draws (a, Θ) above $a^*_{ijs}(\Theta)$ export to destination j, those with draws below $a^*_{ijs}(\Theta)$ and above $a^*_{iis}(\Theta)$ produce only for the domestic market, and those with draws below $a^*_{iis}(\Theta)$ do not produce. Panel B represents the selection mechanism that distortions trigger. Let $\tilde{a}^*_M$ represent the domestic productivity cutoff value in an allocative efficient economy (Melitz economy), and $\tilde{\Lambda}_{ijs}$ the corresponding value of Λ_{ijs}.[28] In such economy, firms with productivity above $\tilde{\Lambda}_{ijs}\tilde{a}^*_M$ export to j, those with productivity between $\tilde{\Lambda}_{ijs}a^*_M$ and a^*_M produce only for the domestic market, and those with productivity less than a^*_M do not produce. Thus, each cutoff function in the allocative inefficient economy creates two effects in the set of firms that sell to each market,

[27] These selection channels are also present in the closed economy models of Bartelsman, Haltiwanger and Scarpetta (2013) and Yang (2017).

[28] In general, a^*_{iis} and Λ_{ijs} are not related to $\tilde{a}^*_M$ and $\tilde{\Lambda}_{ijs}$ respectively. In Figure 2 it is arbitrarily assumed $a^*_{ijs} > \tilde{a}^*_M$.

which can be represented by two sets of areas: the regions under the density function that show firms that as consequence of distortions can no longer produce (light dotted area A) or export to j (light dotted area B) and the regions that display firms that because of distortions operate in the domestic market (dark dashed area A) or in the exporting market (dark dashed area B). The difference between dotted and dashed areas represents the net impact of distortions on the set of firms of country i and sector s, operating in the domestic and country-j markets (differences in A and B respectively).

The timing of information and decisions is as follows. Each time, there is an exogenous probability of exit given by d_{is}. A total of H_{is} potential entrants at country i industry s decide whether to produce and export to each destination conditional on their draws of physical productivity and distortions from G_{is}. All potential entrants pay a fee f_{is}^{e} to draw from G_{is}, which is paid in terms of the composite input bundle. The number of potential entrants is pinned down by the condition in which the expected discounted value of an entry is equal to the cost of entry. As usual in this kind of setup, let us consider no discounting and only stationary equilibria. Hence, the free entry condition is:

$$\sum_{j}^{N}\sum_{m}^{M_{ijs}}\breve{\pi}_{ijm} = \omega_{is} f_{is}^{e} H_{is} \; \forall \, i, s \tag{8}$$

Where M_{ijs} denotes the mass of operating firms in sector s of country i that is selling to country j. Aggregate stability requires that in each destination the mass of effective entrants is equal to the mass of exiting firms:

$$d_{is} M_{ijs} = \left[1 - G_{is}\left(a_{ijs}^{*}\left(\Theta\right), \Theta\right)\right] H_{is} \; \forall \, i, j, s \tag{9}$$

Given CES demand and firms prices, the consumer price index P_{is}^{d} in country i sector s satisfies $\left(P_{is}^{d}\right)^{1-\sigma} = \sum_{k}^{N} P_{kis}^{1-\sigma}$, with:

$$P_{ijs}^{1-\sigma} = \left(\frac{1}{\rho}\omega_{is}\tau_{ijs}\right)^{1-\sigma}\sum_{m}^{M_{ijs}}\left(\frac{a_{im}}{\Theta_{im}}\right)^{\sigma-1} \tag{10}$$

Total expenditure in country i and sector s is $E_{is} = P_{is}^{d} Q_{is}^{d}$. By the upper-level utility function, the overall consumer price index (equal to unit expenditure) is $P_{i}^{d} = \prod_{s}^{S}\left(P_{is}^{d}/\beta_{s}\right)^{\beta_{s}}$ and satisfies $E_{is} = \beta_{s} E_{i}$, with $E_{i} = \sum_{s}^{S} E_{is}$ total country-i expenditure.

Now consider the aggregate variables. Let $X_{ijs} = \sum_{m}^{M_{ijs}} r_{ijm}$ be the value of total exports from country i to destination j in industry s. Analogously as at the firm-level, the total "operational" cost of exporting to country j incurred by all firms of country i in industry s can be written as $C_{ijs} = \rho X_{ijs} + \mathfrak{F}_{ijs}$ where $\mathfrak{F}_{ijs} = \sum_{m}^{M_{ijs}} \omega_{is}\Theta_{im} f_{ijs}$ is the value of total expenditures in fixed costs. Similarly, denote by R_{is}, $\mathfrak{F}_{is}$, C_{is} the same aggregations but at the industry level, with $R_i = \sum_s^S R_{is}$ representing total country i's gross output. Denote the HWA of primary factor-l wedges $(1+\theta_l)$ within industry s as $\left(1+\bar{\theta}_{ils}\right)$, with weights given by the firm's participation in C_{is}. It is possible to show that $\left(1+\bar{\theta}_{ils}\right) = (\rho R_{is} + \mathfrak{F}_{is})\,\alpha_{ls}/w_{il}Z^o_{ils}$ where Z^o_{ils} is the aggregate demand of factor l for "operational" uses in country i in sector s: $Z^o_{ils} \equiv \sum_j^N \sum_m^{M_{ijs}} z_{ijlm}$. Thus, this average wedge is the industry-level analogue of firm-level wedges and allows me to measure the degree of inter-industry misallocation, as in the closed-economy framework of the previous section. The total demand of primary factor l for "operational" uses in country i industry s can be expressed as:

$$Z^o_{ils} = \frac{\alpha_{ls} C_{is}}{w_{il}\left(1+\bar{\theta}_{ils}\right)} \tag{11}$$

Primary factors are used for "operational" (fixed and variable costs) and investment (entry) costs. The sectoral demand of the composite input bundle for entry costs is simply $f^e_{is}H_{is}$. Therefore, the amount of primary factor l allocated to entry costs in country i sector s is $Z^e_{ils} = \alpha_{ls}\omega_{is}f^e_{is}H_{is}/w_{il}$, and the total allocation of the same factor, Z_{ils}, is given by:

$$Z_{ils} = Z^o_{ils} + Z^e_{ils} = \frac{\alpha_{ls} C_{is}}{w_{il}\left(1+\bar{\theta}_{ils}\right)} + \frac{\alpha_{ls}\omega_{is}f^e_{is}H_{is}}{w_{il}} \tag{12}$$

Notice that the inter-industry wedge only appears in the input allocated for operational uses. This is a consequence of the timing of the model, in which firms allocate first real resources (the entry fixed cost) to draw from the joint distribution. Only after this moment is the draw of the vector of distortions known to the firm. Factor-l market clearing condition in country i is then:

$$\bar{Z}_{il} = \sum_s^S Z_{ils} \tag{13}$$

where $\bar{Z}_{il}$ is the total endowment of primary factor l in country i, and Z_{ils} is given by (12). Finally, the balanced trade condition requires equalization of the total revenues to total expenditures plus aggregate deficits:[29]

$$R_i = E_i + D_i \tag{14}$$

[29] By construction, total revenues are the sum of factor payments and profits: $R_i = \sum_s^S \sum_l^L (1+\bar{\theta}_{ils}) w_{il} Z^o_{ils} + \sum_s^S \omega_{is} f^e_{is} H_{is}$. This can be shown decomposing sectoral revenues as:

$$R_{is} = \rho R_{is} + \frac{1}{\sigma} R_{is} = \sum_l^L (1+\bar{\theta}_{ls}) w_{il} Z^o_{ils} - \mathfrak{F}_{is} + \sum_j^N \sum_m^{M_{ijs}} (\pi_{ijms} + \omega_{is}\Omega_{im} f_{ijs})$$

where the second equality is derived from (11) and the aggregation of firms' revenues.

where D_i is the country's trade balance (a positive value means surplus), an exogenous value in the model. Global trade balance requires: $\sum_i^N D_i = 0$. A summary of the whole system of equations and unknowns is given in Table 3. This table also offers the dimensionality of the problem.

3.2 Comparative advantage

Bilateral exports at the industry level can be expressed in terms of sectoral expenditures in the importer country (E_{js}) and trade shares of the importer country (π_{ijs}). The latter term can be re-written in terms of the bilateral price indices as:

$$X_{ijs} = \pi_{ijs} E_{js} = (\frac{P_{ijs}^{1-\sigma}}{\sum_k^N P_{kjs}^{1-\sigma}}) E_{js} \tag{15}$$

The trade share of country i in country-j expenditures in goods of industry s only depends on the value of its bilateral price index P_{ijs}, relative to the same value for all competitors of country i in such market. As I commented earlier, this is so because the price index P_{ijs} is a measure of the unit price incurred by consumers of the destination country, and hence it is an indicator of country-i's competitiveness. To derive the reduced-form of the exporter industry fixed effect, consider the double difference of bilateral flows across exporters i and i' and sectors s and s' for a given importer j, i.e., $\frac{X_{ijs}X_{i'js'}}{X_{ijs'}X_{i'js}}$. It is straightforward to see that this double difference is given by the difference in the relative price index, $(\frac{P_{ijs}P_{i'js'}}{P_{ijs'}P_{i'js}})^{1-\sigma}$. From (10) it is possible to disentangle these bilateral prices indices as follows:

$$P_{ijs} = \tau_{ijs} M_{ijs}^{\frac{1}{1-\sigma}} \frac{\bar{\psi}_{ijs}}{A_{ijs}} \tag{16}$$

where A_{ijs} and $\bar{\psi}_{ijs}$ are the industry-destination analogues of sectoral TFP and sectoral revenue productivity respectively,[30] so A_{ijs} represents the overall efficiency of exporting firms to destination j and $\bar{\psi}_{is}$ depicts the average cost of the factors faced by the same set of exporters. Therefore, equation (16) disentangles the four determinants of exporters' competitiveness: i) their overall efficiency, which is a weighted average of exporters physical productivity and factor market frictions; ii) the average cost of factors for exporters; iii) the mass of exported varieties; and iv) bilateral trade costs. Of these components, factor misallocation has a direct impact on the average TFP and an indirect impact (through general equilibrium channels) on the formation of factor prices and the determination of the number of exported varieties. Notice also that the unit price is a combination of both extensive and intensive margins of

[30]This is: $A_{ijs} = \bar{\Theta}_{ijs}(\frac{1}{M_{ijs}} \sum_m^{M_{ijs}} (\frac{a_{im}}{\Theta_{im}})^{\sigma-1})^{\frac{1}{\sigma-1}}$ and $\bar{\psi}_{ijs} = \frac{\omega_{is}\bar{\Theta}_{ijs}}{\rho}$, where $\bar{\Theta}_{ijs} = \prod_l^L (1+\bar{\theta}_{ijls})^{\alpha_{ls}}$. Here $(1+\bar{\theta}_{ijls})$ denote the HWA of factor-l wedges of firms exporting to destination j in industry s, with weights given by firm's participation in the total cost of factors C_{ijs}.

trade. Thus, the model is very rich about the determinants of competitiveness. It is able to combine the sources of relative export capability in Ricardian and Heckscher-Ohlin models (where comparative advantage is due to differences in efficiency across industries in the first case and the interaction between the sizes of factor endowments and factor intensities across industries that pins down relative factor prices, in the second case) with the motives for intra-industry trade in monopolistic competition models with Dixit-Stiglitz preferences (where the gains-from-variety effect induce reductions in unit costs) in an environment of micro-level resource misallocation, which in turn can also create "artificial" comparative advantage. In the next subsection, I perform numerical simulations to disentangle the effects of both intra- and inter-industry misallocation on each component of the relative unit prices.

At this point I need to impose a functional form for the joint distribution G_{is} to derive the reduced-form equation of the exporter-industry fixed effect from the double difference in unit price. Let $G^a_{is}(a)$ be the univariate margin of G_{is} with respect to a, and $G^{\theta}_{is}(\vec{\theta})$ the multivariate margin of G_{is} with respect to $\vec{\theta}$.[31] Consider the following assumptions:

A. 1. *(Pareto distribution)* $\forall a_i > \bar{a},\ G^a_{is}(a) = 1 - (\frac{\bar{a}_{is}}{a})^{\kappa};\ \kappa > \sigma - 1;$

A. 2. *(Ex-ante independence)* $G_{is} = G_{is}(a, \vec{\theta}) = G^a_{is}(a) G^{\theta}_{is}(\vec{\theta})$

First, regarding Assumption A.1., the Pareto distribution is the common benchmark in the trade literature to model heterogeneity on physical productivity in the Melitz model. Not only does it have a good empirical performance approximating the observed distribution of firm size,[32] but it also makes the model analytically tractable, allowing me to derive a particular expression for the gravity equation. And second, although Assumption A.2. seems problematic given the observed correlation between TFPQ and TFPR in the data, it is worth emphasizing that the assumed independence is only between the latent (ex-ante) marginal distribution of TFPQ and that of the vector of factor distortions. The observed (ex-post) distribution can exhibit any kind of correlation. In fact, given the functional forms of the cutoff functions, endogenous selection in the model implies the positive ex-post correlations between TFPQ and TFPR observed in the data. Furthermore, there is no restriction for the joint distribution of individual factor distortions G^{θ}_{is}, so covariances across factors wedges are completely allowed. I keep Assumptions A.1. and A.2. hereafter unless otherwise indicated.

Under Assumptions A.1. and A.2., the model exhibits an interesting set of features and offers a great simplification, which is done in detail in Appendix B.1 and summarized by the system of equations (21)-(24) below. First, it is possible to show that the property of a constant aggregate profits/revenue ratio of the Pareto-Melitz model still holds under factor misallocation: $R_{is} = \frac{\kappa}{\rho}\Pi_{is} = \frac{\kappa}{\rho}\omega_{is} f^e_{is} H_{is}$ (see equation (B.4) in Appendix B.1). Thus, market

[31] This is, $G^a_{is}(a) = \underset{\vec{\theta}\to\vec{\infty}}{lim}\ G_{is}(a, \vec{\theta})$ and $G^{\theta}_{is}(\vec{\theta}) = \underset{a\to\infty}{lim}\ G_{is}(a, \vec{\theta})$

[32] See for example Cabral and Mata (2003) or Axtell (2001).

clearing conditions can be re-stated as:

$$w_{il}Z_{ils} = \alpha_{ls}\left[\left(1+\bar{\theta}_{ils}\right)^{-1}\left(1-\frac{\rho}{\kappa}\right)+\frac{\rho}{\kappa}\right]R_{is} \tag{17}$$

notice that the HWA wedge $\left(1+\bar{\theta}_{ils}\right)$ affects only the fraction of the total revenue that is allocated to "operational" costs: $1-\frac{\rho}{\kappa}$. Denote the term in curly brackets by v_{ils}. Here, v_{ils} measures the effective extent of inter-industry misallocation for primary factor l, considering all its possible uses (operational and entry costs). Let v_{is} denote the factor-intensity weighted geometric average of these measures: $v_{is} = \prod_{l}^{L} v_{ils}^{\alpha_{ls}}$. Further, aggregate the sectoral demands of primary factors on an industry-level composite input bundle $Z_{is} = \prod_{l}^{L} Z_{ils}^{\alpha_{ls}}$. Thus, we can state $v_{is}R_{is} = \omega_{is}Z_{is}$ and hence $H_{is} = \frac{\rho Z_{is}}{\kappa f_{is}^{e} v_{is}}$, a solution for the mass of entrants similar to that obtained in the multi-sector Pareto-Melitz case (in which the mass of entrants is related to the total allocation of inputs in the sector). The only difference here is the presence of the inter-industry allocative inefficiency measure v_{is}, which affects the total allocation of factors across sectors.

Second, it is possible to derive a relationship between the ex-post HWA wedge and the ex-ante joint distribution of distortions. Appendix B.2 shows that the following relation holds:

$$(1+\bar{\theta}_{ils}) = \frac{\Gamma_{is}}{\Gamma_{ils}} \tag{18}$$

where $\Gamma_{is} = \int_{\theta_{i1}} ... \int_{\theta_{iL}} \Theta_i^{1-\frac{\kappa}{\rho}} dG_{is}^{\theta}(\vec{\theta})$ and $\Gamma_{ils} = \int_{\theta_i} ... \int_{\theta_{iL}} \frac{\Theta_i^{1-\frac{\kappa}{\rho}}}{(1+\theta_{il})} dG_{is}^{\theta}(\vec{\theta})$, terms that only depend on the ex-ante joint distribution of firm-level distortions G_{is}^{θ}. Equation (18) makes evident the interaction between both types of factor misallocation under our assumptions, and depending on the parametric assumptions on the joint distribution G_{is}^{θ}, it allows me to recover some structural parameters from the values of observed HWA wedges.

Third, regarding the gravity equation, I show in Appendix B.3 that relative bilateral exports can be expressed as:

$$\ln\left(\frac{X_{ijs}X_{i'js'}}{X_{ijs'}X_{i'js}}\right) = \ln\left[\frac{\varrho_{is}\varrho_{i's'}}{\varrho_{is'}\varrho_{i's}} \frac{\Gamma_{is}\Gamma_{i's'}}{\Gamma_{is'}\Gamma_{i's}} \frac{R_{is}R_{i's'}}{R_{is'}R_{i's}} \left(\frac{\omega_{is}\omega_{i's'}}{\omega_{is'}\omega_{i's}}\right)^{-\frac{\kappa}{\rho}}\right] + B_{ijs} \tag{19}$$

where B_{ijs} and ϱ_{is} are constants that do not vary when we remove misallocation. The first term of the RHS of equation (19) is what δ_{is} identifies in the regression with fixed effects in (1). I show in Appendix B.3 how it can be decomposed in elements that capture the influence of each source of export capability in the model. Moreover, notice that changes in the extent of allocative inefficiency have a direct effect on the double difference of the term Γ_{is}, and an indirect effect (through general equilibrium channels) on the product of the double differences of the terms R_{is} and $\omega_{is}^{-\frac{\kappa}{\rho}}$. Thus, to figure out the total impact of factor misallocation on

RCA, it is necessary to solve the full model in general equilibrium, which is done in section 4.

3.3 Simulations

To illustrate the effects of both intra- and inter-industry misallocation on comparative advantage, I use numerical simulations under a simple parametrization of the model. Consider a world with two countries, two factors and two sectors, with symmetric factor intensities across sectors. Sector 1 is factor 1-intensive. Country 1 faces factor misallocation in sector 1 (I will simulate distortions on each factor, so the results are totally symmetric for factor misallocation in sector 2). Assume trade costs do not vary across sectors. Two objectives are pursued: first, to show how both types of factor misallocation of country 1 affect its comparative advantage, disentangling the total impact on its determinants; and second, to illustrate how sensitive these effects are to factor intensities and trade costs.

Both sectors in the two countries have the same Pareto TFPQ distribution. Country 1 is relatively abundant in factor 1 with respect to country 2, so in the allocative efficient scenario it has a comparative advantage in sector 1.[33] I am interested in the RCA of country 1 in sector 1 relative to country 2 in sector 2, which I compute using equation (19). Assume also a log-normal distribution for distortions, with location and shape parameters μ_{l1} and σ^2_{l1} for factor l respectively, and to simplify things, zero covariances. I show in Appendix B.4 that using equation (18) under log-normality it is possible to obtain the following relation between the ex-post HWA wedge and those parameters:

$$\ln\left(1+\bar{\theta}_{ils}\right) = \mu_{ils} + \left[\left(1-\frac{\kappa}{\rho}\right)\alpha_{ls} - \frac{1}{2}\right]\sigma^2_{ils} \tag{20}$$

Equation (20) sheds light on the feedbacks between the two types of factor misallocation under endogenous selection of firms. For example, consider the case in which the location parameter is zero. Ex-ante, the average (log) distortion for the firms within the industry is zero. However, for a given value of the dispersion on these frictions (which generates intra-industry misallocation) we obtain $\left(1+\bar{\theta}_{ils}\right) < 1$; that is, ex-post inter-industry misallocation. This result is due to endogenous selection, since firms with both low TFPQ and high distortions exit for sure, pushing the value of the ex-post average of the prevalent distortions below zero, generating inter-industry misallocation.

Only intra-industry misallocation

To represent the impact of only intra-industry misallocation on comparative advantage, I first consider the impact of an increase in the variance of wedges of each factor separately,

[33] Results do not change qualitatively in the case of the opposite relative factor endowments, or if the comparative advantage is countered or enhanced by Ricardian comparative advantage (through differences in the lower bound of the Pareto distribution). In those cases, there is a change in the initial RCA, but the effect of factor misallocation is qualitatively similar.

simultaneously adjusting the location parameter to ensure there is no inter-industry misallocation. Figure 3 displays the results. The first four graphs correspond to the total impact on the comparative advantage of sector 1 (first graph) and the decomposition of the sources of export capability explained above (average efficiency, returns of factors, and number of the mass of exported varieties; second to fourth graphs), following equation (B.8) in Appendix B.3. Each of these graphs plots the difference between the value of the endogenous variable under the parameters assumed for the distribution of distortions, which are displayed in the last graph, and the corresponding values in the allocative efficient equilibrium, so they capture the net effect of the considered allocative inefficiency. The fifth graph illustrates the implicit HWA of the prevalent distortions, following equation (20), to verify the degree of inter-industry misallocation. Blue and red lines correspond to misallocation only in factors 1 and 2, respectively. I consider two trade regimes: free trade, represented by dashed lines,[34] and costly trade, represented by continuous lines. The values for the whole set of parameters used in each simulation are displayed in Table 4.

Introducing only intra-industry misallocation of any factor used in sector 1 reduces its comparative advantage. The effect increases the larger the variance of the (log) wedges and, for the same value of the variance, if the misallocation affects the factor used intensively by industry. The total effect is also marginally larger under free trade for the range of variances considered in the graph. It is worth saying that for larger variances, there is a threshold in which with free trade the system falls in a regime of complete specialization, so the production of sector 1 shuts down. These results are consistent with the intuition that the larger the possibility to substitute goods across countries, the larger the impact of misallocation on industry revenue shares, boosting more reallocation of factors across sectors. Regarding the determinants of relative export capability, intra-industry misallocation creates well-known losses of TFP, as in a closed economy. However, to keep trade balanced, these losses are followed by an adjustment in relative factor prices, absent under autarky. Given endogenous selection, there is relative net exit of exporters in the distorted sector 1, which is a consequence of the reallocation of factors to the undistorted sector 2. The increase in the relative demand of the factor used intensively in sector 2 also reduces the relative price of the factor used intensively in sector 1. The combined effect on factor prices largely counters the effect of the loss in overall efficiency, but the sum of the two forces is still negative. Thus, the total impact on export capability is largely due to the adjustment in the extensive margin of trade, whereas the contribution of the intensive margin is smaller, but not zero.[35]

[34]For free trade I will consider an scenario without iceberg transportation costs but with fixed costs of exporting, since I am interested in keeping endogenous selection on exporting markets.

[35]The prevalence of the extensive margin is probably linked to the Pareto assumption. On the consequences on Pareto's distribution over the two margins of trade, see Fernandes et al. (2018).

Only inter-industry misallocation

Now consider the impact of inter-industry misallocation. For this, I shift the location parameter allowing it to take positive and negative values, keeping the shape parameter equal to zero. Then, there is no dispersion in wedges (and thus no intra-industry misallocation), but the ex-post HWA wedge varies with the location parameter, creating inter-industry misallocation. Figure (4) displays the results with the same graphs and conventions as in the previous exercise. The net impact on comparative advantage is inversely related to the sign on the location parameter. To understand this result, it is useful to think about positive values of the location parameter as an industry-level tax in the cost of the factor, which imply a HWA wedge greater than 1 (or a subsidy for negative values). For instance, consider the effects of introducing an industry-level factor tax. It becomes relatively more expensive to buy the corresponding input for all firms within the taxed industry, raising the average return of the composite input bundle. Some firms whose productivity draws prevent them from paying the new inputs' cost must exit. Here, there is no TFP loss due to within-industry misallocation, because all firms in the industry face the same factor prices, so average TFP depends only on the physical productivities of the incumbents. Instead, there is selection of the more productive firms, so average TFP rises. Both impacts are larger if the taxed factor is the one used intensively in the sector (since it has more weight in the composite bundle) and under free trade (since reallocation of factors is larger). The increase on average TFP entirely compensates the loss in export capability due to the increase in the relative return of the factors, up to the point that net effect on comparative advantage through the intensive margin is positive, but small. Adding the negative effect on the extensive margin due to the exit of firms, which is not very affected by the trade regime or by the intensity in the use of the factors, the overall impact on export capability is negative.

In conclusion, each type of factor misallocation impacts industries' comparative advantage through different general equilibrium channels. The extent of each impact depends on the interaction between factor intensities and the variances of distortions, in the case of intra-industry misallocation, and primarily on whether the HWA wedges are less or greater than one, in the case of inter-industry misallocation. The effect of both types of factor misallocation on the industries' TFP is partially offset by changes in relative factor prices, so the intensive margin contributes less to the adjustment of relative unit prices relative to the extensive margin (the change in the mass of produced varieties due to the reallocation of factors across industries). Therefore, ignoring the general equilibrium effects caused by resource misallocation could lead to misguided conclusions. The next section presents a methodology to solve the model in general equilibrium to produce a counterfactual series of bilateral exports after removing allocative inefficiency in a country, and hence to evaluate its frictionless RCA.

4 Empirical implementation

In this section, I perform the counterfactual exercise of removing both (and separately) the observed intra and inter-industry misallocation in Colombia. I first show how to obtain the counterfactual equilibrium solving the model in relative changes. Next, I comment on the data employed, the method to measure the dispersion in the MRP of the factors under overhead costs, and the baseline results. Finally, I conduct some robustness checks and compare the baseline results with those obtained for the one-sector economy and the closed economy.

4.1 Counterfactual exercise

I show in Appendix B.1 that under assumptions A.1. and A.2. the entire system can be solved in terms of the following system of equations:

$$w_{il} Z_{ils} = \alpha_{ls} v_{ils} R_{is} \tag{21}$$

$$\bar{Z}_{il} = \sum_{s}^{S} Z_{ils} \tag{22}$$

$$R_{is} = \sum_{j}^{N} \pi_{ijs} \beta_{js} \left(\sum_{s}^{S} R_{js} - D_j \right) \tag{23}$$

$$\pi_{ijs} = \frac{\left(\prod_{l}^{L} w_{il}^{\frac{-\kappa \alpha_{ls}}{\rho}} \right) \Gamma_{is} \phi_{ijs} R_{is}}{\sum_{k}^{N} \left(\prod_{l}^{L} w_{kl}^{\frac{-\kappa \alpha_{ls}}{\rho}} \right) \Gamma_{ks} \phi_{kjs} R_{ks}} \tag{24}$$

where $\phi_{ijs} = \frac{f_{ijs}^{\frac{\sigma-1-\kappa}{\sigma-1}} \bar{a}_{is}^{\kappa}}{(\tau_{ijs})^{\kappa} f_{is}^{e} d_{is}}$ and π_{ijs} is the share of country i in total expenditures of country j in sector s. Denote the share of factor l allocated to sector s in country i as $\tilde{Z}_{ils}$, that is: $\tilde{Z}_{ils} \equiv \frac{Z_{ils}}{\bar{Z}_{il}}$. Equations (21) and (22) can be re-stated as: $w_{il} \tilde{Z}_{ils} \bar{Z}_{il} = \alpha_{ls} v_{ils} R_{is}$, with the condition $\sum_{s}^{S} \tilde{Z}_{ils} = 1 \ \forall \ i, l$.

Now I use the methodology of Dekle, Eaton and Kortum (2008), adopted in other papers,[36] to obtain the counterfactual equilibrium in relative changes. This approach, known as exact hat algebra, allows me to solve the model without assuming or estimating parameters that are hard to identify in the data, particularly all those which are embedded in the term ϕ_{ijs} (trade variable and fixed costs, entry costs, lower bounds for TFPQ, probabilities of exit), and the current measures of intra-industry and inter-industry misallocation for all industries and countries. All these values are included in the initial trade shares, and because they do not change in the counterfactual equilibrium, they do not appear in the system in relative changes.

[36] See for example Costinot and Rodríguez-Clare (2014), Caliendo and Parro (2015), Święcki (2017), among others.

For any variable x in the initial equilibrium denote x' its counterfactual value and $\hat{x} \equiv \frac{x'}{x}$ the proportional change. Then, the system in the final equilibrium can be rewritten as:

$$\hat{w}_{il} = \sum_{s}^{S} \tilde{Z}_{ils} \hat{R}_{is} \hat{v}_{ils} \tag{25}$$

$$R_{is}\hat{R}_{is} = \sum_{j}^{N} \pi'_{ijs} \beta_{js} \left(\sum_{s}^{S} R_{js}\hat{R}_{js} - D_j \hat{D}_j \right) \tag{26}$$

$$\pi'_{ijs} = \frac{\pi_{ijs} \left(\prod_{l}^{L} \hat{w}_{il}^{\frac{-\kappa \alpha_{ls}}{\rho}} \right) \hat{\Gamma}_{is} \hat{R}_{is}}{\sum_{k}^{N} \pi_{kjs} \left(\prod_{l}^{L} \hat{w}_{kl}^{\frac{-\kappa \alpha_{ls}}{\rho}} \right) \hat{\Gamma}_{ks} \hat{R}_{ks}} \tag{27}$$

The objective with this system is to analyze the impact of exogenous changes in both intra and inter-industry misallocation (through the terms $\hat{v}_{ils}$ and $\hat{\Gamma}_{is}$) of a country on the equilibrium outcomes $\hat{R}_{is}$ and $\hat{w}_{il}$. For this, the system can be solved for $\hat{R}_{is}$ and $\hat{w}_{il}$ (after imposing the usual normalization $\sum_{i}^{N} R_{is}\hat{R}_{is} = 1$) given values of the observable variables π_{ijs}, $\tilde{Z}_{ils}$ and R_{is}, technological and preference parameters α_{ls} and β_{is} respectively, and assumptions on parameters κ and σ and the variation of aggregate trade deficits $\hat{D}_j$. Since my interest is to remove factor misallocation only in a country, I set $\hat{v}_{ils} = \hat{\Gamma}_{is} = 1$ for all countries different from Colombia, so I only need values of v_{ils} of Γ_{is} for Colombia to derive the corresponding proportional changes.

Once $\hat{R}_{is}$ and $\hat{w}_{il}$ are obtained, it is straightforward to compute the relative changes in aggregate expenditure and trade shares, $\hat{E}_i$ and $\hat{\pi}_{ijs}$. With these variables it is possible to quantify the cost of each type of misallocation in terms of welfare, measured as total real expenditure. In Appendix B.5 I show that the relative change in aggregate real expenditure can be derived from:

$$\frac{\hat{E}_i}{\hat{P}_i^d} = \prod_{s}^{S} \left[\hat{E}_i^{\frac{1}{\kappa} - \frac{1}{\rho}} \left(\frac{\hat{\pi}_{iis}}{\hat{R}_{is} \hat{\Gamma}_{is}} \right)^{\frac{1}{\kappa}} \prod_{l}^{L} \hat{w}_{il}^{\frac{\alpha_{ls}}{\rho}} \right]^{-\beta_s} \tag{28}$$

Notice that in the case of the undistorted economy with one factor of production, equation (28) collapses to the well-known Arkolakis, Costinot and Rodríguez-Clare's (2012) formula ($\prod_{s}^{S} \left[\frac{\hat{\pi}_{iis}}{\hat{Z}_{is}} \right]^{-\frac{\beta_s}{\kappa}}$) to evaluate the increase in welfare in response to any exogenous shock.

4.2 Data and model solution

I collect information on bilateral trade shares, gross output and sectoral factor shares for the same set of countries and manufacturing sectors used in section 2. I use a gross output specification for the production function with capital, materials, skilled and unskilled labor as inputs. I set factor intensities for all countries equal to the US cost shares, under the assumption that US cost shares reflect actual differences in technology across sectors instead of inter-industry misallocation. The primary source of information is the OECD's Trade in Value Added (TiVA) database (2015's release) for the year 1995, but I also use auxiliary information from several other sources; for a detailed description see Appendix A.1. For the calibrated parameters, I use in the baseline results $\kappa = 4.56$ and $\sigma = 3.5$, values consistent with those used in the literature.[37] Section 4.4 verifies how sensitive are the results to changes in those values. Given the static nature of the framework, the model is silent about the adjustment of aggregate trade deficits. Thus, for the counterfactual exercises, I assume that for all countries different from the RoW, trade deficits as a proportion of gross output remain constant in the counterfactual. The trade deficit of the RoW adjusts to ensure global trade balance.

To obtain the proportional changes in the measures of factor misallocation $\hat{v}_{ils}$ and $\hat{\Gamma}_{is}$ for Colombia, I assume that the joint distribution of factor distortions is log-normal. In Appendix B.4 I show how equation (18) can be used to obtain an identity that relates the ex-post HWA wedges to the vector of location parameters and the variance-covariance matrix of the ex-ante joint distribution of the distortions V_{is} (see equation (B.9)). Therefore, I only need measures of the HWA of wedges, which can be inferred from sectoral data using (17), and estimates of V_{is} to obtain the latent location parameters and, consequently, both v_{ils} and Γ_{is}. The counterfactual exercises involve removing: i) both types of misallocation; ii) only intra-; and iii) only inter-industry misallocation for the homogenous production factors: capital, skilled and unskilled labor.[38]

To estimate V_{is}, I use Bils, Klenow and Ruane's (2018) method to compute the dispersion in the factors' MRP in the presence of additive measurement error in revenue and inputs. Since overhead factors are analogous to an unobservable additive term in measured inputs, this approach deals also with the problem of inferring the variance of factors' MRP directly from the observed dispersion of the average revenue products in the presence of fixed costs. The main idea of Bils, Klenow and Ruane's (2018) approach is to estimate a "compression factor" $\hat{\lambda}$ to correct the observed dispersion on TFPR, $\hat{\sigma}^2_{TFPR}$, as a measure of the dispersion in the "true" TFPR, σ^2_{TFPR} ($\hat{\lambda} = \sigma^2_{TFPR}/\hat{\sigma}^2_{TFPR}$), using panel data. The methodology exploits the fact that in the absence of measurement error the elasticity of revenues with respect to inputs should not vary for plants with different average products. Hence, panel data can be used to back out the "true" marginal product dispersion by estimating how such

[37]These values are averages of the ones used by Melitz and Redding (2015) ($\kappa = 4.25$ and $\sigma = 4$) and the ones estimated by Eaton, Kortum and Kramarz (2011) ($\kappa = 4.87$ and $\sigma = 2.98$). Section 4.4 evaluates the sensitivity of the baseline results to changes in these values.

[38]Given the infeasibility of decomposing intermediate consumption into homogeneous inputs, I assume that all observed dispersion in the MRP of materials is due to actual heterogeneity in the input, instead of factor misallocation. Thus, the counterfactual equilibrium preserves both the observed within-industry dispersion and the inter-industry differences in the MRP of intermediate consumption.

elasticity changes for plants with different average products. I estimate $\hat{\lambda}$ by GMM sector by sector, using the panel data from 1991 to 1998. In Appendix A.2, I present details about the methodology and the results of the replication.[39] I correct the observed variance-covariance matrix of the average revenue products of factors by $\hat{\lambda}_s$ to obtain $\hat{V}_{is}$. Table 5 displays for each industry the employed values for the HWA wedges, the corresponding corrected variances and covariances of factors' average revenue products and the "compressions factors" $\hat{\lambda}_s$ used, along with factor intensities. In Appendix A.3 I exemplify how the resulting variances have a relation with possible sources of factor misallocation, as a way to externally validate their relevance. I show that dispersion of capital wedges are related to the intra-industry dispersion of the idiosyncratic cost of capital for firms within the same industry, measured by an appropriate weighted average of the interest rates of their loans.

Finally, the model is constituted by $N \times (S + L) = 1344$ equations. The multiplicity of non-linearities in the model implies that common optimization routines find multiple local solutions. To obtain the global solution, I employ both an algorithm to choose a set of ideal initial conditions and a state-of-the-art solver for large-scale nonlinear systems. Appendix A.4 offers details about these two aspects.

4.3 Baseline results

First, I describe the results of "extreme" reforms that remove the total extent of intra- and inter-industry misallocation in Colombia. The results of gradual reforms are presented in the next section. I compute the RCA measures for each counterfactual equilibrium using PPML. Similar to Figure 1, instead of choosing a pair importer-sector, I normalize by global means. The resulting RCA measures are displayed in Figure 5. All panels plot the actual RCA measures in the horizontal axis and the counterfactuals in the vertical one. Panels A and B show the case of removing both types of misallocation. In Panel A the markers' sizes represent the actual industries' export shares and in Panel B the counterfactual ones.

Once both types of misallocation are removed, the ratio of exports to manufacturing GDP rises from 0.15 to 0.33 and welfare grows 75%. Although the impact of factor misallocation looks at first glance surprisingly large, these results are in line with the findings in much of the literature that assess the gains of similar reforms.[40] Table 6 displays a decomposition of the aggregate results. The boost in exports is due to the increase in the dispersion of the

[39] The point estimates for $\hat{\lambda}_s$ vary in the range [0.75, 0.87], indicating that around 20% of the observable dispersion in TFPR is attributable to measurement error.

[40] For example, HK find that without affecting firms' selection, an intra-industry reform *"would boost aggregate manufacturing TFP by 86%–115% in China, 100%–128% in India, and 30%–43% in the United States"* (Hsieh and Klenow, 2009, pg. 1420). For Indonesia, Yang (2017) computes TFP gains of 207% from removing manufacturing intra-industry misallocation taking into account firms' selection (97% in the case of a comparable reform to HK). All these large magnitudes are in part due to the extreme nature of the counterfactual, which implies a perfect allocation of factors across all firms, perhaps an unrealistic reform. This is the reason why some papers prefer experiments with gradual reforms (for our case see the next section), or with the reduction of misallocation to the levels observed in a reference country (i.e. the United States, as in HK).

Colombian schedule of comparative advantage. This is evident in Figure 6, which compares the location of the Colombian industries in the RCA world distribution for the initial and counterfactual equilibria, where each vertical line represents a single Colombian industry. This figure also evidences the fact that the counterfactual ranking is not related to the actual one. Industrial chemicals, other chemicals, glass and tobacco are the industries with the largest increases with respect to their initial RCA, whereas petroleum, machinery and equipment, transport equipment and computer, electronic and optical products, display the largest drops. The latter industries disappear when both types of misallocation are removed, indicating the presence of a non-interior solution in the counterfactual equilibrium,[41] which explains in part the longer left tail in the counterfactual world distribution.[42] The larger dispersion on the frictionless comparative advantage leads to higher degrees of industrial specialization in the frictionless equilibrium, which is evident comparing the export shares from panel A to panel B. For instance, the whole chemical sector (both industrial chemicals and other chemicals), an industry that ends up in the first percentile of the counterfactual RCA world distribution, concentrates 64% of the counterfactual Colombian exports, from 23% in the actual data.

The total impact on comparative advantage is a non-linear combination of the effects of removing both HWA wedges and the intra-industry dispersion on the returns of the factors. Panel C and Panel D of Figure 5 depict the RCA measures after removing only intra- and inter-industry misallocation respectively, with markers' sizes representing the counterfactual export shares. In each exercise, I compute the counterfactual values v'_{ils} and Γ'_{is} such that the other type of misallocation remains unchanged. Notice that in both cases the dispersion of comparative advantage is lower than in Panels A and B, but larger with respect to the original one. Table 6 shows that in spite of both types of factor misallocation contributing to the total growth in exports, intra-industry misallocation seems quantitatively more important. Removing only intra-industry misallocation leads to an increase in 13 p.p. of the exports to GDP ratio and a rise in 56% in welfare, whereas removing only inter-industry misallocation causes smaller increases (7 pp. and 8% in each variable, respectively).

The directions and the magnitudes of the changes in the RCA due to each type of factor misallocation can be explained by the extent of its respective causes. The simulations performed in section 3.3 suggested that the magnitude of the effect of intra-industry misallocation depends on the interaction between factor intensities and the relative variances of distortions, whereas the impact of inter-industry misallocation depends on whether the HWA wedges are

[41]The feasibility of non-interior solutions in multi-sector Pareto-Melitz type of models is recently evaluated by Kucheryavyy, Lyn and Rodríguez-Clare (2019). These authors show that under the standard formulation of the model in which the elasticities of substitution do not vary between domestic and foreign varieties, as it is the case in this paper, it is guaranteed that the general equilibrium is unique, but not necessarily an interior solution. Besides multiple factors and resource misallocation, the other difference that makes the model here different is the fact that fixed costs of exporting are paid in terms of factors of the source country.

[42]The counterfactual equilibrium also involves large contractions (between 40% and 70%) in some industries of some of the main Colombian trade partners: 4 in Ecuador, 2 in Brazil, 1 in Venezuela and 1 in Hong Kong.

less or greater than 1. Figure 7 confirms this reasoning. Panel A plots the variation in the RCA when removing intra-industry misallocation against the intra-industry dispersion of the TFPR, equal to $\vec{\alpha}_s'\hat{V}_{is}\vec{\alpha}_s$ for sector s, where $\vec{\alpha}_s$ is a L-vector of factor intensities α_{ls}. The positive correlation suggests that sectors in which firms' TFPR is relatively more disperse, have larger gains in comparative advantage. Analogously, Panel B plots the variation in the RCA when removing inter-industry misallocation against the revenue productivity at the industry level. The positive correlation implies that industries with HWA wedges greater than one gain export capability when inter-industry misallocation is removed, otherwise they lose.

A further exploration of the latter results sheds light on the directions and extents of the general equilibrium effects that are present in the model. Similar to section 3.3, I use the decomposition (B.8) in Appendix B.3 to disentangle the effect of each type of misallocation on comparative advantage into the three sources of export capability in the model: average TFP, the cost of inputs and the number of varieties produced in each sector. Panel A of Figure 8 displays the effect of removing all misallocation (in the top graph), only intra (in the middle graph) and only inter-industry misallocation (in the bottom graph), in each sector's RCA. Towards a better understanding of the results for the RCA, Panel B shows the same decomposition when the changes in the three sources of export capability are not compared across industries, but instead are relative only to the same industry in the reference country. Constructed in this way, the decomposition captures a measure that Hanson, Lind and Muendler (2015) denote the "absolute advantage" index.[43] The numbers displayed correspond to the log-differences between the counterfactual values and the initial values of both measures of export capability, and the lengths of the bars represent the strength of each element in the decomposition, so they add up exactly to the number shown.

First, regarding intra-industry misallocation, the gains on average TFP boost "absolute advantage" of all sectors, on average by 0.91 log points. However, these gains are countered by increases in relative factor prices, on average by 0.74 log points (a rise in relative factor prices is shown as a negative contribution). Thus, in spite of the intensive margin plays a role in the total adjustment of the "absolute advantage" measure, this latter is in a large part driven by the extent to which the number of varieties adjusts, i.e., the extensive margin. When we compute the same decomposition for RCA, its variation is almost entirely explained by the number of varieties. This is a result of the low dispersion in the adjustment of the intensive margin of the "absolute advantage" across sectors, contrary to what happens with

[43]Since I choose to normalize by world means, from (19) the log-differences in the measures of export capability are exactly identified by:

$$logR\hat{C}A_{is} = \frac{\hat{\Gamma}_{is}\hat{R}_{is}\hat{\omega}_{is}^{-\frac{\kappa}{\rho}}}{\prod_{s}^{S}(\hat{\Gamma}_{is}\hat{R}_{is}\hat{\omega}_{is}^{-\frac{\kappa}{\rho}})^{1/S}} \Big/ \frac{\prod_{i}^{N}(\hat{\Gamma}_{is}\hat{R}_{is}\hat{\omega}_{is}^{-\frac{\kappa}{\rho}})^{1/N}}{\prod_{i}^{N}\prod_{s}^{S}(\hat{\Gamma}_{is}\hat{R}_{is}\hat{\omega}_{is}^{-\frac{\kappa}{\rho}})^{1/NS}}; \; log\hat{A}A_{is} = \frac{\hat{\Gamma}_{is}\hat{R}_{is}\hat{\omega}_{is}^{-\frac{\kappa}{\rho}}}{\prod_{i}^{N}(\hat{\Gamma}_{is}\hat{R}_{is}\hat{\omega}_{is}^{-\frac{\kappa}{\rho}})^{1/N}}$$

where AA denotes the "absolute advantage" index.

the number of varieties. Second, regarding inter-industry misallocation, industries facing on average low returns of the factors ($\bar{\Theta}_{is} < 1$, see Table 5) increase their inputs' cost, which improves average TFP through the selection of the more productive firms, compensating the adverse effect of factor prices in both RCA and "absolute advantage" measures, and vice versa. In this case, the magnitudes of the adjustments of average TFP and factor prices in the index of "absolute advantage" are lower than those obtained removing MRP dispersions within industries (for example, the median positive change due to average TFP is 0.25 log points). Nevertheless, despite their smaller magnitudes, those changes have a larger dispersion across sectors, enhancing the contribution of the intensive margin in the effect of inter-industry misallocation on the RCA measure.

4.4 Robustness checks and additional results

In this section, I first evaluate the robustness of the previous results to changes in the parameters κ and σ. Next, I present the results of gradually removing misallocation. Finally, I compare the baseline results with those obtained in the cases of taking the whole manufacturing sector as a single industry and in the closed economy.

Changes in κ and σ

Changes in κ or in σ do not importantly alter the ranking of RCA in the counterfactual equilibria and, if any, have a small effect on its dispersion. Figure 9 displays for the case of removing both types of misallocation the ranking of Colombian RCA measures under different values of κ and σ. Changes in the ranking are negligible, and only small variations in the dispersion are noticeable (see column 5 in Table 6). However, for a given MRP distribution and RCA schedule, the extent of factor reallocations across sectors is increasing in κ and decreasing in σ. This is due to the fact that in each industry a fraction $\frac{\rho}{\kappa}$ of the sectoral demand of factors is not affected by firm-level misallocation, the fraction that is allocated to entry. As a result, Table 6 shows that the rise in total exports and in the ratio exports to GDP is lower for $\kappa = 4$ or $\sigma = 4$ and larger for $\kappa = 5$ or $\sigma = 3$.

Gradual reforms

Figure 10 displays the effects of reforms that gradually remove both and separately the two types of misallocation on the welfare gains (Panel A) and exports growth (Panel B). The lines' values in the extreme right - removing 100% misallocation - coincide with the numbers in Table 6. Even the smallest reform, which reduces 10% the extent of both types of misallocation, has a sizable impact on both welfare and exports (6.7% and 11% respectively).[44] Moreover, it

[44]The exports to GDP ratio only begins to increase after removing 20% misallocation, a threshold where the ranking of industries's RCA starts to show alterations.

is noticeable that for any reduction in misallocation, the intra-industry type is quantitatively more important, although its contribution varies with the intensity of the reform.

One-sector vs. multiple sectors

To quantify the importance of industrial specialization in the exports of the frictionless economy, I perform the exercise of removing misallocation, taking the whole manufacturing sector as a single industry. By construction, there is now only intra-industry misallocation, and all industries face the same factor intensities. Thus, I recompute the corresponding US cost shares and the within-industry variances of firm's wedges, values displayed in the last row of Table 5. The increase in welfare is similar to the baseline case (70%), but the increase in nominal exports is only 43%, leading to a decrease in the ratio of exports to GDP of 5 p.p. (see the last row in Table 6).

Closed vs. open economy

Since in the closed economy revenue shares are constant and equal to the expenditure shares in the demand system, there is no change in the industrial composition under the Cobb Douglas demand. However, it is possible to quantify the cost of the same measures of misallocation in terms of welfare. For this, notice that in the closed economy we have $\pi_{iis} = \hat{\pi}_{iis} = 1$ and $\hat{R}_{is} = \hat{E}_{is} = \hat{E}_i$, so we can express (28) as:

$$\left[\frac{\hat{E}_i}{\hat{P}_i^d}\right]^{closed} = \prod_s \left[\hat{\Gamma}_{is}^{-\frac{1}{\kappa}} \prod_l (\sum_s^S \tilde{Z}_{ils} \hat{v}_{ils})^{\frac{\alpha_{ls}}{\rho}}\right]^{-\beta_s} \tag{29}$$

Thus, the welfare cost of misallocation in a closed economy with endogenous selection of firms can be derived only with measures of misallocation and factor shares in autarky. The last column in Table 6 shows the increase in welfare in the case in which Colombia was a closed economy, under the assumption that the measures of misallocation and factor shares were the same. Apart from the case of removing only inter-industry misallocation, the gains on welfare due to removing allocative efficiency are larger under a closed economy, suggesting that in the particular case of Colombia, international trade dampens the welfare cost of resource misallocation.[45]

[45] For the inter-industry case, the results are in line with Święcki (2017), who shows that simultaneously removing intersectoral wedges in labor in 61 countries and 16 industries leads to larger welfare gains in open economies relative to closed ones (for Colombia, the gains are 18% in the open economy case and 11% under autarky). The intuition for his result is that in the closed economy distorted sectors cannot expand beyond the domestic demand for the sector's output. However, adding firms' endogenous selection can make the effect of trade on the cost of misallocation dependent on the joint distribution of TFPQ and wedges. In particular, trade will have a larger impact on welfare in an economy where the exiting plants due to trade contribute relatively more to the total intra-industry misallocation (i.e., where their TFPR dispersion is higher). In that sense, trade could mitigate or exacerbate the cost of misallocation, particularly of the intra-industry type.

5 Conclusions

Resource misallocation at the firm level can alter the relative unit cost of producing a good across sectors, distorting the "natural" comparative advantage of a country. This paper offers a framework to compute for a country the export capabilities of its industries under frictionless factor markets, considering the general equilibrium effects of factors reallocations both within and across sectors. I perform the exercise with a sample of 48 countries, three production factors, and 25 tradable sectors for the observed misallocation in Colombia, a country whose firm-level data provide us with reliable measures of physical productivity.

I find that the reallocation of factors allows Colombia to specialize in industries with "natural" comparative advantage, especially the whole chemical sector (both industrial chemicals and other chemicals). Reallocating factors generates a rise in the ratio of exports to manufacturing GDP by 18 p.p. and an increase in welfare of 75%, for the case of an extreme reform in which factor misallocation is entirely removed. The specialization channel due to comparative advantage, that substantially transforms the industrial composition when removing firm-level factor misallocation, is a omitted mechanism in the workhorse models of firm-level resource misallocation in closed economies.

The impact of allocative efficiency on comparative advantage depends importantly of the adjustment in the extensive margin. In the case of factor misallocation within industries, I find that removing distortions increases comparative advantage for those sectors in which the returns of the factors used intensively are relatively more dispersed. The gains in terms of unit costs are mainly the result of an increase in the relative number of varieties produced because at the intensive margin the increases on average TFP are largely countered by the responses on relative factor prices, and there is not enough variation across industries of the residual effect. And for inter-industry misallocation, industries in which firms on average face factors' returns larger than the allocative efficient values, increase their comparative advantage when misallocation is removed. In this case, the gains in export capability derive from the reduction of average factor costs, which compensates the adverse selection of firms within the sector, plus an increase in the number of varieties produced. The overall effect of factor misallocation on comparative advantage is a combination of these two forces.

These results suggest that the design of mechanisms that smooths the dispersion on factor returns across firms is a desirable policy. It can boost total productivity and welfare allowing for a more efficient pattern of specialization across industries, in which comparative advantage responds more to differences in efficiency across sectors and relative factor endowments, the "natural" sources of export capability. The growing literature exploring the causes of the dispersion on the factors' returns is a fertile field of research to start exploring optimal policy instruments in an open economy.

References

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia. 2017. "Misallocation, selection and productivity: A quantitative analysis with panel data from China." University of Toronto Working Paper 574.

Anderson, James E., and Yoto V. Yotov. 2010. "The changing incidence of geography." *The American Economic Review*, 100(5): 2157–2186.

Anderson, James E., and Yoto V. Yotov. 2016. "Terms of trade and global efficiency effects of free trade agreements, 1990–2002." *Journal of International Economics*, 99: 279 – 298.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New trade models, same old gains?" *The American Economic Review*, 102(1): 94–130.

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

Bagger, Jesper, Bent Jesper Christensen, and Dale T. Mortensen. 2014. "Productivity and wage dispersion: Heterogeneity or misallocation?" Unpublished.

Balassa, Bela. 1965. "Trade liberalisation and "revealed" comparative advantage." *The Manchester School of Economic and Social Studies*, 33(2): 99–123.

Balistreri, Edward J., Russell H. Hillberry, and Thomas F. Rutherford. 2011. "Structural estimation and solution of international trade models with heterogeneous firms." *Journal of International Economics*, 83(2): 95 – 108.

Banerjee, Abhijit V., and Esther Duflo. 2005. "Growth theory through the lens of development economics." In . Vol. 1, Part A of *Handbook of Economic Growth*, , ed. Philippe Aghion and Steven N. Durlauf, 473 – 552. Elsevier.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-country differences in productivity: The role of allocation and selection." *The American Economic Review*, 103(1): 305–34.

Berlemann, Michael, and Jan-Erik Wesselhöft. 2014. "Estimating aggregate capital stocks using the perpetual inventory method - A survey of previous implementations and new empirical evidence for 103 countries." *Review of Economics*, 65(1): 1–34.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2007. "Comparative Advantage and Heterogeneous Firms." *The Review of Economic Studies*, 74(1): 31–66.

Berthou, Antoine, Jong Hyun Chung, Kalina Manova, and Charlotte Sandoz Dit Bragard. 2018. "Productivity, (Mis)allocation and Trade." Unpublished.

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2018. "Misallocation or mismeasurement?" Unpublished.

Bombardini, Matilde, Christopher J. Kurz, and Peter M. Morrow. 2012. "Ricardian trade and the impact of domestic competition on export performance." *Canadian Journal of Economics/Revue canadienne d'économique*, 45(2): 585–612.

Bombardini, Matilde, Giovanni Gallipoli, and Germán Pupato. 2012. "Skill dispersion and trade flows." *The American Economic Review*, 102(5): 2327–2348.

Brandt, Loren, Trevor Tombe, and Xiaodong Zhu. 2013. "Factor market distortions across time, space and sectors in China." *Review of Economic Dynamics*, 16(1): 39 – 58.

Cabral, Luís M B, and José Mata. 2003. "On the evolution of the firm size distribution: Facts and theory." *American Economic Review*, 93(4): 1075–1090.

Caliendo, Lorenzo, and Fernando Parro. 2015. "Estimates of the trade and welfare effects of NAFTA." *The Review of Economic Studies*, 82(1): 1–44.

Caliendo, Lorenzo, Fernando Parro, and Aleh Tsyvinski. 2017. "Distortions and the structure of the world economy." National Bureau of Economic Research Working Paper 23332.

Chaney, Thomas. 2008. "Distorted gravity: The intensive and extensive margins of international trade." *The American Economic Review*, 98(4): 1707–1721.

Chang, Ha-Joon. 2006. *The East Asian development experience: The miracle, the crisis and the future.* Palgrave Macmillan.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2007. "Business cycle accounting." *Econometrica*, 75(3): 781–836.

Chen, Kaiji, and Alfonso Irarrazabal. 2015. "The role of allocative efficiency in a decade of recovery." *Review of Economic Dynamics*, 18(3): 523 – 550.

Chung, Jong Hyun. 2018. "Firm heterogeneity, misallocation, and trade." Unpublished.

Cole, Harold L., and Lee E. Ohanian. 2002. "The U.S. and U.K. great depressions through the lens of neoclassical growth theory." *The American Economic Review*, 92(2): 28–32.

Costa-Scottini, Lucas. 2018. "Firm-level distortions, trade, and international productivity differences." Unpublished.

Costinot, Arnaud, and Andrés Rodríguez-Clare. 2014. “Trade theory with numbers: Quantifying the consequences of globalization.” In *Handbook of International Economics*. Vol. 4 of *Handbook of International Economics*, , ed. Elhanan Helpman Gita Gopinath and Kenneth Rogoff, 197 – 261. Elsevier.

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer. 2012. “What goods do countries trade? A quantitative exploration of Ricardo’s ideas.” *The Review of Economic Studies*, 79(2): 581–608.

David, Joel M, and Venky Venkateswaran. 2017. “The sources of capital misallocation.” National Bureau of Economic Research Working Paper 23129.

Deardorff, Alan V. 1998. “Determinants of bilateral trade: Does gravity work in a neoclassical World?” In *The Regionalization of the World Economy.* , ed. Jeffrey A. Frankel, Chapter 1, 7–32. University of Chicago Press.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2008. “Global rebalancing with gravity: Measuring the burden of adjustment.” National Bureau of Economic Research Working Paper 13846.

De Loecker, Jan, and Pinelopi Koujianou Goldberg. 2014. “Firm performance in a global market.” *Annual Review of Economics*, 6(1): 201–227.

de Sousa, José, Thierry Mayer, and Soledad Zignago. 2012. “Market access in global and regional trade.” *Regional Science and Urban Economics*, 42(6): 1037 – 1052.

Eaton, Jonathan, and Samuel Kortum. 2001. “Trade in capital goods.” *European Economic Review*, 45(7): 1195 – 1235.

Eaton, Jonathan, and Samuel Kortum. 2002. “Technology, geography, and trade.” *Econometrica*, 70(5): 1741–1779.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2011. “An anatomy of international trade: Evidence from French firms.” *Econometrica*, 79(5): 1453–1498.

Edmond, Chris, Virgiliu Midrigan, and Xu Daniel Yi. 2015. “Competition, markups, and the gains from international trade.” *The American Economic Review*, 105(10): 3183–3221.

Epifani, Paolo, and Gino Gancia. 2011. “Trade, markup heterogeneity and misallocations.” *Journal of International Economics*, 83(1): 1 – 13.

Eslava, Marcela, and John Haltiwanger. 2018. “The life-cycle growth of plants: The role of productivity, demand and "distortions".” Unpublished.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler. 2004. "The effects of structural reforms on productivity and profitability enhancing reallocation: Evidence from Colombia." *Journal of Development Economics*, 75(2): 333 – 371.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler. 2013. "Trade and market selection: Evidence from manufacturing plants in Colombia." *Review of Economic Dynamics*, 16(1): 135 – 158.

Fernandes, Ana M, Peter J Klenow, Sergii Meleshchuk, Denisse Pierola, and Andrés Rodríguez-Clare. 2018. "The intensive margin in trade." National Bureau of Economic Research Working Paper 25195.

Fernández-Val, Iván, and Martin Weidner. 2016. "Individual and time effects in non-linear panel models with large N,T." *Journal of Econometrics*, 192(1): 291 – 312.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?" *The American Economic Review*, 98(1): 394.

French, Scott. 2017. "Revealed comparative advantage: What is it good for?" *Journal of International Economics*, 106: 83 – 103.

Gollin, Douglas, and Christopher R. Udry. 2019. "Heterogeneity, measurement error and misallocation: Evidence from African agriculture." National Bureau of Economic Research Working Paper 25440.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2017. "Capital allocation and productivity in South Europe." *The Quarterly Journal of Economics*, 132(4): 1915–1967.

Hanson, Gordon H., Nelson Lind, and Marc-Andreas Muendler. 2015. "The dynamics of comparative advantage." National Bureau of Economic Research Working Paper 21753.

Head, Keith, and Thierry Mayer. 2014. "Gravity equations: Workhorse,toolkit, and cookbook." In *Handbook of International Economics*. Vol. 4 of *Handbook of International Economics*, , ed. Elhanan Helpman Gita Gopinath and Kenneth Rogoff, 131 – 195. Elsevier.

Ho, Giang T. 2012. "Trade liberalization with size-dependant distortions: Theory and evidence from India." Unpublished, IMF.

Hopenhayn, Hugo A. 2014*a*. "Firms, misallocation, and aggregate productivity: A Review." *Annual Review of Economics*, 6(1): 735–770.

Hopenhayn, Hugo A. 2014*b*. “On the measure of distortions.” National Bureau of Economic Research Working Paper 20404.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. “Misallocation and manufacturing TFP in China and India.” *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Krugman, Paul. 1980. “Scale economies, product differentiation, and the pattern of trade.” *The American Economic Review*, 70(5): 950–959.

Kucheryavyy, Konstantin, Gary Lyn, and Andrés Rodríguez-Clare. 2019. “Grounded by gravity: A well-behaved trade model with external economies.” Unpublished.

Kugler, Maurice, and Eric Verhoogen. 2012. “Prices, plant size, and product quality.” *The Review of Economic Studies*, 79(1): 307–339.

Lahiri, Amartya, and Kei-Mu Yi. 2009. “A tale of two states: Maharashtra and West Bengal.” *Review of Economic Dynamics*, 12(3): 523 – 542.

Lane, Nathan. 2017. “Manufacturing revolutions.” Unpublished.

Leromain, Elsa, and Gianluca Orefice. 2014. “New revealed comparative advantage index: Dataset and empirical distribution.” *International Economics*, 139: 48 – 70.

Levchenko, Andrei A., and Jing Zhang. 2016. “The evolution of comparative advantage: Measurement and welfare implications.” *Journal of Monetary Economics*, 78: 96 – 111.

Liu, Jing, Nico van Leeuwen, Tri Thanh Vo, Rod Tyers, and Thomas Hertel. 1998. “Disaggregating labor payments by skill level in GTAP.” Purdue University GTAP Technical Paper 11.

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

Melitz, Marc J., and Stephen J Redding. 2014. “Heterogeneous firms and trade.” In *Handbook of International Economics*. Vol. 4 of *Handbook of International Economics*, , ed. Elhanan Helpman Gita Gopinath and Kenneth Rogoff, 1–54. Elsevier.

Melitz, Marc J., and Stephen J. Redding. 2015. “New trade models, new welfare implications.” *The American Economic Review*, 105(3): 1105–1146.

Mulligan, Casey B. 2005. “Public policies as specification errors.” *Review of Economic Dynamics*, 8(4): 902 – 926.

Neyman, J., and Elizabeth L. Scott. 1948. “Consistent estimates based on partially consistent observations.” *Econometrica*, 16(1): 1–32.

Oberfield, Ezra. 2013. "Productivity and misallocation during a crisis: Evidence from the Chilean crisis of 1982." *Review of Economic Dynamics*, 16(1): 100 – 119.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic Dynamics*, 11(4): 707 – 720.

Restuccia, Diego, and Richard Rogerson. 2013. "Misallocation and productivity." *Review of Economic Dynamics*, 16(1): 1 – 10.

Rodrik, Dani. 1995. "Getting interventions right: How South Korea and Taiwan grew rich." *Economic Policy*, 10(1): 55 – 107.

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. "The log of gravity." *The Review of Economics and Statistics*, 88(4): 641–658.

Święcki, Tomasz. 2017. "Intersectoral distortions and the welfare gains from trade." *Journal of International Economics*, 104: 138 – 156.

Timmer, Marcel P., Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries. 2015. "An illustrated user guide to the world input-output database: The case of global automotive production." *Review of International Economics*, 23(3): 575–605.

Tombe, Trevor. 2015. "The missing food problem: Trade, agriculture, and international productivity differences." *American Economic Journal: Macroeconomics*, 7(3): 226–58.

Yang, Mu-Jeung. 2017. "Micro-level misallocation and selection: Estimation and aggregate implications." Unpublished.

Tables

Table 1 – Alternative explanations for dispersion in revenue productivity

Source	Variable	Contribution*	Countries	Paper
I. Model misspecification				
Adjustment costs	σ^2_{MRPK}	1%	China, Colombia and Mexico	David and Venkateswaran (2017)
Uncertainty about firms' TFP		7%		
Variable markups		5%	China	
Heterogeneity in technology		17%		
Heterogeneity in workers' ability	σ^2_{MRPL}	9%	Denmark	Bagger, Christensen and Mortensen (2014)
II. Measurement error (M.E.)				
M.E additive in revenues and inputs	σ^2_{TFPR}	45%	India	Bils, Klenow and Ruane (2018)

Notes: σ^2_{TFPR} corresponds to the variance of revenue productivity (TFPR), which is a function of the variances (and covariances) of the marginal revenue products (MRP), σ^2_{MRPz} for factor z. The table displays the contribution of causes different to misallocation to the corresponding variances of the MRP (for capital (K) and labor (L)) or directly to the TFPR. *Average contribution if the number of countries is greater than one.

Table 2 – RCA explained by misallocation measures and determinants of export capability

	(1) $dRCA_{ist}$	(2) $dRCA_{ist}$	(3) $dRCA_{ist}$	(4) $dRCA_{ist}$
Intra-ind. allocative efficiency	0.358***		0.575***	0.339***
	(0.082)		(0.088)	(0.084)
Intra-ind. variance of TFPR		-0.145**		
		(0.060)		
Inter-industry wedges	-0.351***	-0.241***	-0.202**	-0.371***
	(0.081)	(0.088)	(0.063)	(0.085)
Efficient TFP	0.244**	0.234**	0.218**	0.272***
	(0.090)	(0.098)	(0.103)	(0.088)
Factor prices	-0.318***	-0.197**	-0.263***	-0.306***
	(0.066)	(0.076)	(0.077)	(0.067)
Observations	208	208	208	208
R-square	0.327	0.266	0.551	0.23

Notes: * p<0.10, ** p<0.05 and *** p<0.01. The results correspond to the second-stage of the econometric strategy, where in the first stage the exporter-industry FE are estimated by PPML. The dependent variable is $dRCA_{ist}$, the change in the RCA measure with respect to the first period. All independent variables are transformed to be changes with respect to the first period relative to the reference industry, normalized by the corresponding changes in the US PPI. (1) and (2) are the baseline results. (3) changes reference industry (to min. number of zeros), (4) changes set of countries (to 19). Standardized coefficients and heteroskedastic robust errors.

Table 3 – Equilibrium conditions and endogenous variables

Equilibrium condition	Equation	Dimension
Factor clearing	(13)	$N \times L$
Industry factor demand	(12)	$N \times L \times S$
Zero profit	(7)	$N \times N \times S$
Aggregate stability	(9)	$N \times N \times S$
Free entry	(8)	$N \times S$
Industry price	(10)	$N \times S$
Industry demand	$Q^d_{is} = (\sum_k^N \sum_m^{M_{kis}} q^{\rho}_{kim})^{\frac{1}{\rho}}$	$N \times S$
Aggregate price	$P^d_i = \prod_s^S (\frac{P^d_{is}}{\beta_s})^{\beta_s}$	N
Trade balance	(14)	N

Endogenous variable	Notation	Dimension
Primary factor price	w_{il}	$N \times L$
Industry-level primary factor	Z_{ils}	$N \times L \times S$
Cutoffs for undistorted firms by dest.	a^*_{ijs}	$N \times N \times S$
Mass of firms by destination	M_{ijs}	$N \times N \times S$
Mass of entrants	H_{is}	$N \times S$
Industry-level consumer price & demand	P^d_{is}, Q^d_{is}	$2 \times N \times S$
Aggregate consumer price & demand	P^d_i, Q^d_i	$2 \times N$

Table 4 – Parameters used in simulations

Parameter	Description	Value
α_{ls}	Factor intensities	$\begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$
β_{is}	Expenditure shares	$0.5\ \forall\ i, s$
σ	Varieties' elasticity of substitution	3.8
κ	Pareto's shape parameter	4.58
$\bar{Z}_{il}$	Factor endowments	$\begin{bmatrix} 100 & 90 \\ 90 & 100 \end{bmatrix}$
$\bar{a}_{is}$	Pareto's location parameter	$1\ \forall\ i, s$
δ_{is}	Exogenous probability of exit	$0.025\ \forall\ i, s$
f^e_{is}	Fixed entry cost	$2\ \forall\ i, s$
f_{ijs}	Fixed trade cost	$2\ \forall\ i, j, s$
τ_{ijs}	Iceberg trade cost	Free trade: $1\ \forall\ i, j, s$ Costly trade: $2\ \forall\ s \wedge i \neq j$; $1\ \forall\ s \wedge i = j$
σ_{l1}	Log-normal shape par. in sector 1	For figure 3: $[0, 0.5]\ \forall\ l$ For figure 4: $0\ \forall\ l$
μ_{l1}	Log-normal location par. sector 1	For figure 3: $(\frac{1}{2} - (1 - \frac{\kappa}{\rho})\alpha_{l1})\sigma^2_{l1}\ \forall\ l$ For figure 4: $[-0.5, 0.5]\ \forall\ l$

Table 5 – Factor intensities and misallocation measures used in counterfactuals

Sector	Number of firms (in 1995)	Factor intensities (GO specification)			Inter-industry wedges (HWA of firm-level wedges)				Corrected** intra-industry variances of log-wedges			Corrected** intra-industry covariances of log-wedges			BKR's (2017) "compression"	
		α_k	α_s	α_u	$(1+\bar{\theta}_k)$	$(1+\bar{\theta}_s)$	$(1+\bar{\theta}_u)$	$\bar{\Theta}$	σ_k^2	σ_s^2	σ_u^2	σ_{ks}	σ_{ku}	σ_{su}	$\hat{\lambda}_s{}^*$	s.e.
Food	1435	0.31	0.06	0.09	1.90	1.01	1.14	1.15	1.07	1.09	1.20	0.19	0.19	0.86	0.81a	0.13
Beverage	142	0.36	0.06	0.06	1.05	0.98	1.14	1.33	0.90	0.76	0.75	0.00	-0.07	0.49	0.79	1.74
Tobacco	9	0.73	0.02	0.04	1.67	1.64	0.39	1.28	0.53	1.24	1.62	0.28	-0.34	0.94	0.76a	0.02
Textiles	465	0.22	0.08	0.18	0.81	1.08	0.88	1.02	1.33	0.71	0.69	-0.06	0.08	0.43	0.82	0.76
Apparel	944	0.23	0.10	0.17	1.25	0.40	0.26	0.72	1.27	0.65	0.61	0.11	0.16	0.29	0.87a	0.04
Leather	118	0.32	0.12	0.16	1.38	1.00	0.47	0.73	0.89	0.73	0.46	-0.01	-0.06	0.46	0.84a	0.09
Footwear	254	0.21	0.12	0.20	1.51	1.00	0.59	0.97	1.09	0.66	0.46	0.08	0.12	0.34	0.80	0.73
Wood	196	0.13	0.07	0.18	0.25	0.37	0.48	0.51	1.43	0.45	0.37	0.27	0.15	0.29	0.86a	0.12
Furniture	270	0.18	0.11	0.25	0.70	0.27	0.32	0.50	1.45	0.40	0.40	0.12	0.01	0.20	0.85	0.58
Paper	170	0.21	0.09	0.18	0.64	2.40	2.62	1.17	0.94	0.80	1.10	0.05	-0.03	0.68	0.79c	0.44
Printing	434	0.23	0.15	0.26	1.02	0.83	1.62	1.02	0.74	0.50	0.50	-0.05	-0.09	0.20	0.85a	0.03
Chemicals	177	0.37	0.07	0.08	1.23	1.96	1.77	1.08	1.43	0.78	0.76	0.11	-0.06	0.54	0.83a	0.06
Other chemicals	356	0.36	0.12	0.09	2.50	1.13	1.49	1.53	1.02	0.71	0.85	-0.07	-0.11	0.50	0.81	0.98
Petroleum	46	0.15	0.02	0.02	0.65	0.98	0.86	1.28	2.02	1.14	1.47	0.82	0.97	1.20	0.76a	0.01
Rubber	93	0.20	0.12	0.22	0.63	2.01	1.64	1.05	0.68	0.61	0.48	0.20	0.20	0.33	0.83	1.24
Plastic	428	0.10	0.08	0.28	0.38	0.95	1.74	1.04	0.83	0.61	0.59	-0.01	-0.04	0.39	0.83a	0.02
Pottery	13	0.27	0.13	0.30	1.16	1.19	1.38	1.11	0.18	0.46	0.73	-0.06	-0.08	0.56	0.80a	0.01
Glass	82	0.26	0.29	0.12	0.91	4.59	0.70	1.38	0.97	0.53	0.49	-0.15	0.02	0.33	0.80	2.72
Other non-metallic	365	0.21	0.07	0.14	0.46	1.36	1.11	1.05	1.28	0.72	0.91	0.02	-0.01	0.64	0.80	2.59
Iron and steel	86	0.18	0.10	0.21	0.50	2.74	3.01	1.28	0.91	1.08	1.35	-0.15	-0.12	1.07	0.78a	0.01
Non-ferrous metal	42	0.18	0.10	0.27	0.38	0.56	0.94	0.39	0.44	0.78	1.22	-0.14	-0.40	0.89	0.82a	0.03
Metal products	664	0.21	0.12	0.17	1.09	1.20	0.72	0.99	1.27	0.58	0.55	0.09	0.08	0.39	0.84b	0.35
Mach. & equipment	374	0.25	0.11	0.09	1.50	0.83	0.36	1.04	0.94	0.43	0.46	0.02	0.12	0.28	0.83a	0.02
Electric. / Profess.	276	0.19	0.02	0.08	1.00	1.27	0.74	1.01	0.94	0.59	0.62	0.05	0.06	0.43	0.78	0.58
Transport	274	0.24	0.15	0.13	2.23	0.45	0.91	1.20	0.93	0.48	0.73	0.19	0.23	0.38	0.84a	0.02
One-sector	7713	0.24	0.09	0.13	1.00	1.00	1.00	1.00	1.13	1.05	0.86	0.08	0.08	0.63	0.85a	0.33

Notes: *Point estimates for λ_s using Bils, Klenow and Ruane (2018) (see Appendix A.2). Levels of significance: c $p < 0.1$, b $p < 0.05$, a $p < 0.01$.

**"Corrected" values correspond to the product of the observed dispersion (after removing outliers and trimming 1% tails) and the corresponding value for λ_s. For non-significant values of λ_s, the value of the last row is used, a specification that controls for industry×years fixed effects.

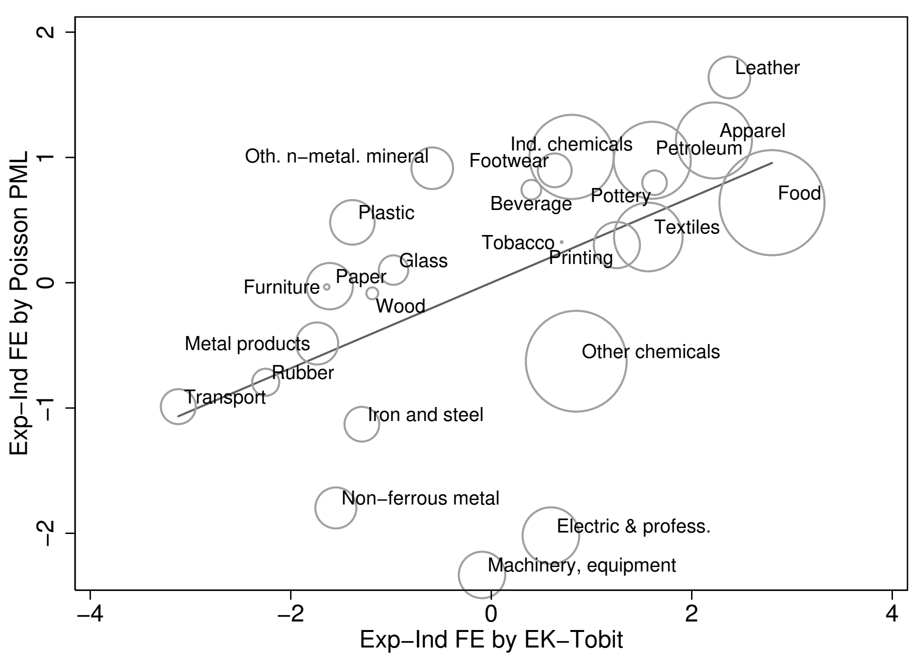
Table 6 – Counterfactuals

	Change in each variable after removing factor misallocation in Colombia						
Variable	Revenue	Value added	Exports	Exports /GDP*	RCA s.d.*	Welfare	Welfare - autarky
Counterfactual	$\hat{R}_{Col}$	$G\hat{D}P_{Col}$	$\hat{X}_{Col}$	$\Delta(\frac{X}{GDP})_{Col}$	$\Delta\sigma_{RCA_{Col}}$	$\frac{\hat{E}_{Col}}{\hat{P}_{Col}}$	$\left[\frac{\hat{E}_{Col}}{\hat{P}_{Col}}\right]^{closed}$
Baseline results							
Both types	1.54	2.22	4.78	0.18	2.60	1.75	1.85
Only intra-industry	1.41	1.92	3.59	0.13	1.95	1.56	1.72
Only inter-industry	1.04	1.09	1.57	0.07	1.69	1.08	1.07
Robustness: Both types							
Decreasing σ (to 3)	1.59	2.35	5.22	0.19	2.68	1.90	1.99
Increasing σ (to 4)	1.50	2.14	4.51	0.17	2.69	1.67	1.76
Decreasing κ (to 4)	1.44	2.01	4.14	0.16	2.40	1.64	1.75
Increasing κ (to 5)	1.61	2.38	5.36	0.19	2.61	1.84	1.92
One-sector							
Only intra-industry	1.58	2.32	1.43	-0.05	-	1.70	1.87

Note: Each cell shows the proportional change in each variable between the counterfactual equilibrium and the actual data. For variables marked by *, the simple difference in the measure is displayed.

Figures

Figure 1 – Revealed comparative advantage (RCA) measures for Colombia



Notes: Markers' sizes represent export shares, and the line the best linear fitting.

Figure 2 – Cutoff functions and selection effects of distortions

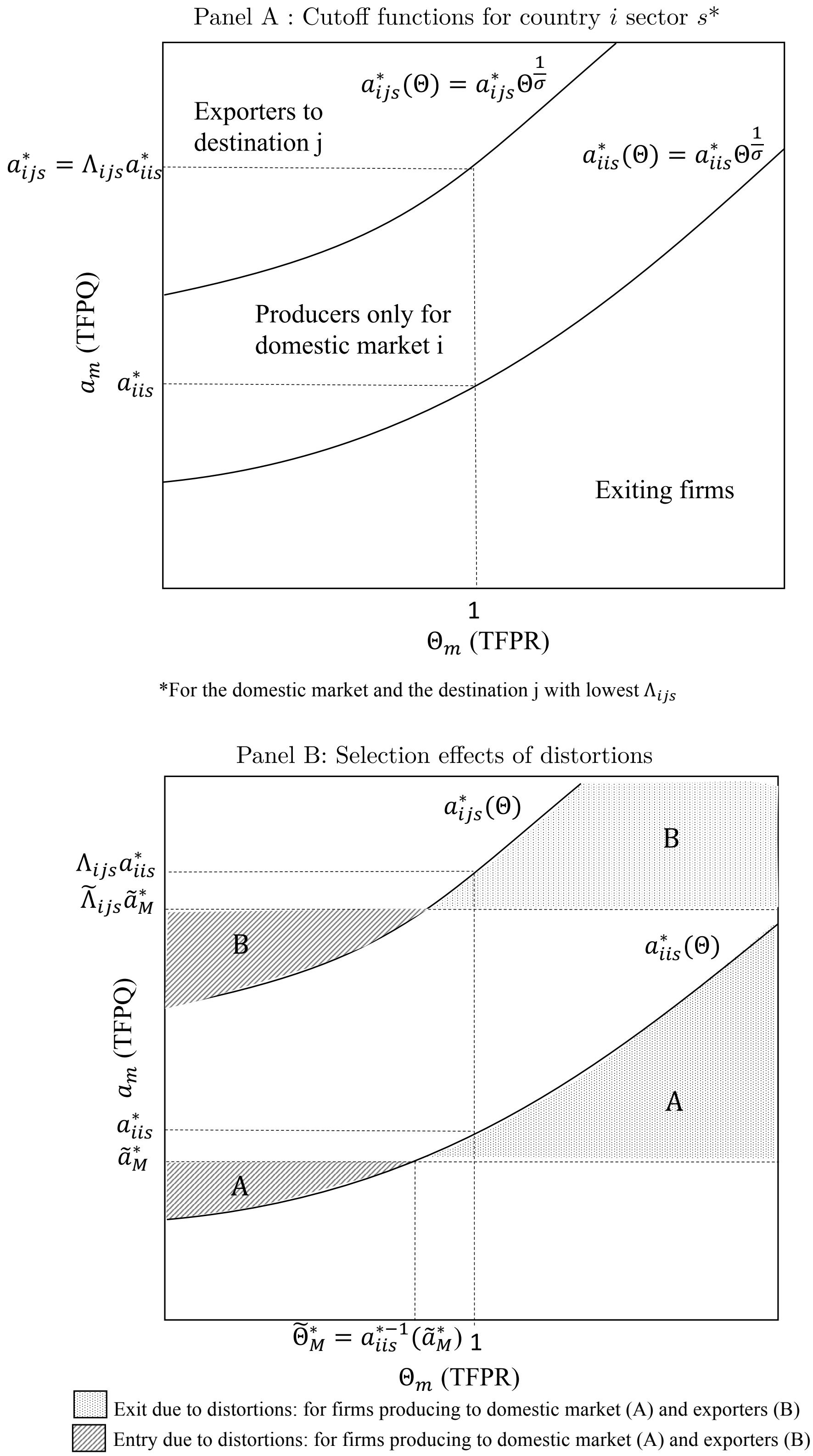


Figure 3 – Effects of factor misallocation within industries on RCA and its determinants

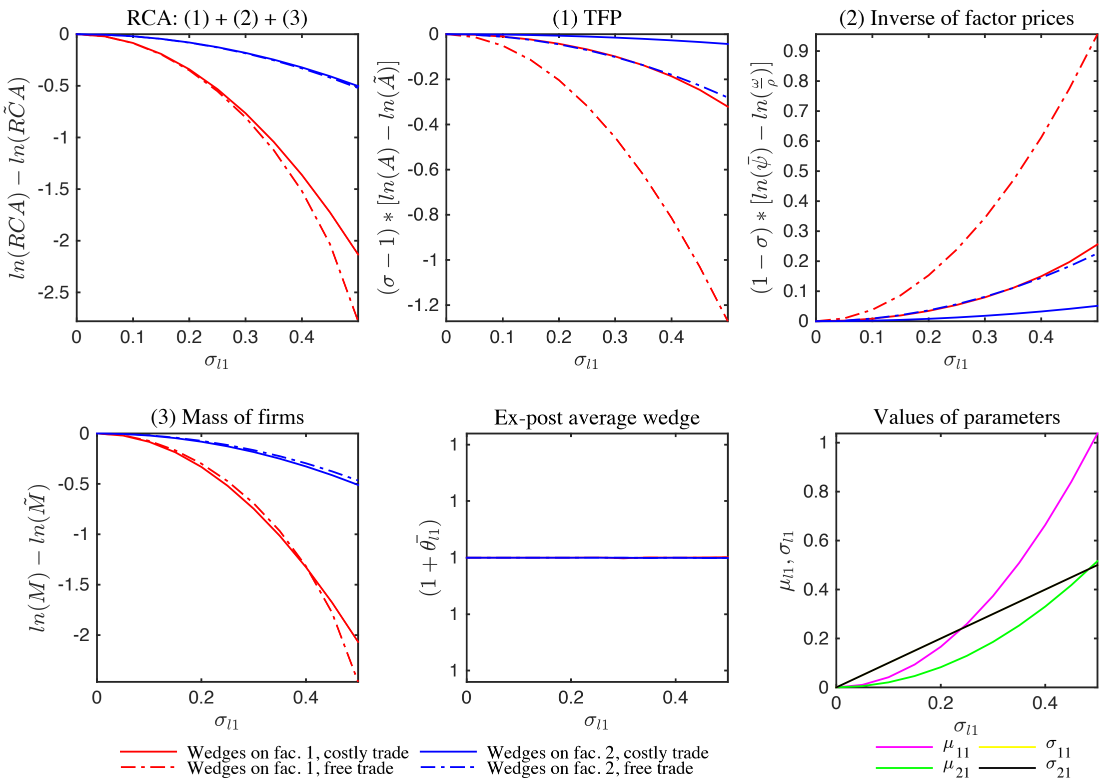


Figure 4 – Effects of factor misallocation across industries on RCA and its determinants

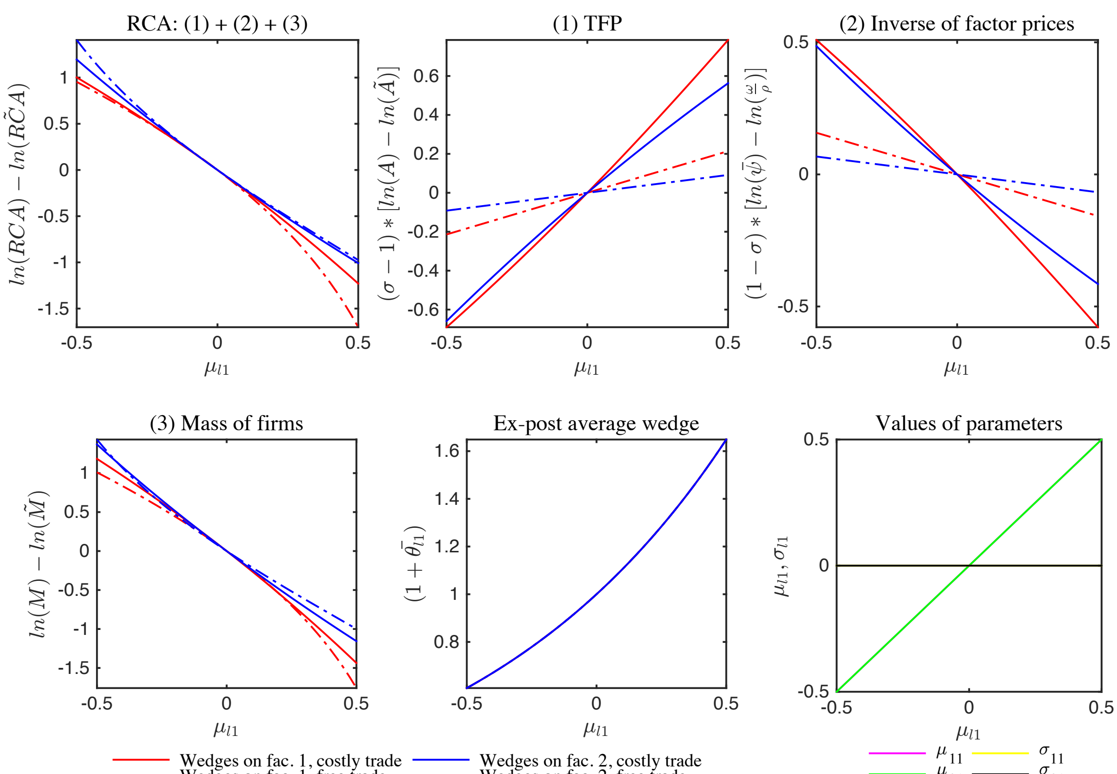
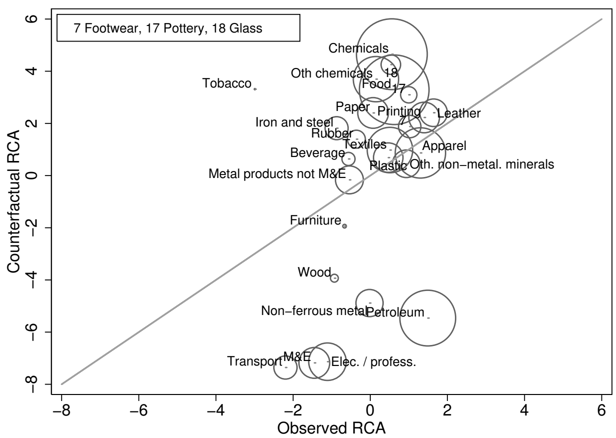
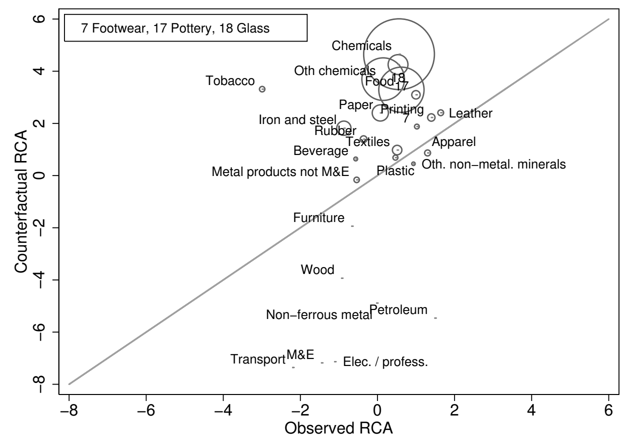


Figure 5 – Allocative efficient RCA and observed RCA for Colombia

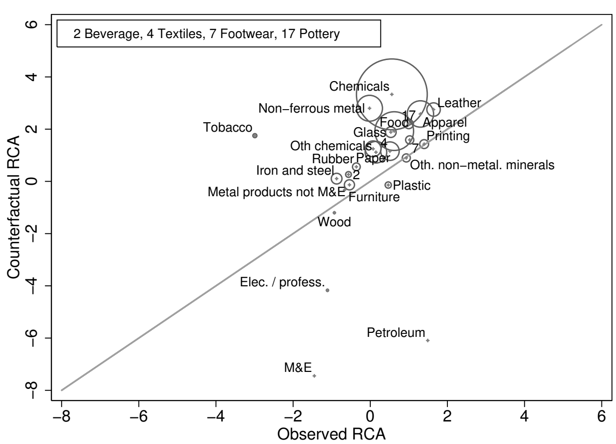
Panel A: Intra- and inter-industry allocative efficient RCA and observed RCA (observed export shares)



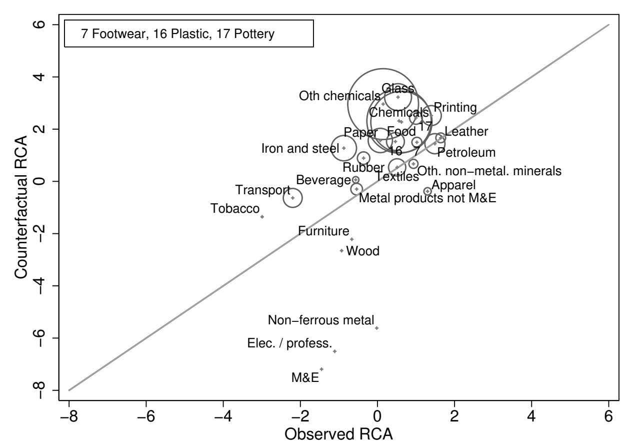
Panel B: Intra- and inter-industry allocative efficient RCA and observed RCA (counterfactual export shares)



Panel C: Only intra-industry allocative efficient RCA and observed RCA (counterfactual export shares)



Panel D: Only inter-industry allocative efficient RCA and observed RCA (counterfactual export shares)

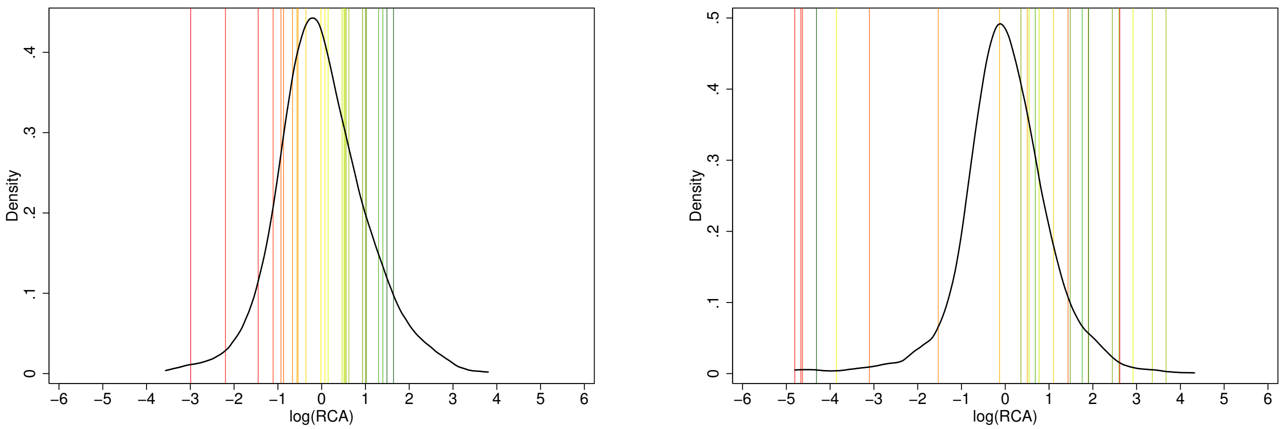


Notes: Each panel compares the RCA measures in the corresponding counterfactuals to the observed RCA measures. Markers' sizes represent the indicated export shares.

Figure 6 – Colombian industries in the world distribution of RCA

Panel A: Distribution under observed data

Panel B: Distribution under Colombia's efficient allocation

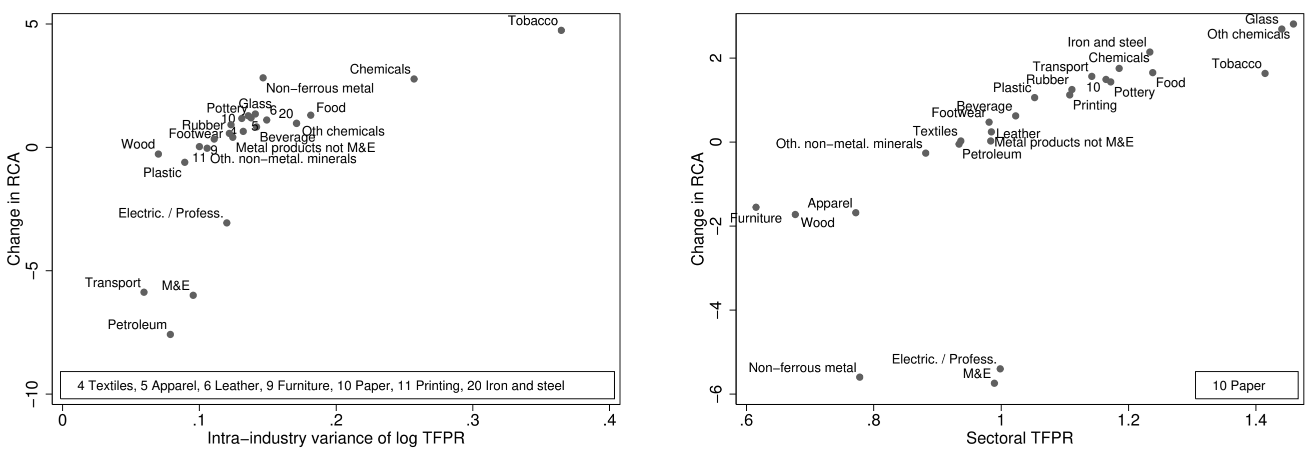


Note: Each vertical line represents the location of a Colombian industry in the RCA world distribution.

Figure 7 – Changes in Colombian RCA and their causes

Panel A: Change in RCA by removing intra-industry misallocation and within-industry variance of TFPR

Panel A: Change in RCA by removing inter-industry misallocation and sectoral TFPR for Colombia



Notes: Intra-industry variance of log TFPR in Panel A is constructed as the weighted average of the within-industry dispersion of the factors' MRP that face misallocation: capital, skilled and unskilled labor. Similarly, sectoral TFPR in Panel B is computed using only capital, skilled and unskilled labor as inputs.

Figure 8 – Changes in determinants of Colombian RCA

Panel A: Changes in comparative advantage determinants Panel B: Changes in absolute advantage determinants

1. Removing intra- and inter-industry misallocation

2. Removing only intra-industry misallocation

3. Removing only inter-industry misallocation

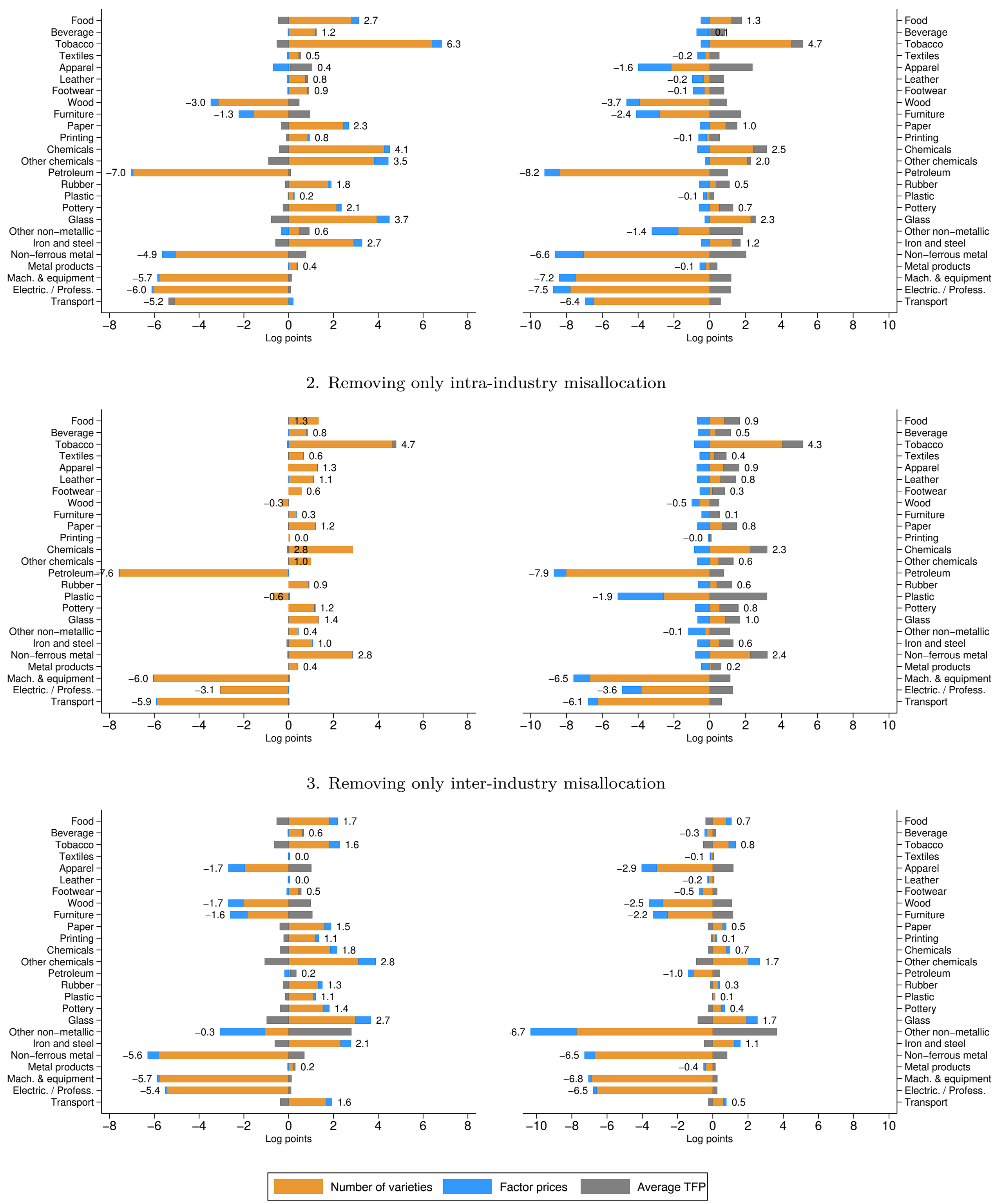


Figure 9 – Rankings of RCA for different values of κ and σ

Panel A: Changes in σ | Panel B: Changes in κ

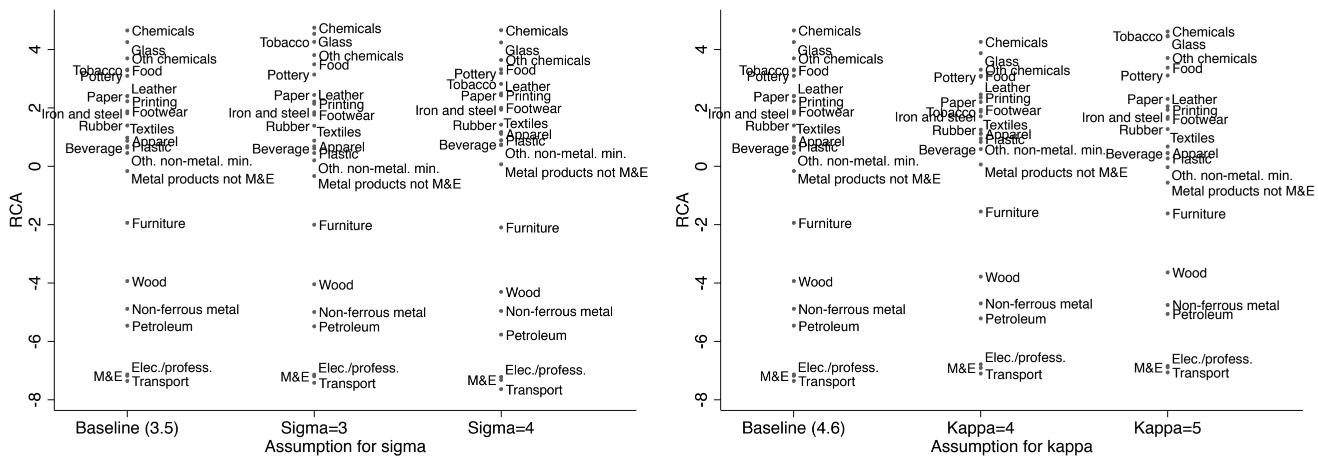
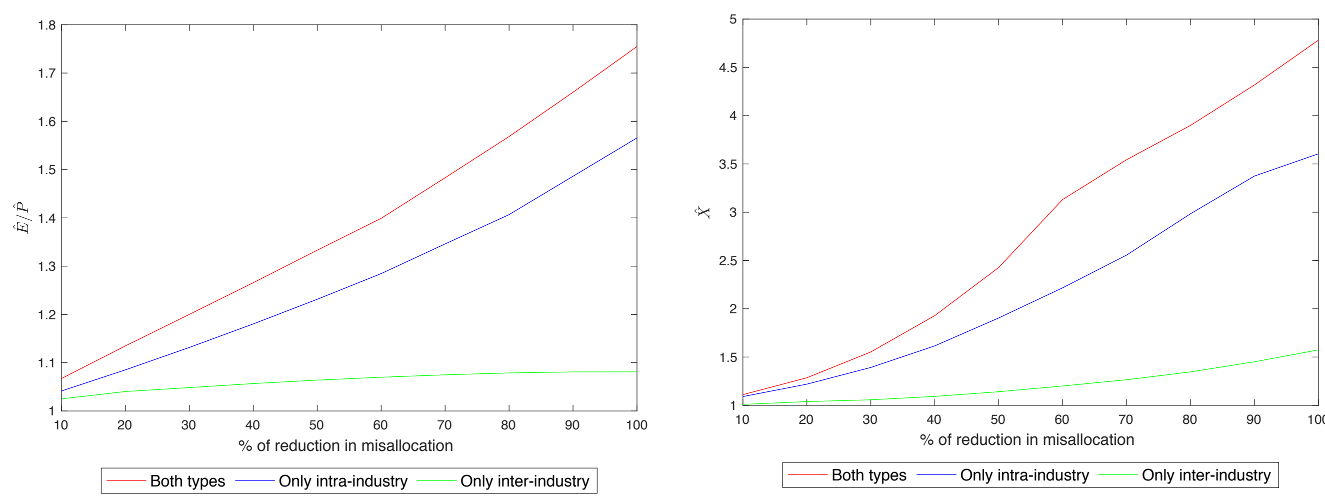


Figure 10 – Welfare gains and export growth from gradual reforms

Panel A: Welfare gains | Panel B: Export growth



Appendix

A Data and solution of the model

A.1 Description of the dataset

This paper uses two types of data: A "macro" dataset with information at the country-sectoral level, and a "micro" dataset, with information at the firm level for Colombia.

The "macro" dataset collects sectoral information of gross output, bilateral trade flows, intermediate consumption and shares of employment and capital for a sample of 48 countries and 25 manufacturing industries (3-digit ISIC rev. 2 level), for the year 1995. Table A.1 and A.2 at the end of this section display the considered countries and industries respectively.

Data for sectoral gross output, bilateral trade flows and intermediate consumption come from OECD's Trade in Value Added (TiVA) database (2015's release). This dataset contains a range of indicators derived from the OECD's Inter-Country Input-Output (ICIO) database. The latter is constructed by OECD from various national and international data sources, all drawn together and balanced under constraints based on official National Accounts (SNA93).[46] Information on gross output and trade flows was collected for all available manufacturing sectors in TiVA (16), and an imputation scheme was implemented to obtain output and bilateral flows for the remaining sectors and for two countries not available in TiVA (Venezuela and Ecuador, which were included given their relevance as Colombia's trade partners), based on production and trade shares computed from the CEPII database (de Sousa, Mayer and Zignago, 2012).

I derive imports from home from the difference between gross output and total exports. As it is known in the literature, this procedure could generate negative values for some country-industry pairs (for instance if the country-sector has high amount of reexports). To solve this issue, I follow Costinot and Rodríguez-Clare (2014) and Święcki (2017), adjusting those negative flows rescaling exports to all destinations until the ratio total exports to gross output is as in the sector with the highest ratio still less than one in that country. This adjustment was needed in the case of six country-industry observations.

Factors shares were constructed using information from several sources. For materials, I compute the shares using the series of intermediate consumption from TiVA. Data for the remaining industries and for Venezuela and Ecuador was imputed using shares from UNIDO's INDSTAT2 database (2015's release), which contains information at the 2-digit ISIC rev. 3 level only for manufacturing industries. The information was gathered adjusting each country's available aggregation to the one used here. For labor, ICIO database contains information of employment (measured in number of persons engaged) for 42 of the 48 countries considered here. For the remaining sectors and countries, data was collected using UNIDO's INDSTAT2 database. Skilled and unskilled labor shares were allocated using GTAP-5 database, which are draw on labor force surveys and national censuses where they are available, or the statistical model proposed by Liu et al. (1998) otherwise.

For capital, shares were constructed as follows. First, the Social Economics Accounts of the World Input Output Database (WIOD, see Timmer et al. (2015)) contain calculations

[46] The underlying sources used are notably: i) National supply and use tables; ii) National and harmonized Input-Output Tables, iii) Bilateral trade in goods by industry and end-use category; and iv) Bilateral trade in services. For more information, see www.oecd.org/trade/valueadded

of the stocks of capital at the two-digit ISIC rev. 3 level or groups thereof for 36 countries of the 48 countries considered here (in the 2013's release). For the remaining countries, I apply the steady-state approach on the calculation of the initial stock of capital in the perpetual inventory method,[47] using information of gross fixed capital formation (GCFC) from INDSTAT2 database. For country i-industry s the share of capital γ_{iks} was imputed as:

$$\gamma_{iks} = \frac{\frac{GCFC_{is}}{g_{is}+\delta^r_{is}}}{\sum_s^S \frac{GCFC_{is}}{g_{is}+\delta^r_{is}}}$$

where $GCFC_{is}$ is the average GCFC over the five-year window centered on the reference year, g_{is} is the growth rate of the GDP of the sector in the same period, and δ^r_{is} is an exogenous depreciation rate, which are computed using the NBER-CES Manufacturing Industry database for US.[48] I compute capital shares using this methodology even for the countries with available information from WIOD, to assess the fit of the imputation procedure. I evaluate the imputation results in terms of cross correlations and mean absolute errors using three approximations: i) Setting $g_{is} = \delta^r_{is} = 0\ \forall i, s$ (thus I use only information on GCFC); ii) Setting $g_{is} = 0\ \forall i, s$ (hence I use information on GCFC and US depreciation rates); iii) Using the full set of information. I found the best adjustment under the second approach. Therefore, capital shares for the remaining countries were imputed using only series of GCFC and US depreciation rates.

For the "micro" dataset I use the panel of manufacturing plants created by Eslava et al. (2004) (hereafter EHKK) for the period 1984-1998 from the Colombian Annual Manufacturing Survey (AMS), collected by the Departamento Administrative Nacional de Estadística (DANE), the Colombian national statistical agency. The AMS is a census of plants with 10 or more workers or annual sales above certain limit, which is adjusted over time.[49] A unique feature of the AMS is that, in conjunction with the main variables of standard surveys (output and sales values, overall cost, energy consumption, payroll, number of workers and book values of equipment and structures), the DANE collects information at the product level (with a disaggregation comparable to the 6-digit HS) on the value and physical quantities of outputs and inputs (valued at factory-gate prices). This allows EHKK to obtain prices as unit values for each output and input produced and used by every plant, and hence to construct specific firm prices of total output and materials using Tornqvist indices (see EHHK Appendix for details).

I perform the detailed cleaning procedure of Kugler and Verhoogen (2012) to reduce the influence of measurement error and outliers (see their data Appendix). Next, I follow HK and remove 1% tails of the distributions of $log(\psi_m/\bar{\psi}_s)$ and $log(M_s^{\frac{1}{\sigma-1}} a_m/\widetilde{A}_s)$ to drop remaining influential observations.[50] Following the misallocation literature, to obtain TFP measures I use as factor intensities average U.S. cost shares at the corresponding aggregation levels from the NBER-CES Manufacturing Industry Database during the same period of time. Since for

[47]For reference, see for example Berlemann and Wesselhöft (2014).

[48]I use five-year windows to prevent that short-run volatility in the GCFC bias the imputation results. Notice that since I only need sectoral factor shares, a temporal shock that affects homogeneously the whole economy does not affect the imputation results.

[49]For 1998, the last year of the panel, was around US$35000. This criterium was introduced in the AMS in 1992 to increase coverage.

[50]For the definitions of $\bar{\psi}_s$, M_s and $\widetilde{A}_s$ see Appendix C.

the selected years the AMS uses ISIC rev-2 adapted for Colombia, I match the NAICS97 US code with the ISIC rev-3, and afterwards with the Colombian one. The purpose of using US cost shares is to employ factor intensities that reflect true technological differences across industries instead of frictions in factor markets, since domestic cost shares can be affected by the extent of inter-industry factor misallocation.

The final panel contains around 4700 plants on average in a typical year. On average, around 390 firms enter each year while 450 exit, which corresponds to an entry/exit rate of 8 and 9 percent respectively. For the computation of the misallocation measures in the counterfactual exercise, I use information only for the reference year (1995). Despite its coverage, EHHK's dataset does not include exports. Thus, I use the panel employed by Bombardini, Kurz and Morrow (2012) for 1978-1991, which has been used extensively in the literature, to obtain exports. I merge both panels using variables in quantities (year, 4-digit ISIC, production and non-production workers and energy consumption). For the overlapping period, plants representing between 2% and 3% of the original nominal production were unmatched, and therefore dropped from the sample. I also keep only plants with positive and non-missing values for production and inputs. Up to 1991, on average around 13 of each 100 firms were exporters, while the total value exported represents in average 8% of industry's gross revenue, with a large variation across sectors.

With the goal to ensure consistency between the macro and the micro dataset, two procedures were executed. First, since the calculation of factor shares in the macro dataset is independent on the series of gross output and bilateral trade flows, factor shares for Colombia were taken directly from the AMS. It is worth to say that the factor shares computed by both sources are very similar, minor differences occur due to the exclusion of outliers in the micro dataset. Second, revenues of all firms within each industry were re-scaled to ensure that the revenue share included in the TiVA database coincide with the corresponding shares on the AMS. Once again, revenue shares from the two sources are very alike, and the small discrepancies also occur for the exclusion of outliers.

Table A.1 – Countries in the sample

OECD Country (I)	Code	OECD Country (II)	Code	Non-OECD Country	Code
Australia	AUS	Korea	KOR	Argentina	ARG
Austria	AUT	Mexico	MEX	Brazil	BRA
Belgium	BEL	Netherlands	NLD	China	CHN
Canada	CAN	New Zealand	NZL	Colombia	COL
Chile	CHL	Norway	NOR	Ecuador	ECU
Denmark	DNK	Poland	POL	Hong Kong	HKG
Finland	FIN	Portugal	PRT	India	IND
France	FRA	Czech Republic	CZE	Indonesia	IDN
Germany	DEU	Spain	ESP	Malaysia	MYS
Greece	GRC	Sweden	SWE	Philippines	PHL
Hungary	HUN	Switzerland	CHE	Rest of the World	ROW
Ireland	IRL	Turkey	TUR	Romania	ROU
Israel	ISR	United Kingdom	GBR	Russia	RUS
Italy	ITA	United States	USA	Saudi Arabia	SAU
Japan	JPN			Singapore	SGP
				South Africa	ZAF
				Thailand	THA
				Taiwan	TWN
				Venezuela	VEN

Table A.2 – Sectors in the sample

No.	Sector	Sector Description	ISIC Rev. 2
1	Food	Food manufacturing	311-312
2	Beverage	Beverage industries	313
3	Tobacco	Tobacco manufactures	314
4	Textiles	Manufacture of textiles	321
5	Apparel	Wearing apparel, except footwear	322
6	Leather	Leather and products of leather and footwear	323
7	Footwear	Footwear, except vulcanized or moulded rubber or plastic footwear	324
8	Wood	Wood and products of wood and cork, except furniture	331
9	Furniture	Furniture and fixtures, except primarily of metal	332
10	Paper	Paper and paper products	341
11	Printing	Printing, publishing and allied industries	342
12	Chemicals	Industrial chemicals	351
13	Other chemicals	Other chemicals (paints, medicines, soaps, cosmetics)	352
14	Petroleum	Petroleum refineries, products of petroleum and coal	353-354
15	Rubber	Rubber products	355
16	Plastic	Plastic products	356
17	Pottery	Pottery, china and earthenware	361
18	Glass	Glass and glass products	362
19	Other non-metallic	Other non-metallic mineral products (clay, cement)	369
20	Iron and steel	Iron and steel basic industries	371
21	Non-ferrous metal	Non-ferrous metal basic industries	372
22	Metal products	Fabricated metal products, except machinery and equipment	381
23	Mach. & equipment	Machinery and equipment except electrical	382
24	Electric. / Profess.	Electrical machinery apparatus, appliances and supplies & professional and scientific, measuring and controlling equipment	383-385
25	Transport	Transport equipment	384

A.2 Bils, Klenow and Ruane's (2018) method and results for Colombia

Here I succinctly introduce Bils, Klenow and Ruane's (2018) method to estimate the dispersion in the factors' MRP in the presence of additive measurement error in revenue and inputs, which in the latter case can be also interpreted as overhead factors. Define measured revenues and inputs for firm producing variety m as the sum of the "real" values plus an idiosyncratic measurement error: $\hat{R}_m = R_m + f_m$ and $\hat{I}_m = I_m + g_m$. Denote Δ the log-difference and $\blacktriangle$ the absolute difference. Bils, Klenow and Ruane (2018) find, under some reasonable assumptions, that the elasticity of $\Delta\hat{R}$ with respect to $\Delta\hat{I}$, $\hat{\beta} = \frac{\sigma_{\Delta\hat{R},\Delta\hat{I}}}{\sigma^2_{\Delta\hat{I}}}$, satisfies:

$$E\left\{\hat{\beta} \mid ln(TFPR_m)\right\} = \left[\Psi + \Lambda(ln(TFPR_m))^2\right]\left[1 - (1-\lambda)ln(TFPR_m)\right]$$

with $\lambda = \frac{\sigma^2_{ln\Theta}}{\sigma^2_{TFPR}}$, the ratio between the dispersion of the factor's MRP and the dispersion of the observed TFPR, our measure of interest; $\Psi = 1 + \Omega_\Theta - \Omega_{f'}$, where $\Omega_\Theta = \frac{\sigma_{\Delta\Theta,\Delta I}}{\sigma^2_{\Delta I}}$, $\Omega_{f'} = \frac{\sigma_{\blacktriangle f',\Delta\hat{I}}}{\sigma^2_{\Delta\hat{I}}}$, $\blacktriangle f' = \frac{\blacktriangle f_m}{\hat{I}_m}$; and Λ a constant that depends on the stochastic process of Θ, which is assumed is stationary. In the absence of measurement error ($\lambda = 1$) the elasticity of revenues with respect to inputs should be the same (Ψ) for plants with different average products. The quadratic term $\Lambda(ln(TFPR_m))^2$ is included to reflect the possibility of mean reversions in the stochastic process of Θ, given the stationary assumption. Therefore, λ can be estimated by GMM through the non-linear regression:

$$\begin{aligned}\Delta\hat{R}_m =& \phi ln(TFPR_m) + \Psi\Delta\hat{I}_m - \Psi(1-\lambda)ln(TFPR_m)\Delta\hat{I}_m \\ &+ \Gamma\left(ln(TFPR_m)\right)^2 + \Lambda(1-\lambda)\left(ln(TFPR_m)\right)^2\Delta\hat{I}_m \\ &+ \Upsilon\left(ln(TFPR_m)\right)^3 + \Lambda(1-\lambda)\left(ln(TFPR_m)\right)^3\Delta\hat{I}_m + \epsilon_m\end{aligned} \tag{A.1}$$

With Colombian data, I follow closely Bils, Klenow and Ruane (2018) for the construction of the variables. I estimate equation (A.1) by GMM sector by sector, controlling for year fixed effects, in the panel from 1991 to 1998. Standard errors are clustered at the firm-level. The last two columns in Table 5 show the point estimates for $\hat{\lambda}_s$ and its standard errors. For sectors in which the method does not deliver significative values, probably due to the influence of remaining outliers, I use the results from estimating (A.1) in the whole manufacturing sector controlling for a full set of sector-year fixed effects (as in Bils, Klenow and Ruane (2018)), values that are displayed in the last row.

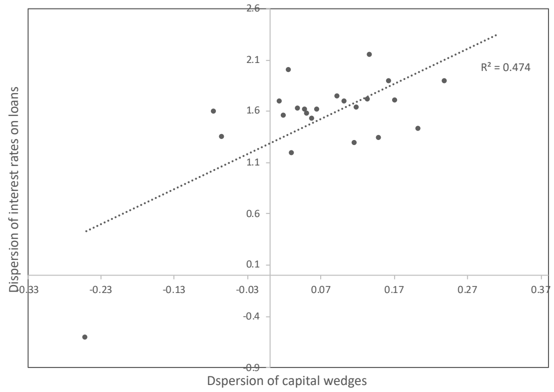
I use the estimated values of $\hat{\lambda}_s$ to compress the observed dispersions in the average revenue products of the factors to obtain variances and covariances of the MRP, and hence to derive $\hat{V}_{is}$.

A.3 External validation of misallocation measures

Here I use external data to check whether the computed dispersion in the intra-industry misallocation measures across sectors are related to a possible quantifiable source of such misallocation. For example, the intra-industry dispersion of capital wedges should be related to the dispersion of the idiosyncratic cost of capital for firms within the same industry. Thus, I can use data of the credit registry in Colombia to estimate the dispersion in the interest rates of new corporate loans by year in each manufacturing industry. This registry is done by the Colombian Financial Superintendency (*Superfinanciera*) and is available from 2002 to 2015.[51] The registry provides information at the bank-firm-loan level about the issuance date, amount disbursed, interest rate, maturity, among other variables for each corporate loan issued by each of the 38 commercial banks in the country. For each firm, I compute a weighted-average (by amounts disbursed) of the interest rates of firms' new loans, normalized by the term-premium of the Colombian sovereign debt to make comparable the different maturities of the loans across firms. I compute the dispersion of the interest rates on loans for each manufacturing industry and year. Due to the differences in the time periods covered by my panel of manufacturing plants and the credit registry, I control the measures of dispersion in both datasets for year fixed effects and compute the residual averages. Figure A.1 shows how the two measures are related. The correlation coefficient is 0.69, and the R-squared of the linear fit is around 0.47. The high correlation between the two measures validates the intuition that the magnitude of the dispersion of the computed factor wedges is associated to the extent of possible sources of factor misallocation.

[51] The data was made available to me by the Central Bank of Colombia. I'm grateful to Stefany Moreno who was in charge of the data cleaning and processing.

Figure A.1 – Intra-industry dispersion of capital wedges $(1 + \theta_k)$ and of interest rates on loans



Notes: The line represents the best linear fitting.

A.4 Solution of the model

To obtain the global solution of the system of equations, I employ both an algorithm to choose ideal initial conditions and a state-of-the-art solver for large-scale nonlinear systems. The proposed algorithm consists of the following three steps:

1. *Step 1:* I start solving the model for a two-country world composed by Colombia and an aggregate adding the rest of countries up (the number of equations is $N \times (S+L) = 56$). The purpose of this step is to find ideal initial conditions for Colombia and the rest of the world in step 2. To solve this two-country model I perform first a global search using particles swarm optimization a sufficient large number of times (500), to remove the influence of randomness in the initial position of the particles. Next, I use a local solver initialized in each of the 50 best solutions of the global search. For the local solver, I use auto-differentiation to obtain information about the gradient and the hessian of the objective function, and Knitro, a solver that implements both novel interior-point and active-set methods for solving large-scale nonlinear optimization problems.[52] The final solution is the best point of those 50 local solutions. It is worth to say that the obtained solution behaves according to the predictions of a small-open economy model, where the small country cannot influence foreign factor prices.
2. *Step 2:* Next, I solve the model $N-1$ times, in each case for a small-scale version of the world with the following three countries: Colombia, each country in the dataset and an aggregate adding the remaining countries up (the model is solved for $N \times (S+L) = 84$ equations 47 times). The objective of this step is to find ideal initial points for every country to solve the full model in step 3. In each of the $N-1$ times I initialize the local solver using for Colombia the solution found in step 1, and for the remaining two countries the solution for the rest of the world in step 1. I use the same local-solver and auto-differentiation as in step 1.
3. *Step 3:* Finally, I collect the solution for each country in step 2 to initialize the local solver for the model with the full set of countries; while for Colombia I initialize with a median of its $N-1$ solutions found in step 2 (such solutions have low dispersion). I use the same local-solver and auto-differentiation as in steps 1 and 2. The number of equations in this case is $N \times (S+L) = 1344$.

[52] I use auto-differentiation and the Knitro solver through the Tomlab optimization environment in Matlab.

B Mathematical derivations

B.1 Model solution under assumptions A.1 and A.2

Under assumptions A.1 and A.2, it is possible to express:

$$\sum_m^{M_{ijs}} (\frac{a_{im}}{\Theta_{im}})^{\sigma-1} = \frac{H_{is}}{d_{is}} \int_{\theta_i} \cdots \int_{\theta_{iL}} \int_{a^*_{ijs}(\Theta)}^{\infty} (\frac{a_{im}}{\Theta_{im}})^{\sigma-1} dG_{is} = \frac{H_{is}\kappa \bar{a}^{\kappa}_{is}}{d_{is}} \int_{\theta_{i1}} \cdots \int_{\theta_{iL}} \int_{a^*_{ijs}(\Theta)}^{\infty} a_{im}^{\sigma-\kappa-2} \Theta_{im}^{1-\sigma} dG^{\theta}_{is}$$

Using the formula of the cutoff function in (7), the last expression can be simplified as:

$$\sum_m^{M_{ijs}} (\frac{a_{im}}{\Theta_{im}})^{\sigma-1} = \frac{H_{is}}{d_{is}} \frac{\kappa}{1+\kappa-\sigma} (\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa} a^{*\sigma-1}_{ijs} \Gamma_{is} \tag{B.1}$$

with Γ_{is} defined as in the text. Applying the formulas for firm-level profits and revenues, the free entry condition can be restated as:

$$\sum_j^N \sum_m^{M_{ijs}} \frac{1}{\sigma} (\frac{\tau_{ijs}\Theta_{im}}{\rho a_{im}})^{1-\sigma} \omega_{is}^{-\sigma} E_{js} P_{js}^{\sigma-1} - \sum_j^N \sum_m^{M_{ijs}} \Theta_{im} f_{ijs} = f^e_{is} H_{is}$$

Notice that $\sum_m^{M_{ijs}} \Theta_{im} = \frac{H_{is}}{d_{is}} (\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa} \Gamma_{is}$. Combining with equation (B.1), it is possible to obtain:

$$\sum_j^N \frac{1}{\sigma} (\frac{\tau_{ijs}}{\rho})^{1-\sigma} \omega_{is}^{-\sigma} E_{js} P_{js}^{\sigma-1} \frac{1}{d_{is}} \frac{\kappa}{1+\kappa-\sigma} (\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa} a^{*\sigma-1}_{ijs} \Gamma_{is} - \sum_j^N \frac{f_{ijs}}{d_{is}} (\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa} \Gamma_{is} = f^e_{is}$$

Using the definition of the productivity cutoff value for the undistorted firms in (7) to substitute in $a^{*\sigma-1}_{ijs}$, the expression can be simplified to:

$$\sum_j^N (\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa} f_{ijs} = \frac{d_{is} f^e_{is} (1+\kappa-\sigma)}{\Gamma_{is}(\sigma-1)} \tag{B.2}$$

On the other hand, applying again (B.1) and the definition of the productivity cutoff value, bilateral exports $X_{ijs} = \sum_m^{M_{ijs}} r_{ijm}$ are given by:

$$X_{ijs} = \sum_m^{M_{ijs}} (\frac{\tau_{ijs}\Theta_{im}\omega_{is}}{\rho a_{im}})^{1-\sigma} E_{js} P_{js}^{\sigma-1} = \frac{\omega_{is} H_{is}}{d_{is}} \frac{\sigma\kappa}{1+\kappa-\sigma} (\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa} \Gamma_{is} f_{ijs} \tag{B.3}$$

Hence, from (B.2), sectoral revenues $R_{is} = \sum_j^N X_{ijs}$ are given by:

$$R_{is} = \frac{\kappa}{\rho} \omega_{is} f^e_{is} H_{is} \tag{B.4}$$

Free entry requires that the aggregate sectoral profits, Π_{is}, are equal to the expenditures in entry, $\omega_{is} f^e_{is} H_{is}$. This means the Pareto property of a constant profits/revenue ratio is not affected by distortions: $R_{is} = \frac{k}{\rho} \Pi_{is}$. From equations (11) and (12), the sectoral demand of primary factor l for both operational (fixed and variable costs) and entry uses is given by:

$$Z_{ils} = Z^o_{ils} + Z^e_{ils} = \frac{\rho \alpha_{ls} R_{is}}{w_{il}(1+\bar{\theta}_{ils})} + \frac{\alpha_{ls} \mathfrak{F}_{is}}{w_{il}(1+\bar{\theta}_{ils})} + \frac{\alpha_{ls} \omega_{is} f^e_{is} H_{is}}{w_{il}}$$

Substituting the expression for $\sum_m^{M_{ijs}} \Theta_{im}$ from above in the definition of $\mathfrak{F}_{is}$ and using again equation (B.4), it is straightforward to obtain equation (17), the total demand of primary factor l in terms of sector revenue, underlying factor prices and the HWA wedges. With the definition of v_{ils} as in the text, equation (21) is evident.

Finally, combining (B.3) with the gravity equation, I obtain:

$$X_{ijs} = \frac{X_{ijs}}{\sum_k^N X_{kjs}} E_{js} = \frac{\frac{\omega_{is} H_{is}}{d_{is}} (\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa} \Gamma_{is} f_{ijs}}{\sum_k^N \frac{\omega_{ks} H_{ks}}{d_{ks}} (\frac{\bar{a}_{ks}}{a^*_{kjs}})^{\kappa} \Gamma_{ks} f_{kjs}} E_{js}$$

By definition of the cutoff function in (7), it is possible to show the following relation between the cutoffs for the undistorted firms of country i and country i' for the same destination j:

$$\frac{a^*_{ijs}}{a^*_{i'js}} = (\frac{\tau_{ijs}}{\tau_{i'js}})(\frac{\omega_{is}}{\omega_{i's}})^{\frac{1}{\rho}}(\frac{f_{ijs}}{f_{i'js}})^{\frac{1}{\sigma-1}} \tag{B.5}$$

Using the formula in (B.5) into the denominator of bilateral exports, I obtain:

$$X_{ijs} = \frac{\frac{1}{d_{is}} \omega_{is}^{1-\frac{\kappa}{\rho}} H_{is} \bar{a}_{is}{}^{\kappa} (\frac{1}{\tau_{ijs}})^{\kappa} (f_{ijs})^{1-\frac{\kappa}{\sigma-1}} \Gamma_{is}}{\sum_k^N \frac{1}{d_{ks}} \omega_{ks}^{1-\frac{\kappa}{\rho}} H_{ks} \bar{a}^{\kappa}_{ks} (\frac{1}{\tau_{kjs}})^{\kappa} (f_{kjs})^{1-\frac{\kappa}{\sigma-1}} \Gamma_{ks}} E_{js}$$

Using (B.4) to substitute for the mass of entrants in terms of sectoral revenue, it simplifies to:

$$X_{ijs} = \frac{\omega_{is}^{-\frac{\kappa}{\rho}} R_{is} \phi_{ijs} \Gamma_{is}}{\sum_k^N \omega_{ks}^{-\frac{\kappa}{\rho}} R_{ks} \phi_{kjs} \Gamma_{ks}} E_{js} \tag{B.6}$$

where ϕ_{ijs} is as in the text. Hence, trade shares are given by (24). The model is closed combining (B.6) with the definitions of sectoral and aggregate revenues ($R_{is} = \sum_j^N X_{ijs}$ and $R_i = \sum_s^S R_{is}$), the Cobb-Douglas solution for sectoral expenditures, $E_{js} = \beta_{js} E_j$ and the trade balance condition: $E_j = \sum_s^S R_{js} - D_j$, which results on equation (23).

The system can be solved for the values of R_{is} for a given set of values of factor intensities α_{ls}, factor endowments $\bar{Z}_{il}$, expenditure shares β_{js}, aggregate trade deficits D_j, deep parameters ϕ_{ijs}, κ and ρ, and misallocation measures Γ_{is} and v_{ils}. Once the solution of R_{is} is computed, the values of all remaining variables can be found following the next sequence: i) factor prices and sectoral factor allocations from (21) and (22); ii) expenditures from the trade balance condition; iii) bilateral exports from (B.6); iv) mass of entrants from (B.4); v) bilateral cutoffs values for the undistorted firms from (B.3); vi) mass of operating firms from (9).

B.2 Demonstration of equation (18)

Here I deduce the formula for the ex-post HWA wedge in equation (18).

Proof. Starting by the definition of the HWA wedge:

$$(1+\bar{\theta}_{ils}) \equiv (\sum_j^N \sum_m^{M_{ijs}} \frac{1}{(1+\theta_{ilm})} \frac{c_{ijm}}{C_{is}})^{-1} = (\sum_j^N \sum_m^{M_{ijs}} \frac{1}{(1+\theta_{ilm})} \frac{\rho r_{ijm} + \omega_{is} \Theta_{im} f_{ijs}}{\rho R_{is} + \mathfrak{F}_{is}})^{-1}$$

Substituting firm level exports from i to j and after few algebraic manipulations we can write:

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}} = (\sum_{j}^{N}\sum_{m}^{M_{ijs}}\tfrac{\rho}{(1+\theta_{ilm})}(\tfrac{\tau_{ijs}\Theta_{im}\omega_{is}}{\rho a_{im}})^{1-\sigma}E_{js}P_{js}^{\sigma-1}+\tfrac{\omega_{is}\Theta_{im}f_{ijs}}{(1+\theta_{ilm})})^{-1}$$

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}} = (\rho(\tfrac{\omega_{is}}{\rho})^{1-\sigma}\sum_{j}^{N}\tau_{ijs}^{(1-\sigma)}E_{js}P_{js}^{\sigma-1}\sum_{m}^{M_{ijs}}\tfrac{1}{(1+\theta_{ilm})}(\tfrac{\Theta_{im}}{a_{im}})^{1-\sigma}+\omega_{is}\sum_{j}^{N}f_{ijs}\sum_{m}^{M_{ijs}}\tfrac{\Theta_{im}}{(1+\theta_{ilm})})^{-1}$$

Similar to how it is done in the precedent section, it is possible to show that: $\sum_{m}^{M_{ijs}}\frac{1}{(1+\theta_{ilm})}(\frac{\Theta_{im}}{a_{im}})^{1-\sigma} = \frac{M_{is}^{e}}{d_{is}}(\frac{\bar{a}_{is}}{a_{ijs}^{*}})^{\kappa}a_{ijs}^{*\sigma-1}\frac{\kappa\Gamma_{ils}}{1+\kappa-\sigma}$ and $\sum_{m}^{M_{ijs}}\frac{\Theta_{im}}{(1+\theta_{ilm})} = \frac{M_{is}^{e}\Gamma_{ils}}{d_{is}}(\frac{\bar{a}_{is}}{a_{ijs}^{*}})^{\kappa}$, with Γ_{ils} as in the text. Thus:

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}} = (\rho(\tfrac{\omega_{is}}{\rho})^{1-\sigma}\tfrac{M_{is}^{e}}{d_{is}}\sum_{j}^{N}\tau_{ijs}^{(1-\sigma)}E_{js}P_{js}^{\sigma-1}(\frac{\bar{a}_{is}}{a_{ijs}^{*}})^{\kappa}a_{ijs}^{*\sigma-1}\frac{\kappa\Gamma_{ils}}{1+\kappa-\sigma}+\omega_{is}\tfrac{M_{is}^{e}}{d_{is}}\sum_{j}^{N}f_{ijs}(\tfrac{\bar{a}_{is}}{a_{ijs}^{*}})^{\kappa}\Gamma_{ils})^{-1}$$

Substituting the definition of the productivity cutoff value for the undistorted firms in (7) in $a_{ijs}^{*\sigma-1}$, I obtain:

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}} = (\omega_{is}\tfrac{M_{is}^{e}}{d_{is}}\tfrac{(\sigma-1)\kappa\Gamma_{ils}}{1+\kappa-\sigma}\sum_{j}^{N}f_{ijs}(\tfrac{\bar{a}_{is}}{a_{ijs}^{*}})^{\kappa}+\omega_{is}\tfrac{M_{is}^{e}\Gamma_{ils}}{d_{is}}\sum_{j}^{N}f_{ijs}(\tfrac{\bar{a}_{is}}{a_{ij}^{s*}})^{\kappa})^{-1}$$

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}} = (\omega_{is}\tfrac{M_{is}^{e}}{d_{is}}\Gamma_{ils}\tfrac{\sigma\kappa+1-\sigma}{(1+\kappa-\sigma)}\sum_{j}^{N}f_{ijs}(\frac{\bar{a}_{is}}{a_{ijs}^{*}})^{\kappa})^{-1}$$

Using the free entry condition in (B.2):

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}} = (\omega_{is}M_{is}^{e}f_{is}^{e}\tfrac{\Gamma_{ils}}{\Gamma_{is}}\tfrac{\sigma\kappa+1-\sigma}{(\sigma-1)})^{-1}$$

Substituting the expression for $\sum_{m}^{M_{ijs}}\Theta_{im}$ given in Appendix B.1. in the definition of $\mathfrak{F}_{is}$ and using again equation (B.4) it is possible to show $\rho R_{is}+\mathfrak{F}_{is} = \omega_{is}M_{i}^{e}f_{i}^{e}\frac{\sigma\kappa+1-\sigma}{(\sigma-1)}$ and hence:

$$(1+\bar{\theta}_{ils}) = \frac{\Gamma_{is}}{\Gamma_{ils}} \qquad \square$$

It is possible to repeat the proof to derive an expression for the HWA wedge of the firms able to sell in each market j. Doing so, it follows $(1+\bar{\theta}_{ijls}) = (1+\bar{\theta}_{ils})$, this is, the HWA wedge does not vary across destinations. Even though this result looks at first glance counterintuitive, since this average it is not computed for the same set of firms (for example, $(1+\bar{\theta}_{iils})$ includes the firms that only sell in the domestic market, who must have, conditional on TFPQ, higher wedges than the firms exporting to j), the fact that in the HWA the inverse of the wedge is weighted by the cost share (firms that only sell in the domestic market have higher cost shares), makes possible this equalization.

B.3 Decomposition of industry-exporter fixed effect

From the definition of bilateral price index in equation (16), the double difference across sectors and exporters of the unit prices in each destination can be re-written in terms of the relative bilateral iceberg costs, number of exporters, average TFP and factor returns as:

$$(\frac{P_{ijs}P_{i'js'}}{P_{ijs'}P_{i'js}})^{1-\sigma} = (\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}})^{1-\sigma}(\frac{M_{ijs}M_{i'js'}}{M_{ijs'}M_{i'js}})(\frac{\bar{\psi}_{ijs}\bar{\psi}_{i'js'}}{\bar{\psi}_{ijs'}\bar{\psi}_{i'js}})^{1-\sigma}(\frac{A_{ijs}A_{i'js'}}{A_{ijs'}A_{i'js}})^{\sigma-1} \tag{B.7}$$

My interest is twofold. First, I will provide a proof of equation (19), and second I will decompose the industry-exporter fixed effect on single components that come from each of the mentioned sources. For this reason, in the next lines I develop the RHS of (B.7) keeping each term separated in square brackets, without simplifying across terms. Using the definitions of $\bar{\psi}_{ijs}$ and A_{ijs} in the text, equation (B.7) can be written as:

$$\begin{aligned}(\frac{P_{ijs}P_{i'js'}}{P_{ijs'}P_{i'js}})^{1-\sigma} =&[\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}}]^{1-\sigma}[\frac{M_{ijs}M_{i'js'}}{M_{ijs'}M_{i'js}}][\frac{\omega_{is}\omega_{i's'}\bar{\Theta}_{ijs}\bar{\Theta}_{i'js'}}{\omega_{is'}\omega_{i's}\bar{\Theta}_{ijs'}\bar{\Theta}_{i'js'}}]^{1-\sigma}\\ &[\frac{\bar{\Theta}_{ijs}M_{ijs}^{\frac{1}{1-\sigma}}(\sum_{m}^{M_{ijs}}(\frac{a_{im}}{\Theta_{im}})^{\sigma-1})^{\frac{1}{\sigma-1}}}{\bar{\Theta}_{ijs'}M_{ijs'}^{\frac{1}{1-\sigma}}(\sum_{m}^{M_{ijs'}}(\frac{a_{im}}{\Theta_{im}})^{\sigma-1})^{\frac{1}{\sigma-1}}}\frac{\bar{\Theta}_{i'js'}M_{i'js'}^{\frac{1}{1-\sigma}}(\sum_{m}^{M_{i'js'}}(\frac{a_{i'm}}{\Theta_{i'm}})^{\sigma-1})^{\frac{1}{\sigma-1}}}{\bar{\Theta}_{i'js}M_{i'js}^{\frac{1}{1-\sigma}}(\sum_{m}^{M_{i'js}}(\frac{a_{i'm}}{\Theta_{i'm}})^{\sigma-1})^{\frac{1}{\sigma-1}}}]^{\sigma-1}\end{aligned}$$

Using the expression for $\sum_{m}^{M_{ijs}}(\frac{a_{im}}{\Theta_{im}})^{\sigma-1}$ in equation (B.1) and the fact $\bar{\Theta}_{ijs} = \bar{\Theta}_{is}$ derived in Appendix C.2, this reduces to:

$$\begin{aligned}(\frac{P_{ijs}P_{i'js'}}{P_{ijs'}P_{i'js}})^{1-\sigma} =&[\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}}]^{1-\sigma}[\frac{M_{ijs}M_{i'js'}}{M_{ijs'}M_{i'js}}][\frac{\omega_{is}\omega_{i's'}\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\omega_{is'}\omega_{i's}\bar{\Theta}_{is'}\bar{\Theta}_{i's'}}]^{1-\sigma}\\ &[(\frac{\bar{\Theta}_{is}\bar{\Theta}_{i's'}a^*_{ijs}a^*_{i'js'}}{\bar{\Theta}_{is'}\bar{\Theta}_{i's'}a^*_{ijs'}a^*_{i'js}})^{\sigma-1}(\frac{M_{ijs'}M_{i'js}}{M_{ijs}M_{i'js'}})(\frac{\Gamma_{is}\Gamma_{i's'}}{\Gamma_{is'}\Gamma_{i's}})\frac{\frac{H_{is}}{d_{is}}(\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa}}{\frac{H_{is'}}{d_{is'}}(\frac{\bar{a}_{is'}}{a^*_{ijs'}})^{\kappa}}\frac{\frac{H_{i's'}}{d_{i's'}}(\frac{\bar{a}_{is'}}{a^*_{i'js'}})^{\kappa}}{\frac{H_{i's}}{d_{i's}}(\frac{\bar{a}_{is}}{a^*_{i'js}})^{\kappa}}]\end{aligned}$$

Under assumptions A.1 and A.2. the aggregate stability condition (9) can be solved to obtain $M_{ijs} = \frac{H_{is}\Upsilon_{is}}{\delta_{is}}(\frac{\bar{a}_{is}}{a^*_{ijs}})^{\kappa}$ with $\Upsilon_{is} = \int_{\theta_{i1}} ... \int_{\theta_{iL}} \Theta_i^{-\frac{k}{\rho}} dG^{\theta}_{is}(\vec{\theta})$, an expected value that depends only on the joint distribution of distortions. Substituting this expression in the first and third terms, and using equation (B.5), I obtain for the RHS:

$$\begin{aligned}=&[\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}}]^{1-\sigma}[\frac{d_{is'}d_{i's}}{d_{is}d_{i's'}}\frac{H_{is}H_{i's'}}{H_{is'}H_{i's}}\frac{\Upsilon_{is}\Upsilon_{i's'}}{\Upsilon_{is'}\Upsilon_{i's}}(\frac{\bar{a}_{is}\bar{a}_{i's'}}{\bar{a}_{is'}\bar{a}_{i's}})^{\kappa}(\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}})^{-\kappa}(\frac{\omega_{is}}{\omega_{is'}}\frac{\omega_{i's'}}{\omega_{i's}})^{-\frac{\kappa}{\rho}}(\frac{f_{ijs}f_{i'js'}}{f_{ijs'}f_{i'js}})^{\frac{-\kappa}{\sigma-1}}]\\ &[\frac{\omega_{is}\omega_{i's'}\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\omega_{is'}\omega_{i's}\bar{\Theta}_{is'}\bar{\Theta}_{i's'}}]^{1-\sigma}[(\frac{\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\bar{\Theta}_{is'}\bar{\Theta}_{i's'}})^{\sigma-1}\frac{\Gamma_{is}\Gamma_{i's'}}{\Gamma_{is'}\Gamma_{i's}}\frac{\Upsilon_{is'}\Upsilon_{i's}}{\Upsilon_{is}\Upsilon_{i's'}}(\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}})^{\sigma-1}(\frac{\omega_{is}\omega_{i's'}}{\omega_{is'}\omega_{i's}})^{\sigma}(\frac{f_{ijs}f_{i'js'}}{f_{ijs'}f_{i'js}})]\end{aligned}$$

Using $H_{is} = \frac{R_{is}}{\omega_{is}f^e_{is}}$ and applying logs to separate the components that only depend on exporter-industry terms and simplifying, I finally obtain for the RHS of (B.7):

$$\begin{aligned}=&log[\frac{\varrho_{is}\varrho_{i's'}}{\varrho_{is'}\varrho_{i's}}\frac{R_{is}R_{i's'}}{R_{is'}R_{i's}}\frac{\Upsilon_{is}\Upsilon_{i's'}}{\Upsilon_{is'}\Upsilon_{i's}}(\frac{\omega_{is}}{\omega_{is'}}\frac{\omega_{i's'}}{\omega_{i's}})^{-\frac{\kappa}{\rho}-1}] + log[\frac{\omega_{is}\omega_{i's'}\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\omega_{is'}\omega_{i's}\bar{\Theta}_{is'}\bar{\Theta}_{i's'}}]^{1-\sigma}\\ &+ log[(\frac{\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\bar{\Theta}_{is'}\bar{\Theta}_{i's'}})^{\sigma-1}(\frac{\omega_{is}}{\omega_{is'}}\frac{\omega_{i's'}}{\omega_{i's}})^{\sigma}\frac{\Gamma_{is}\Gamma_{i's'}}{\Gamma_{is'}\Gamma_{i's}}\frac{\Upsilon_{is'}\Upsilon_{i's}}{\Upsilon_{is}\Upsilon_{i's'}}] + B_{ijs}\end{aligned} \tag{B.8}$$

where $B_{ijs} = ln[(\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}})^{-\kappa}(\frac{f_{ijs}f_{i'js'}}{f_{ijs'}f_{i'js}})^{1-\frac{\kappa}{\sigma-1}}]$ and $\varrho_{is} = \frac{\bar{a}^{\kappa}_{is}}{d_{is}f^e_{is}}$. Canceling out the double differences of $\bar{\Theta}_{is}$ and Υ_{is} across terms and simplifying the double differences of ω_{is} it is straightforward to derive the gravity equation in (19). Furthermore, equation (B.8) offers a decomposition of the exporter-industry fixed effect on the three sources of interest: number of exporters (first term in log), average factor returns (second term in log) and TFP (third term in log).

This decomposition is used in section 3.3 as follows. Denote $\tilde{x}$ the value in the allocative efficient equilibrium of x, and $\check{x} \equiv \frac{x}{\tilde{x}}$ the proportional change when we introduce distortions. Thus figure 3 plots in each chart the following terms:

$$log(\frac{\check{X}_{ijs}\check{X}_{i'js'}}{\check{X}_{ijs'}\check{X}_{i'js}}) = log\frac{\check{R}_{is}\check{R}_{i's'}}{\check{R}_{is'}\check{R}_{i's}}\frac{\Upsilon_{is}\Upsilon_{i's'}}{\Upsilon_{is'}\Upsilon_{i's}}(\frac{\check{\omega}_{is}}{\check{\omega}_{is'}}\frac{\check{\omega}_{i's'}}{\check{\omega}_{i's}})^{-\frac{\kappa}{\rho}-1} + log(\frac{\check{\omega}_{is}\check{\omega}_{i's'}\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\check{\omega}_{is'}\check{\omega}_{i's}\bar{\Theta}_{is'}\bar{\Theta}_{i's'}})^{1-\sigma}$$
$$+ log(\frac{\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\bar{\Theta}_{is'}\bar{\Theta}_{i's'}})^{\sigma-1}(\frac{\check{\omega}_{is}}{\check{\omega}_{is'}}\frac{\check{\omega}_{i's'}}{\check{\omega}_{i's}})^{\sigma}\frac{\Gamma_{is}\Gamma_{i's'}}{\Gamma_{is'}\Gamma_{i's}}\frac{\Upsilon_{is'}\Upsilon_{i's}}{\Upsilon_{is}\Upsilon_{i's'}}$$

with $i = 1$, $i' = 2$, $j = 2$, $s = 1$, $s' = 2$.

B.4 Solution for Γ_{is} under log-normal

By definition of Γ_{ils} in the text:

$$\Gamma_{is} = \int_{\theta_i} \dots \int_{\theta_{iL}} \Theta_i{}^{1-\frac{\kappa}{\rho}} dG^{\theta}_{is} = E(\prod_{l}^{L}(1+\theta_{il})^{(1-\frac{\kappa}{\rho})\alpha_{ls}})$$

Assume $\vec{\theta_{is}} = \{\theta_{i1s}, \theta_{i2s}, \dots\theta_{iLs}\}$ has a multivariate log-normal distribution, such the transformed vector $\vec{\theta^*_{is}} = \{\ln(\theta_{i1s}), \ln(\theta_{i2s}), \dots \ln(\theta_{iLs})\}$ has a multivariate normal distribution with expected value $\vec{\mu_{is}}$ ($1{\times}L$ vector) and variance V_{is} ($L{\times}L$ matrix). Let $\vec{\alpha_s}$ a (column) vector with elements: $\vec{\alpha}_s = \{(1-\frac{\kappa}{\rho})\alpha_{1s}, (1-\frac{\kappa}{\rho})\alpha_{2s}, \dots, (1-\frac{k}{\rho})\alpha_{Ls}\}'$. Then the product $\prod_{l}^{L}(1+\theta_{il})^{(1-\frac{\kappa}{\rho})\alpha_{ls}}$ is log-normal distributed with location parameter $(\vec{\alpha}_s)'\vec{\mu_{is}}$ and shape parameter $(\vec{\alpha_s})'V_{is}\vec{\alpha_s}$. Under log-normality, the required expected value is then:

$$\Gamma_{is} = \exp\left[(\vec{\alpha_s})'\vec{\mu_{is}} + \frac{1}{2}(\vec{\alpha_s})'V_{is}\vec{\alpha_s}\right]$$

On the other hand, the definition of Γ_{ils} in the text:

$$\Gamma_{ils} = \int_{\theta_i} \dots \int_{\theta_{iL}} \frac{\Theta_i^{1-\frac{\kappa}{\rho}}}{(1+\theta_{ils})} dG^{\theta}_{is} = E[(1+\theta_{il})^{(1-\frac{\kappa}{\rho})\alpha_{ls}-1}\prod_{h\neq l}^{L}(1+\theta_{ih})^{(1-\frac{\kappa}{\rho})\alpha_{hs}}]$$

By the same token, let $\vec{\alpha_{ls}}$ a (column) vector with elements: $\vec{\alpha_{ls}} = \{(1-\frac{\kappa}{\rho})\alpha_{1s}, \dots, (1-\frac{\kappa}{\rho})\alpha_{ls} - 1, \dots, (1-\frac{\kappa}{\rho})\alpha_{Ls}\}'$. This is, $\vec{\alpha_{ls}}$ has the same elements of $\vec{\alpha_s}$ with exception to the element in position l, which is $(1-\frac{k}{\rho})\alpha_{ls} - 1$. Thus the product $(1+\theta_{il})^{(1-\frac{\kappa}{\rho})\alpha_{ls}-1}\prod_{h\neq l}^{L}(1+\theta_{ih})^{(1-\frac{\kappa}{\rho})\alpha_{hs}}$ is log-normal distributed with location parameter $(\vec{\alpha_{ls}})'\vec{\mu_{is}}$ and shape parameter $(\vec{\alpha_{ls}})'V_{is}\vec{\alpha_{ls}}$. Accordingly, its expected value is:

$$\Gamma_{ils} = \exp\left[(\vec{\alpha_{ls}})'\vec{\mu_{is}} + \frac{1}{2}(\vec{\alpha_{ls}})'V_{is}\vec{\alpha_{ls}}\right]$$

Now, using the formula for $(1 + \bar{\theta}_{ils})$ in (18) we obtain:

$$\ln(1 + \bar{\theta}_{ils}) = (\vec{\alpha_s})'\vec{\mu}_{is} + \frac{1}{2}(\vec{\alpha_s})'V_{is}\vec{\alpha_s} - (\vec{\alpha_{ls}})'\vec{\mu_{is}} - \frac{1}{2}(\vec{\alpha_{ls}})'V_{is}\vec{\alpha_{ls}}$$
$$= \mu_{ils} + \frac{1}{2}[(\vec{\alpha_s})'V_{is}\vec{\alpha_s} - (\vec{\alpha_{ls}})'V_{is}\vec{\alpha_{ls}}] \tag{B.9}$$

B.5 Welfare

Combining the formula of the consumer price index in sector s and equation (B.1) we obtain:

$$\left(P_{is}^d\right)^{1-\sigma} = \sum_k^N P_{kis}^{1-\sigma} = \sum_k^N \frac{\tau_{kis}}{\rho}\omega_{ks} \sum_m^{M_{kis}} (\frac{a_{km}}{\Theta_{km}})^{\sigma-1} = \sum_k^N \frac{\tau_{kis}}{\rho} \frac{\omega_{ks}H_{ks}}{d_{ks}} \frac{\kappa}{1+\kappa-\sigma}(\frac{\bar{a}_{ks}}{a^*_{kis}})^{\kappa} a_{kis}^{*\sigma-1}\Gamma_{ks}$$

Inserting the definition of the productivity cutoff value for the undistorted firms in (7) in the term $a_{kis}^{*\sigma-1-\kappa}$, the price index can be written as:

$$\left(P_{is}^d\right)^{-\kappa} = E_{is}^{\frac{-\kappa}{1-\sigma}-1} \sum_k^N (\frac{\tau_{kis}}{\rho})^{-\kappa} \omega_{ks}^{1-\frac{\kappa}{\rho}} \frac{H_{ks}}{d_{ks}} \frac{\kappa}{1+\kappa-\sigma} (\bar{a}_{ks})^{\kappa} (\sigma f_{kis})^{1-\frac{\kappa}{\sigma-1}} \Gamma_{ks}$$

Using the country i's share of expenditure on itself within sector s from equation (B.6), we obtain:

$$\left(P_{is}^d\right)^{-\kappa} = \varsigma_{ijs} E_{is}^{\frac{-\kappa}{1-\sigma}-1} \omega_{is}^{-\frac{\kappa}{\rho}} R_{is}\Gamma_{is}(\frac{1}{\pi_{iis}})$$

where $\varsigma_{ijs} = (\frac{\rho \bar{a}_{is}}{\tau_{ijs}})^{\kappa} \frac{1}{d_{is} f_i^e} (\frac{1}{f_{iis}})^{1-\frac{\kappa}{\sigma-1}} (\frac{\kappa}{1+\kappa-\rho})$ a term that does not vary in the counterfactual exercise. Hence, the proportional change of the price index from the initial equilibrium to the counterfactual one can be written as:

$$\hat{P}_{is}^d = \hat{E}_{is}^{\frac{1}{1-\sigma}+\frac{1}{k}} \hat{\omega}_{is}^{\frac{1}{\rho}} \hat{R}_{is}^{-\frac{1}{\kappa}} \hat{\Gamma}_{is}^{-\frac{1}{\kappa}} (\hat{\pi}_{iis}^{\frac{1}{k}})$$

Using the fact that $\hat{P}_i^d = \prod_s (\hat{P}_{is}^d)^{\beta_s}$, $\hat{E}_{is} = \hat{E}_i$ and equation (25) to substitute $\hat{\omega}_{is}$, the derivation of equation (28) is straightforward. Moreover, notice that in the case of the undistorted economy with one factor production, $\hat{R}_{is} = \hat{\omega}_{is}\hat{Z}_{is}$ and $\hat{\omega}_{is} = \hat{w}_i = \hat{E}_i$ so the increase in the sectoral price index is $\hat{P}_{is}^d = \hat{w}_i(\frac{\hat{\pi}_{iis}}{\hat{Z}_{is}})^{\frac{1}{k}}$, which leads to the Arkolakis, Costinot and Rodríguez-Clare (2012)'s formula to compute the increase in welfare in response to any exogenous shock.

C Intra- and inter-industry misallocation in the HK's economy

In this Appendix I compute the TFP gains from removing intra- and inter-industry factor misallocation in the case of a closed economy with a fixed number of operating firms, following the same notation as in Section 2.

C.1 Contribution of inter-industry misallocation

Denote the TFPQ and TFPR of firm producing variety m as a_m and ψ_m, respectively, and ξ_{lm} the MRP of the input l. Let $\bar{\xi}_{ls}$ denote the HWA of ξ_{lm}, with weights given by the participations of firm's revenues in total industry revenue. Note that $\bar{\xi}_{ls} = (1+\bar{\theta}_{ls})\frac{w_l}{\rho}$. Using the cost minimization condition of the CD aggregator across sectors, total demand of factor-l in industry s can be expressed as:

$$Z_{ls} = \frac{\alpha_{ls}\beta_s/\bar{\xi}_{ls}}{\sum\limits_s^S \alpha_{ls}\beta_s/\bar{\xi}_{ls}}\bar{Z}_l \tag{C.1}$$

where $\bar{Z}_l \equiv \sum\limits_s^S Z_{ls}$ correspond to the fixed endowment of factor-l in the economy.Standard aggregation under monopolistic competition leads to an industry production of the form $Q_s = A_s M_s^{\frac{1}{\sigma-1}} \prod\limits_l^L {Z_{ls}}^{\alpha_{ls}}$, where sectoral TFP A_s can be derived from firm-level data from:

$$A_s^{\sigma-1} = \frac{1}{M_s}\sum_m^{M_s}\left(\frac{a_m\bar{\psi}_s}{\psi_m}\right)^{\sigma-1} \tag{C.2}$$

where $\bar{\psi}_s$ is the sectoral revenue productivity. If a reform equalizes TFPR across firms, the sectoral (efficient) TFP is simply the power mean of physical productivities: $\widetilde{A}_s^{\sigma-1} = \tilde{M}_s^{-1}\sum\limits_m^{\tilde{M}_s} a_m^{\sigma-1}$. With the assumption of no self-selection of firms, $\tilde{M}_s = M_s$ and the percentage gains on sectoral TFP due to TFPR equalization are:

$$Gains_s^{intra} = 100(\frac{\widetilde{A}_s}{A_s} - 1) = 100((\sum_m^{M_s}(\frac{a_m\bar{\psi}_s}{\widetilde{A}_s\psi_m})^{\sigma-1})^{\frac{1}{1-\sigma}} - 1) \tag{C.3}$$

Equation (C.3) is the cornerstone of HK's counterfactual exercise, and the description until here provided summarizes the main features of HK's model. The gains from removing intra-industry misallocation in (C.3) are the same if the reform equalizes firms' TFPR to $\bar{\psi}_s$, so the factors' MRP are equal to their HWA in the industry, or to the inter-industry efficient allocation, in which case the factors' MRP are equated to $\frac{w_l}{\rho}$. However, only in the first case it is ensured there are no factor reallocations across sectors (which is evident from equation C.1), so the sectoral TFP gains in equation (C.3) are identical to the gains in industry output, $100(\frac{\widetilde{Q}_s}{Q_s} - 1)$. In this specific case, total output gains in the economy can be computed simply by aggregating sectoral productivities up using the CD aggregator across industries:

$$Gains^{intra} = 100(\prod_s^S(\frac{\widetilde{A}_s}{A_s})^{\beta_s} - 1) \tag{C.4}$$

Clearly, total gains in (C.4) are only due to resource reallocation within industries: by assumption, there are not factor reallocations across sectors. In this case, there is MRP equalization within industries, but not necessarily across them. In the more general case in which I impose MRP equalization not only within but across industries (i.e. removing all wedges), sectoral TFP gains are the same as in (C.3), but output gains in each industry are no longer equal to the corresponding TFP gains, due to factor reallocation across sectors. From (C.1), the allocative efficient demand of factors at the industry level is given by $\widetilde{Z}_{ls} = \alpha_{ls}\beta_s\bar{Z}_l/\sum_s^S \alpha_{ls}\beta_s$.[53] Industry's output in frictionless factor markets is given by $\tilde{Q}_s = \tilde{A}_s\tilde{M}_s^{\frac{1}{\sigma-1}}\prod_l^L \tilde{Z}_{ls}{}^{\alpha_{ls}}$. Thus, the variation in sectoral output due to a reform that removes all wedges is a consequence of both a rise in the TFP and a variation in the use of factors in the whole sector, which depends exclusively on the sign of $\bar{\theta}_{ls}$ (the extent of inter-industry misallocation). At the aggregate level, factor endowments between the distorted economy and the allocative efficient counterfactual are kept constant. So any change in aggregate output Q is attributable to variations in the aggregate TFP, and it is due to resource reallocation, both within and between industries. Gains in aggregate TFP can be caused by increases in sectoral TFP, term denoted $Gains^{intra}$ above, or by reallocation of factors between industries, given by:

$$Gains^{inter} = 100\left(\prod_s^S\prod_l^L \frac{\widetilde{Z}_{ls}{}^{\alpha_{ls}\beta_s}}{Z_{ls}{}^{\alpha_{ls}\beta_s}} - 1\right) = 100\left(\prod_s^S\prod_l^l \left[\frac{\sum_s^S(\alpha_{ls}\beta_s/\bar{\xi}_{ls})}{(\sum_s^S\alpha_{ls}\beta_s)/\bar{\xi}_{ls}}\right]^{\alpha_{ls}\beta_s} - 1\right) \tag{C.5}$$

Where I use equation (C.1) and the expression for $\widetilde{Z}_{ls}$ to obtain the explicit closed-form solution. Thus, inter-industry gains only depend on the industry average MRP interacted with technological parameters, a plain consequence of the sectoral demand of factors in equation (C.1). These gains can be computed only with industry-level data, a fact that allows me to make cross-country comparisons to evaluate whether this component also explains the TFP gaps observed across countries, an exercise that is performed below. Finally, total gains in the economy, given by the variation on total output (or aggregate TFP), are a combination of both sources of gains:

$$Gains = 100\left(\frac{\widetilde{Y}}{Y} - 1\right) = 100\left[\left(\frac{Gains^{inter}}{100} + 1\right)\left(\frac{Gains^{intra}}{100} + 1\right) - 1\right] \tag{C.6}$$

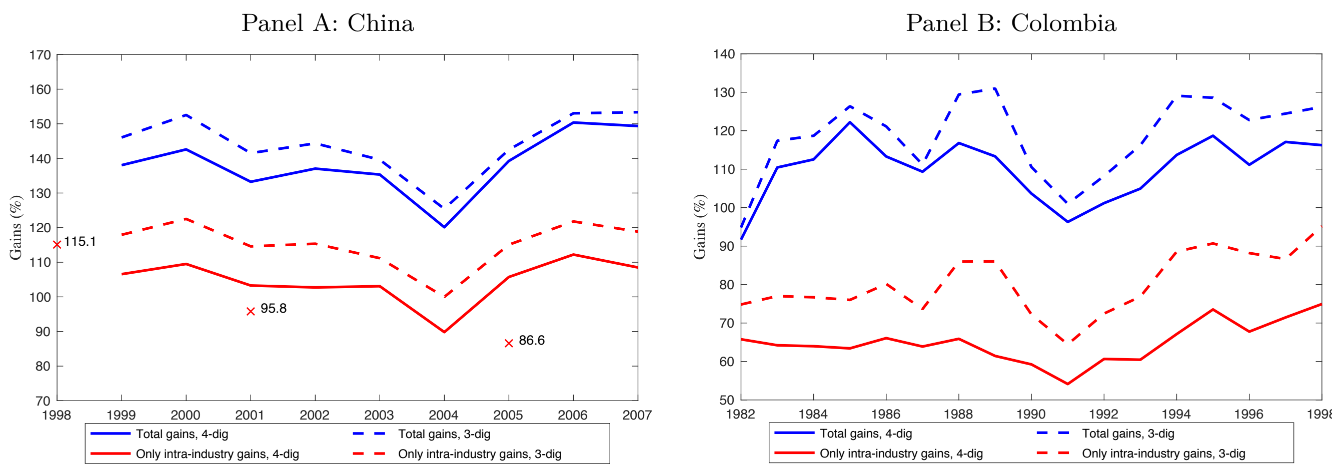
The importance of each type of misallocation depends, of course, on the considered industry aggregation. For example, in the extreme case in which the whole manufacturing sector is represented as a single industry, the entire TFP loss due to allocative inefficiency proceeds from the intra-industry type, whereas in the opposite extreme, the whole loss proceeds from the inter-sectoral type. Using a 4-digit ISIC industry classification,[54] a value added specifica-

[53] This is, in the case that all sectors have the same revenue shares, the efficient allocation of factors across sectors implies that more intensive industries should have a larger proportion of the corresponding factor. Similarly, in the case that all sectors have the same factor intensities, the factors should be allocated in proportion only on sectoral revenue shares. The efficient factor allocation across industries is the combination of these two forces.

[54] For the 4-digit classification in the Colombian case, due to small number of observations, 14 industries were reclassified to its closest 4-digit industry or to the 4-digit sector within the same 3-digit industry that merges the products not elsewhere classified.

tion for the production function, and average US cost shares at the corresponding aggregation level from the NBER-CES Manufacturing Industry Database during the same period, the same set of specifications than the used in HK's baseline, I find that the inter-sectoral component contributes on average up to 35% of the total reallocation gains of a comprehensive reform that removes all factor misallocation in Colombia, for the period 1982-1998. As a robustness check, I replicate the exercise with firm-level data from China, a country that offers external validation using the calculations provided by HK.[55] In Figure C.2 I report using continue lines the total gains (blue) and the intra-sectoral gains (red) from removing distortions for both countries, when the 4-digit ISIC industry aggregation is used. The difference between both lines is due to the gains from inter-sectoral reallocation. For China I find similar TFP gains as in HK in the case of removing only intra-industry misallocation, and an average contribution of 30% of the inter-sectoral component for the complete reform.

Figure C.2 – TFP gains from factor reallocation in a closed economy



Note: In Panel A, × correspond to the values found by HK.

In general, gains from removing distortions are larger for China, although the time periods are not comparable. The graph shows that over time in both countries there are not significant improvements in allocative efficiency in the considered periods; indeed, there is a slight worsening at the end of each one. When I move to the 3-digit ISIC classification, the predictions from the decomposition seem to hold. The dashed lines in Figure C.2 report once again the total gains (blue) and the intra-sectoral gains (red) from removing distortions, but now at the 3-digit ISIC classification. Both total gains fluctuate around a similar range. However, the intra-industry gains rise in a larger proportion than the total gains, so their average contribution is now 68% and 73% for Colombia and China, respectively. This confirms that as the level of disaggregation increases, the intra-industry gains are lower.

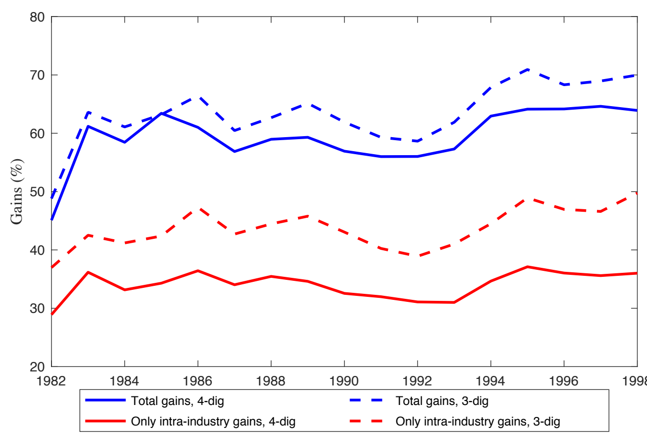
[55]For China, I use the panel from the Annual Survey of Industrial Production collected by the Chinese government's National Bureau of Statistics, for the period 1999-2007

C.2 Robustness checks

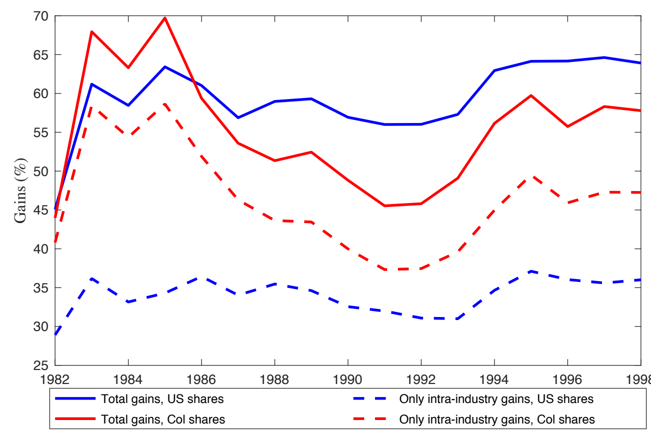
The source of inter-industry gains is neither related to the use of US cost shares instead of domestic factor intensities in the sectoral production function nor to the use of a value-added specification. For example, Figure C.3 displays for the Colombian case that using a gross-output specification (Panel A) or changing the production function coefficients for Colombian cost shares (Panel B) does not alter importantly the key insights. In the latter case, factor intensities are now equal to the observed share costs, but they are still different to the optimal share cost in monopolistic competition (where the total cost is ρ times the revenue), which is what matters in the efficient allocation. However, the use of Colombian cost shares reduces the relative importance of inter-sectoral reallocation: its average contribution shrinks to 23%.

Figure C.3 – Sensitivity to production function specification and factor intensities

Panel A : TFP gains using gross output specification (Colombia, US cost shares)



Panel B : TFP gains by set of cost shares (Colombia, 4-dig, gross output specification)



Further, the total gains and the contribution of the inter-sectoral component increase using a higher elasticity of substitution across sectors. This is completely in line with the HK prediction that when sectors outputs are better substitutes, inputs are reallocated toward sectors with bigger productivity gains, so there are larger TFP gains. We can show this with a CES demand across sectors. In this case, there is not a closed-form solution for each component, but it is possible to implement a numerical procedure to obtain both gains[56].

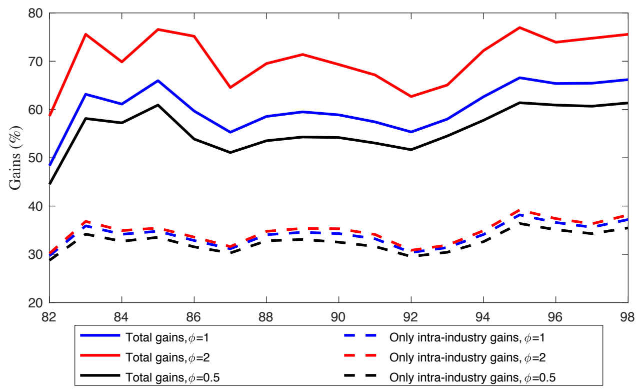
[56]With a CES aggregator of the form $Y^{\varphi} = \sum_{s}^{S} \beta_s {Y_s}^{\varphi}$, where $\varphi = \frac{\phi-1}{\phi}$ and ϕ is the elasticity of substitution across sectors, the sectoral factor demand is now:

$$Z_{ls} = \frac{\alpha_{ls}\beta_s^{\phi}P_s^{1-\phi}/\bar{\xi}_{ls}}{\sum_{s}^{S}\alpha_{ls}\beta_s^{\phi}P_s^{1-\phi}/\bar{\xi}_{ls}}\bar{Z}_l$$

Thus, in the allocative efficient inter-industry allocation, not only factor intensities and revenue shares play a role, but also the efficient sectoral price indexes as indicators of productivity. The direction and strength of their influence depends on the magnitude of ϕ. For $\phi > 1$ ($\phi < 1$), if factor intensities and shares of sectoral revenue are constant across sectors, factors should be allocated to more (less) productive sectors. The interaction of these three sectoral forces (factor intensities, revenue shares and aggregate productivities) is what determines the efficient inter-sectoral allocation. Notice that to find $\widetilde{Z}_{ls}$ it is necessary to solve for $\tilde{P}_s$,

Figure C.4 shows that for different values of the elasticity of substitution across sectors (ϕ), the components of the gains behave as predicted. The numerical procedure replicates the results of the close-form solutions for the CD aggregator for both components in the case $\phi = 1$, whereas total gains and the contribution of the inter-sectoral component increases when $\phi = 2$ (up to 50% from 43% in the latter case) and decreases when $\phi = 0.5$ (to 36% in the latter case). In those exercises the change in the intra-sectoral gains is negligible.

Figure C.4 – Sensitivity to elasticity of substitution across sectors



C.3 Inter-industry misallocation and development

Another important question about the relevance of inter-industry misallocation is whether its associated TFP loss is larger in less developed economies, as is the case with intra-industry misallocation, the core result of HK's paper. If the inter-sectoral gains vary systematically across countries, omitting the inter-sectoral component implies an under-estimation of the TFP gap attributed to factor misallocation, if the latter is computed only with intra-industry reforms, as in HK. In the case of the CD aggregator across sectors, the closed form solution for the TFP gains of removing inter-industry misallocation only requires information at the industry level. Thus, I use information from the socio-economic accounts of the World Input Output Database - WIOD (Timmer et al. (2015)), which contains industry-level data for 40 countries and 35 industries mostly at the 2-digit ISIC level, covering the overall economy, to compute those gains.

Figure C.5 presents how the gains from inter-sectoral reallocation vary with the GDP per capita by country.[57] For this calculation, I use a gross output specification for the sectoral production function with 3 inputs (hours worked, capital and materials) and US cost shares. The linear correlation between both variables in this baseline is -0.75 (Figure C.5 also shows the best linear fit). The negative correlation is robust to the use of value added specification

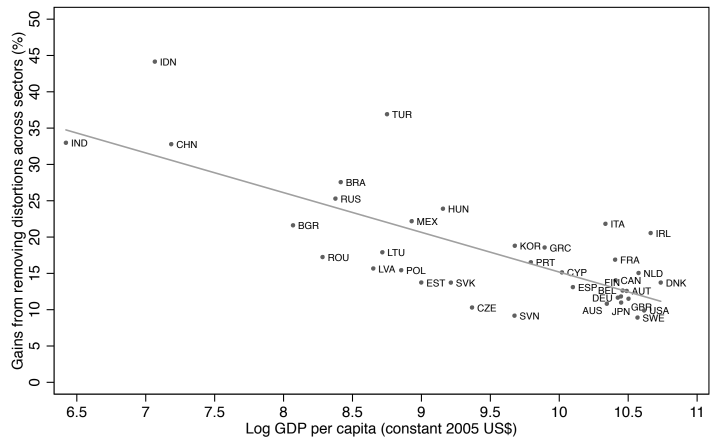
---

which implies to find firm's output prices in the efficient allocation. These prices can be obtained by solving the non-linear system that includes all firm-level prices, through numerical optimization. Once $\widetilde{Z}_{ls}$ are obtained, it is simple to calculate both gains from removing misallocation, using the counterfactual aggregate output generated by $\widetilde{A}_s$ and Z_{ls}.

[57]Each dot corresponds to the average value between 1995 and 2007 of the intersectoral gains calculated using (C.5) for each country and the average GDP per capita in constant 2005 US dollars obtained from the World Bank. The results are very similar if median values are used. Two small countries with many zeros in sectoral data were dropped from the WIOD sample (Luxembourg and Malta). Likewise, Taiwan was dropped to make comparable WIOD and World Bank data.

or own country's cost shares in the production function; to restrict the set of sectors to only manufacturing industries and to measure labor with the wage bill and materials in nominal values to control for heterogeneity in labor and for differences in quality of intermediate inputs respectively, graphs shown in Figure C.6. Therefore, there is evidence that less developed economies tend to have greater inter-sectoral gains for removing distortions. This is consistent with the insights of multi-country studies as Tombe (2015) or Święcki (2017) which focus on inter-sectoral misallocation, that find larger intersectoral distortions in poor countries. Thus, omitting the inter-sectoral component of the total gains from removing distortions understates the common TFP gaps attributed to firm-level misallocation.

Figure C.5 – TFP gains from removing inter-industry misallocation and GDP per capita



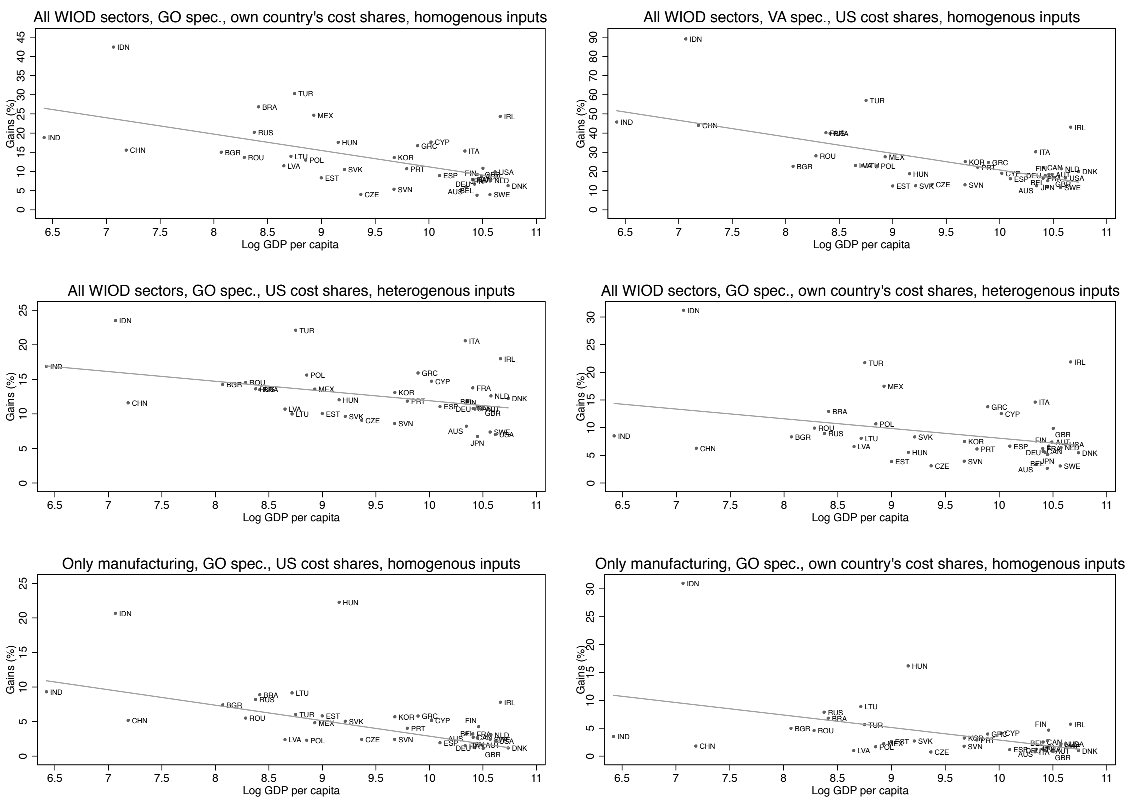
Note: Each dot corresponds to the average gains from removing inter-industry misallocation and the corresponding average GDP per capita in the period 1991-2007. The source of the data is WIOD and the World Bank development indicators.

If the inter-sectoral gains vary systematically across countries, omitting the inter-sectoral component implies an under-estimation of the TFP gap attributed to factor misallocation, if the latter is computed only with intra-industry reforms, as in HK. Since the inter-industry gains could be calculated with sectoral data, I use information from the socio-economic accounts of the World Input Output Database - WIOD (Timmer et al., 2015), which contains industry-level data for 40 countries and 35 industries mostly at the 2-digit ISIC level, covering the overall economy, to compute this dimension. Figure C.5 presents how the gains from inter-sectoral reallocation vary with the GDP per capita by country[58]. For this calculation, I use a gross output specification for the sectoral production function with 3 inputs (hours worked, capital and materials) and US cost shares. The linear correlation between both variables in this baseline is -0.75 (Figure C.5 also shows the best linear fit). The negative correlation is robust to the use of value added specification or own country's cost shares in the production function; to restrict the set of sectors to only manufacturing industries and to measure labor with the wage bill and materials in nominal values to control for heterogeneity in labor and for differences in quality of intermediate inputs respectively, graphs shown in Figure C.6 below.

[58]Each dot corresponds to the average value between 1995 and 2007 of the intersectoral gains calculated using (C.5) for each country and the average GDP per capita in constant 2005 US dollars obtained from the World Bank. The results are very similar if median values are used. Two small countries with many zeros in sectoral data were dropped from the WIOD sample (Luxembourg and Malta). Likewise, Taiwan was dropped to make comparable WIOD and World Bank data.

Therefore, there is evidence that less developed economies tend to have greater inter-sectoral gains for removing distortions. This is consistent with the insights of multi-country studies as Tombe (2015) or Święcki (2017) which focus on inter-sectoral misallocation, that find larger intersectoral distortions in poor countries. Thus, omitting the inter-sectoral component of the total gains from removing distortions understates the common TFP gaps attributed to firm-level misallocation.

Figure C.6 – Inter-sectoral gains and GDP per capita: Alternative specifications



Note: Averages 1994-2007. Data source: WIOD (Timmer et al., 2015), World Bank Development Indicators