Intra- and inter-industry misallocation and comparative advantage*

José Pulido†

February 2021
[Latest version here]

Abstract

Micro-level resource misallocation, both within and across industries, can affect the relative unit costs of production across sectors, distorting comparative advantage. After presenting evidence on how changes in factor misallocation of a particular country (Colombia) relate to the dynamics of its revealed comparative advantage, I use a mis-allocation model with international trade to evaluate how its specialization patterns would change if resources were used efficiently. The new specialization would allow Colombia to raise its ratio of exports to manufacturing GDP by 18 pp. This industrial composition effect is absent in the workhorse models of misallocation under closed economies.

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

*I am especially grateful for the advice and guidance of my supervisors at the University of British Columbia (UBC), Matilde Bombardini and Tomasz Święcki. I also thank Vanessa Alviarez, Amartya Lahiri, Yaniv Yedid-Levi, Jesse Perla, Victoria Hnatkovska, Giovanni Gallipoli and Nouri Najjar from UBC for their helpful comments and collaboration, as well as Rodrigo Adão, Lorenzo Caliendo, Andrés Rodríguez-Claire, Svetlana Demidova, Marcela Eslava, Chang Tai Hsieh and Loukas Karabarbounis for their feedback. I also thank seminar participants at the 2017 CEA, Central Bank of Colombia, BIS-CCA, Sauder-VSE trade group, Universidad del Rosario and Universidad EAFIT for their comments and suggestions.

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

1 Introduction

What are the implications of micro-level resource 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 total factor productivity (TFP). Under standard assumptions on the demand and production structure, and regardless of the underlying cause of the inefficient use of resources – regulations, financial constraints, information asymmetries, crony capitalism, and so on – 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 (HK), 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 aggregate economic performance has focused on closed economies.[2] If the extent of factor misallocation varies not only across countries but also across industries, in open economies it could also shape the relative unit costs of production across sectors, distorting the "natural" comparative advantage of a country.[3] For example, consider the broad range of industrial policies to promote strategic industries that several East Asian countries introduced during the post-war period. 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, and more.[4] Thus, the likely improvement in the

[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 welfare effects of trade liberalization in economies with factor misallocation; these references are commented 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 (2019).

export capability of targeted sectors due to reductions in the average factor costs compared to neglected industries, could have been countered by decreases in their sectoral TFP, due to their larger extent of intra-industry resource misallocation. A relevant question then is 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 study explores how resource 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 two questions. First, does resource misallocation explain observed industries' export capabilities once I 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 considering all general equilibrium effects?

To verify the role of factor misallocation as a determinant of comparative advantage, I first present empirical evidence on how standard metrics of resource misallocation of a given country (Colombia), both within and across industries, are related to measures of the country's sectoral export capability, once its "natural" determinants of comparative advantage are controlled for.[5] As a metric of sectoral 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 measures relative to the United States (US) on indicators of both intra- and inter-industry misallocation, exploiting their variation over time in a lapse of 20 years. The specification controls for the "natural" sources of Colombian comparative advantage, i.e. its total factor endowments interacted with factor intensities and its efficient sectoral productivities; which capture Heckscher-Ohlin and Ricardian forces respectively. All variables are expressed relative to the sectoral producer price index in the US, the natural indicator of the total opportunity cost of production in a sector (French, 2017).

The results suggest that the metrics of resource misallocation have a quantitative relevance in shaping Colombian RCA, with similar magnitudes to the ones observed for the "natural" determinants. These results are robust to accounting for measurement error, one of

[5]Colombian manufacturing data, considered one of the richest in the world (De Loecker and Goldberg, 2014), have been extensively used in the empirical literature. One of its main features is the availability of product-level prices for firms' inputs and outputs in several years. Those prices have allowed researchers to disentangle the measures of physical productivity used in the misallocation literature into efficiency and demand components. However, since in the empirical application of my model I take a stance on the distribution of firm's physical productivities, those prices are not directly used.

the main concerns in the misallocation literature (White, Reiter and Petrin, 2018; Rotemberg and White, 2019; Bils, Klenow and Ruane, 2020; Gollin and Udry, 2021). Particularly, I use Bils, Klenow and Ruane's (2020) method, which considers additive measurement error in both revenue and inputs. This method utilizes 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 (revenue) products. Hence, panel data can be used to back out the "true" dispersion in the marginal returns of factors, by estimating how such elasticity changes for plants with different levels of average products.

Next, I examine the different channels through which resource misallocation can shape relative industries' unit costs and hence RCA. This exploration considers several general equilibrium 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 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 HK, the gains in sectoral efficiency from removing intra-industry misallocation do not generate reallocation of factors across sectors under the standard two-tier [upper-level Cobb Douglas (CD) and lower-level constant elasticity of substitution (CES)] demand system.[6] 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 sector-specific productivity shocks in standard Ricardian models. 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), Adamopoulos et al. (2017) or Yang (2021), 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

[6]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 relative factor 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 for details).

incumbent firms: a withdrawal of low-efficiency firms that survived because of factor misallocation and the addition of potential high-efficiency firms that were not able to operate under allocative inefficiency. In monopolistically competitive industries, this recomposition of firms affects the scale of the sectors and the average productivities through firms' selection effects, impacting also 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 adjustments on relative factor prices, which produces further adjustments on average productivities under self-selecting heterogeneous firms.

To consider all these general equilibrium channels, I explore the consequences of factor misallocation in an open economy using a tractable multi-country, multi-factor, and multi-sector model of international trade à la Melitz (2003). In the model, the allocation of factors across heterogeneous firms is inefficient; and because either the industry as whole faces different returns of the factors or the degree of intra-industry misallocation varies across sectors, inter-industry inefficiencies are also generated. I employ wedge analysis to characterize the observed dispersion in the marginal returns of the factors at the firm-level, 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.[7] Under this approach, each firm is represented by a draw of "true" efficiency and a 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.

To illustrate the consequences of factor misallocation on the RCA of a country and its bilateral exports, I consider again the Colombian case and investigate the effect of removing its observed resource misallocation, after factoring out measurement error. 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

[7] 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*b*), Tombe (2015) and Yang (2021) among others.

is not demanding in terms of data requirements. I perform the exercise using a world composed of 47 countries and an aggregate rest of the world (RoW), three production factors, and 25 tradable sectors.

The results of the counterfactual exercises suggest that in Colombia, resource misallocation plays a major role in shaping its 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 gross domestic product (GDP) rises by 18 p.p. and welfare, measured by real expenditure, grows 75%.[8] 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, and 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 to only to factor misallocation, particularly computer, electronic and optical products, transportation equipment, petroleum, and machinery and equipment. These four industries shrink and practically disappear, indicating a non-interior solution in the counterfactual equilibrium.[9]

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 costs: 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 RoW are offset in large part by the rise in the relative factor prices, and the residual effect does not vary much across industries.

Related literature

This study 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*b*), Caliendo, Parro and Tsyvinski (2017), Costa-Scottini (2018), Bai, Jin and Lu (2019), Berthou et al. (2020) and Chung (2020). My approach is different from the one used in most of these

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

[9]The feasibility of non-interior solutions in multi-sector Pareto-Melitz type of models is established by Kucheryavyy, Lyn and Rodríguez-Clare (2020). Under a simplified setup to the one used in this paper, it is guaranteed that the general equilibrium is unique, but not necessarily an interior solution.

papers and is aligned to Caliendo, Parro and Tsyvinski (2017): 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 studies with the closest models to the one used here are Ho (2012), Costa-Scottini (2018), Bai, Jin and Lu (2019), Berthou et al. (2020) and Chung (2020), which use different variations of open-economy models with firm-level factor distortions and selection effects of heterogeneous firms. My multi-country, multi-sector, and multi-factor model shares some features with those studies, but it differs in several aspects.[10] My empirical implementation is also different, since I obtain counterfactuals without relying on the combination of estimating and calibrating large sets of structural parameters. Instead, I use "exact hat algebra" that is not demanding in terms of data requirements. Further, unlike Tombe (2015) and Święcki (2017*b*), who use a type of Eaton and Kortum's (2002) model to study welfare and the gains from trade under the presence of sectoral distortions, and thus only inter-industry misallocation, my model is able to generate ex-post misallocation across industries as a result of differences in the first and second moments of the underlying distributions of factor distortions across sectors. This 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 study is also related to the trade literature focusing 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

[10] 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 productivity cut-off. Thus, the selection effects of distortions do not produce rank-reversals, which are necessary to obtain the large TFP gaps attributed to factor misallocation (Hopenhayn, 2014*a*).

(2016), and French (2017). I use the same approach to obtain RCA measures, which is the main metric of interest in my counterfactual exercises.

The rest of this paper is organized 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 verify whether different measures of Colombian factor misallocation are related to the RCA metrics. 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 explain how this measure is linked to the relative producer price index. Second, I decompose the price index in terms of the "natural" sources of comparative advantage and metrics of factor misallocation. Next, I propose a strategy to evaluate the relation between the metrics of factor misallocation and 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 wherein 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 factors that represent: i) the capabilities of exporter i as a supplier to all destinations; ii) the demand for foreign goods of importer j; iii) the 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 to reduce aggregation bias.[11] With cross-sectional data, the standard procedure involves taking logs and estimating

[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).

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), the 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), the 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 exporter-industry fixed effect characterizes the relative country's productive potential in an industry and, given the structure of the gravity equation, 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 RCA, an approach that has increasingly gained relevance in the trade literature (Costinot, Donaldson and Komunjer, 2012; Hanson, Lind and Muendler, 2015; 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 the RCA measures of the 25 manufacturing industries listed in Table A.1 of Appendix A.1 for Colombia in 1995. To compute these measures, I use bilateral trade flows among 47 countries plus an RoW aggregate, plus an estimation of imports from home for each country-sector. The set of countries is listed in Table A.2 of Appendix A.1, where more details about data sources and procedures are given. Similar to Hanson, Lind and Muendler (2015), I use the mean of all countries and industries as a reference country and industry; hence, the RCA can be interpreted as a measure of Colombian industries' capabilities relative to a "typical" country and "typical" sector.[12] The logarithmic transformation in equation (1) poses two well-known econometric issues for an estimation by ordinary least squares (OLS). First, zeros in bilateral exports are not likely to be random in the data, and since

[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:

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

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) and Poisson pseudo-maximum-likelihood (PPML, proposed by 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). Notably, 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 the 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 technology requirements. In the Krugman (1980) model, it depends only on the relative number of varieties produced, quantifying the effect of differences in industries' scales on the unit costs. In the Pareto version of the Melitz (2003) model, the ratio is analogous to that in Krugman (1980), adjusted by the lower bound of Pareto's productivity distribution, implying that the support of the firms' physical productivities also plays a role. Multi-factor models with heterogeneous firms, as in Bernard, Redding and Schott (2007) or in this study, combine all mentioned sources of 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 considering endogenous entry and

[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)).

selection of firms, features that will provide a rich theoretical grounding to the RCA measure. However, at this point, I can use the insights from the standard misallocation framework –HK–, to decompose the producer price index in its different determinants and empirically test whether the components due to resource misallocation are related to the metrics of RCA, once I control for the remaining sources of export capability.

2.2 Decomposing the price index under factor misallocation

To evaluate the implications of firm-level factor misallocation, the HK framework 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 (TFPQ) 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, 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 intra-industry factor misallocation and leads to a loss in sectoral TFP.

Now consider the link between the RCA measure obtained in the last section and the factor misallocation measures. To do so, assume firms producing a variety m in a manufacturing industry s of country i use a CD production technology with L homogenous factors z_l, TFPQ a_m and factor intensities α_{ls}. I omit industry and country subscripts for firm-specific variables and denote sectoral aggregates with capital letters. Denote as σ the constant elasticity of substitution of the demand side that produces a constant mark-up $\frac{1}{\rho}$. 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, that is $(A^e_{is})^{\sigma-1} = \sum_m a_m^{\sigma-1}$, 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 $\frac{1}{\rho}\prod_l^L w_l^{\alpha_{ls}}$.

Denote sectoral revenue as R_{is}. Since the sectoral price index can be expressed as the ratio of the sectoral revenue productivity, $TFPR_{is}$, and the industry TFP A_{is}, it can, in turn,

[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.

be decomposed in terms of "natural" sources of comparative advantage and metrics of factor misallocation as:

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

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 $\frac{w_l}{\rho}$; and AEM_{is} corresponds to the ratio of the sectoral TFP to the efficient one, $AEM_{is} \equiv A_{is}/A_{is}^{e}$. Those two ratios quantify the extent of resource misallocation. In the first case, the sectoral wedge $\left(1+\bar{\theta}_{ils}\right)$ characterizes how much factor l is misallocated in the whole sector related to other sectors, and thus, the first term on the right-hand side of (2) is a factor-intensity weighted measure of total inter-industry resource misallocation. This sectoral wedge can also be computed as the harmonic weighted average (HWA) of analogue wedges at the firm level, with weights given by firms' shares in sectoral revenue. In the second case, AEM_{is} characterizes the amount of intra-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 intra-industry variance of firms' TFPR.[17]

Therefore, the decomposition in equation (2) reveals the four theoretical determinants of the RCA measure under resource misallocation: i) the efficient TFP A_{is}^{e}, 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 can be recovered in equilibrium as the interaction between factor endowments and intensities;[18] iii) the geometric average of inter-industry wedges $\prod_{l}^{L}(1+\bar{\theta}_{ils})^{\alpha_{ls}}$, which is a measure of inter-industry misallocation; and iv) the measure of intra-industry misallocation AEM_{is}. Since the first component is related to technical efficiency and the second component is related to factor abundance, they represent the "Ricardian" and "Heckscher-Ohlin" sources of comparative advantage, respectively. The third and fourth determinants summarize the effects of inter- and intra-industry resource misallocation, respectively. I use these four components (in logs) as explanatory variables in a regression of the RCA measure derived from the estimation of fixed effects, in order to test whether resource misallocation, once I control for "Ricardian" and "Heckscher-Ohlin" sources of export capability, is related to the metrics of RCA.

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

[18] 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 and β_s the sectoral expenditure (revenue) shares.

2.3 Relation between RCA and misallocation measures

Ideally, the suggested regression would require measures of the four variables for 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 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 to vary over time. That is, with data for the same set of countries and sectors as in Section 2.1, I run the regression:

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

for the period 1992-2012, 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 $dRCA_{ist}$, I take the US as the reference country i' (instead of global means), the first year in the panel (1992) as the reference year t', and the sector with the median number of zero bilateral flows in the data (footwear) as the reference industry s'.[19] In the second stage, I regress the estimates of $dRCA_{ist}$ for Colombian industries on the four theoretical determinants of comparative advantage (i.e. those in the right hand side of equation (2)), constructed using the Colombian micro-level data. Each variable is transformed to be expressed as a double difference, first with respect to the reference industry, and second with respect to the reference year, and then 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.[20]

For the implementation of this strategy there are two main concerns to be addressed. First, the introduction of the time dimension poses an additional challenge for the fixed effects es-

[19]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:

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

where i' =US, $t' = 1992$ and s' = Footwear. The results that follow below are not sensitive to the choice of t' or s'.

[20]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.

timators. Particularly, I 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 relatively small. Fernández-Val and Weidner (2016) prove that under exogenous regressors, in a Poisson model this bias is zero, which make PPML preferable to EK-Tobit as the estimating method in the first stage. Thus, the following results use only PPML to compute the exporter-industry-year fixed effects in the first stage.

Second, there is an important body of research raising concerns about the reliability of the HK metrics as true measures of resource misallocation. Particularly, interpreting all the dispersion in TFPR as misallocation might be problematic because differences in the average revenue products of factors can obey either to other causes that are not specified in the HK's representation of the economy (e.g., variable markups, heterogeneity in inputs, adjustments costs, differences in technologies, etc.) or to the presence of measurement errors in revenues or inputs. Regarding the first set of causes, some recent studies employ different extensions of the HK framework to quantify the individual contribution of other possible sources of dispersion in TFPR that should not be accounted as resource misallocation. The common finding is that, at least for developing countries, the individual contribution of those other sources is relatively small.[21] Instead, measurement error in revenues or inputs, attributable for instance to the difficulties to measure capital or inventories or to the (lack-of) cleaning/imputation procedures of statistical offices, has been regarded as a possible major thread to the quality of standard metrics of resource misallocation (White, Reiter and Petrin, 2018; Rotemberg and White, 2019; Gollin and Udry, 2021). Given the latter evidence, and for tractability, I abstract from the possible sources of TFPR dispersion related to model misspecification and proceed to correct the Colombian misallocation metrics for the possible influence of measurement errors.[22]

For this purpose I use the method recently developed by Bils, Klenow and Ruane (2020)

[21]For example, David and Venkateswaran (2019) find that for China adjustment costs and uncertainty about firms' TFP, while are significant in explaining investment dynamics, contribute only about 1 and 10 percent respectively to the dispersion of the average revenue products of capital. Other firm-specific factors as heterogeneous markups or technologies also account for a modest proportion of the total dispersion: 4 and 23 percent respectively. Their results are similar when replicating their methodology for both Colombia and Mexico. For Colombia, Eslava and Haltiwanger (2020) show in a model that accounts for idiosyncratic demand shocks at the firm and firm-product levels and heterogeneity in factor prices, that under the assumption of imposing the right demand elasticities and returns to scale, the TFPR is still a good proxy for actual distortions, especially for those correlated with fundamentals. For developed countries, where factor misallocation should be less prominent, the contribution of misspecification to the observed dispersion in TFPR would be larger (see for the US David and Venkateswaran, 2019 and Haltiwanger, Kulick and Syverson, 2018).

[22]Moreover, it is worth to remind that my strategy exploits the variation in time of the misallocation measures across sectors. Thus, unless the causes of misspecification have heterogeneous impacts across sectors and over time, the results are robust to their omission.

which considers additive measurement error in both revenue and inputs. The method utilizes the fact that in the absence of additive measurement error, the elasticity of revenues with respect to inputs should not vary for plants with different levels of TFPR. Hence, by estimating in a panel how such elasticity changes for plants with different levels of TFPR, it is possible to quantify in which proportion the variances of both observed TFPR and the average returns of the factors are affected by measurement error. In Appendix A.2, I present the details on the methodology and results of its implementation for Colombia. I find that the contribution of measurement error to the variance of TFPR in Colombia is around 30% on average from 1992 to 2012, between to what the authors find for India (around 26% on average from 1985 to 2013) and for the US (around 63% on average for 1978-2013). Finally, as a way to externally validate the economic relevance of the obtained metrics, in Appendix A.3 I check for a single production factor whether its resulting measures of intra-industry misallocation are related to indicators that might suggest misallocation of the considered factor. Particularly, using an external dataset of the credit registry in Colombia, I show that the obtained intra-industry variances of the average revenue products of capital are related to the observed dispersion of the idiosyncratic cost of capital for firms within the same industry, measured by an appropriate weighted average of the interest rates payed for their loans.

The results of the two-stage strategy using PPML to obtain RCA measures in the first stage and the correction for measurement error in the metrics of intra-industry misallocation in the second stage are displayed in Table 1. Standardized coefficients of the regression in the second stage are presented to make comparable the magnitudes across determinants. Bootstrapped standard errors are shown in parentheses based on 1000 replications, to account for uncertainty in the estimation of both stages.[23] In the first column, I present the results for the four determinants that are shown in the right hand side of equation (2), so AEM_{is} is used as a measure of intra-industry misallocation, implying that larger values reflect less intra-industry misallocation. In the second column, as a robustness check, I use instead the within-industry variance of firms' TFPR as an alternative metric of intra-industry misallocation, so in this case larger values reflect more misallocation (the remaining determinants are equal to those in the first column).

Columns (1) and (2) show that both metrics of intra- and inter-industry misallocation are significantly correlated with the RCA measure and display the expected signs: more intra-industry misallocation reduces the RCA of the sector, and sectors with larger sectoral wedges $\prod_{l}^{L}\left(1+\bar{\theta}_{ils}\right)^{\alpha_{ls}}$ have lower RCA metrics. Moreover, the magnitudes of the standardized coef-

[23]Since the value of bilateral trade flows is obtained from the imports of reporting countries, clusters by importer-year are used for resampling in the first stage.

ficients suggest that both types of misallocation have an impact on shaping Colombian RCA that, although is smaller than the found for "Ricardian" comparative advantage, is not negligible compared to the magnitudes found for both "natural" determinants. Finally, as an additional robustness check, columns (3) and (4) replicate the exercise considering only the 20 Colombia's main trade partners to reduce the influence of zeros in the first stage. The results are quantitatively and qualitatively very similar. Therefore, the empirical evidence suggests that resource misallocation can play a role in shaping the schedule of comparative advantage in Colombia. The model in the next section offers a theoretical grounding for 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 misallocations 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 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 $G_{is}(a, \vec{\theta})$. 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_{ijs}^{x} 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} in the market of primary factor l (given by the l-th element of $\vec{\theta}_{im}$) such that the input price perceived by the firm is $(1+\theta_{ilm})\, w_{il}$. Define $f_{ijs} = f_{ijs}^{x}$ if $j \neq i$; $f_{ijs} = f_{iis}^{x} + f_{is}$ 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 $\Theta_{im} = \prod_{l}^{L}(1+\theta_{ilm})^{\alpha_{ls}}$ is a factor-intensity weighted geometric average of firm wedges and $\omega_{is} = \prod_{l}^{L}(w_{il}/\alpha_{ls})^{\alpha_{ls}}$ 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.

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):

$$q_{ijm} = p_{ijm}^{-\sigma}E_{js}P_{js}^{d\sigma-1}\ ;\ r_{ijm} = p_{ijm}^{1-\sigma}E_{js}P_{js}^{d\sigma-1};\ \breve{\pi}_{ijm} = \frac{1}{\sigma}r_{ijm} - \omega_{is}\Theta_{im}f_{ijs} \tag{5}$$

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: $c_{ijm} = \rho r_{ijm} + \omega_{is}\Theta_{im}f_{ijs}$.

Define z_{ijlm} as 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. Revenue productivity (TFPR) of selling variety m in destination j, denoted by ψ_{ijm}, is the ratio between revenue and the variable input used in such production: $\psi_{ijm} \equiv r_{ijm}/\left(z_{ijm} - f_{ijs}\right) = \Theta_{im}\omega_{is}/\rho$. Notice that although this destination-specific TFPR is not directly observable, 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 z_{ijlm}. 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 $(1+\theta_{ilm}) w_{il} z_{ijlm} = \alpha_{ls} c_{ijm}$, which derives into $\xi_{ijlm} = (1+\theta_{ilm}) w_{il} / \rho$. 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,[24] 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 $a^*_{ijs}(\Theta)$ for each destination j, such that $\breve{\pi}_{ijs}\left(a^*_{ijs}(\Theta), \Theta\right) = 0 \ \forall\, i, j, s$. For a given combination of factor wedges Θ of firms in country i industry s, i.e., a given value of TFPR, $a^*_{ijs}(\Theta)$ indicates the minimum TFPQ required to earn non-negative profits in destination j. Define a^*_{ijs} as the cutoff value of TFPQ for 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:

$$a^*_{ijs}(\Theta) = a^*_{ijs} \Theta^{\frac{1}{\rho}} \ with \ a^*_{ijs} \equiv a^*_{ijs}(1) = \frac{\tau_{ijs}}{\rho} \left(\frac{E_{js} P_{js}^{\sigma-1}}{\sigma f_{ijs}} \right)^{\frac{1}{1-\sigma}} \omega_{is}^{\frac{1}{\rho}} \ \forall\, i, j, s. \tag{6}$$

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

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

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.[25]

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 triggered by distortions. 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}.[26] 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, 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

[25]These selection channels are also present in the closed economy models of Bartelsman, Haltiwanger and Scarpetta (2013) and Yang (2021).

[26]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 $a^*_{ijs} > \tilde{a}^*_M$.

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:

$$\sum_{j}^{N}\sum_{m}^{M_{ijs}}\breve{\pi}_{ijm}=\omega_{is}f_{is}^{e}H_{is}\,\forall\, i,s \tag{7}$$

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:

$$d_{is}M_{ijs}=\left[1-G_{is}\left(a_{ijs}^{*}\left(\Theta\right),\Theta\right)\right]H_{is}\,\forall\, i,j,s \tag{8}$$

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

$$P_{ijs}^{1-\sigma}=\left(\frac{1}{\rho}\omega_{is}\tau_{ijs}\right)^{1-\sigma}\sum_{m}^{M_{ijs}}\left(\frac{a_{im}}{\Theta_{im}}\right)^{\sigma-1} \tag{9}$$

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 $P_{i}^{d}=\prod\limits_{s}^{S}\left(P_{is}^{d}/\beta_{s}\right)^{\beta_{s}}$ and satisfies $E_{is}=\beta_{s}E_{i}$, with $E_{i}=\sum\limits_{s}^{S}E_{is}$ total country-i expenditure.

Finally, consider the aggregate variables. Let $X_{ijs}=\sum\limits_{m}^{M_{ijs}}r_{ijm}$ 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 of all firms of country i industry s can be written as $C_{ijs}=\rho X_{ijs}+\mathfrak{F}_{ijs}$, where $\mathfrak{F}_{ijs}$ is the value of total expenditures in fixed costs, i.e.:

$$\mathfrak{F}_{ijs}=\sum_{m}^{M_{ijs}}\omega_{is}\Theta_{im}f_{ijs} \tag{10}$$

Similarly, denote by R_{is}, $\mathfrak{F}_{is}$, C_{is} the same aggregations but at the industry level, and R_{i} 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 $\left(1+\bar{\theta}_{ils}\right)=\left(\rho R_{is}+\mathfrak{F}_{is}\right)\alpha_{ls}/w_{il}Z_{ils}^{o}$ where Z_{ils}^{o} is the aggregate demand of factor l for "operational" uses in country i in sector s. 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 section 2.2. 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 $f^e_{is} H_{is}$, where a proportion $\alpha_{ls}\omega_{is}/w_{il}$ corresponds to primary factor l. Thus, the total allocation of 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 first allocate 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:

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

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 2. This table also offers the dimensionality of the problem.

3.2 Measure of RCA

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:

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

The trade share of country i in country j's expenditures on 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.[27] 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, that is, $\frac{X_{ijs}X_{i'js'}}{X_{ijs'}X_{i'js}}$. 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 (9) it is possible to disentangle these bilateral price 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.[28] Thus, A_{ijs} represents the overall efficiency of exporting firms to destination j and $\bar{\psi}_{is}$ depicts the average factor cost for 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 factor costs 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 determination of the number of exported varieties. Thus, the model is very rich in explaining the determinants of comparative advantage. It is able to combine the sources of relative export capability in Ricardian, Heckscher-Ohlin, and intra-industry trade models in an environment of micro-level resource misallocation, which, in turn, acts as a distortion to those sources of comparative advantage. In the next subsection, I perform numerical simulations to disentangle the effects of both intra- and inter-industry misallocations 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, and hence, the RCA measure. 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 $\vec{\theta}$.[29] Consider the following assumptions:

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

[27] As it was stated 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.

[28] That 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}.

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

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

First, for Assumption 1, the Pareto distribution is a 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[30] but it also makes the model analytically tractable, allowing me to derive a particular expression for the gravity equation. Second, although Assumption 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 a positive ex-post correlation between TFPQ and TFPR, such that the observed in the data. Furthermore, there is no restriction for the joint distribution of individual factor distortions G^{θ}_{is}; hence, covariances across factors wedges are completely allowed. I keep Assumptions 1 and 2 hereafter, unless otherwise indicated.

Under Assumptions 1 and 2, the model exhibits a noteworthy set of features and offers a great simplification, which is presented 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: $R_{is} = \frac{\kappa}{\rho}\Pi_{is} = \frac{\kappa}{\rho}\omega_{is} f^{e}_{is} H_{is}$. Thus, market clearing conditions can be re-stated as:

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

Note 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 square 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, I can state $v_{is} R_{is} = \omega_{is} Z_{is}$ and hence $H_{is} = \rho Z_{is} / \kappa f^{e}_{is} v_{is}$ 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 moments of the ex-ante joint distribution of distortions. Appendix B.2 shows that the following

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

relation holds:

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

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

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

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

where B_{ijs} and ρ_{is} are constants which do not vary when I remove misallocation. The first term of the right-hand side 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 into elements that capture the influence of each source of export capability distorted by resource misallocation: the average efficiency, returns of factors, and mass of exported varieties, which represents the effect of sector's scale. Moreover, note 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 presented in section 4.

3.3 Simulations

To illustrate the effects of both intra- and inter-industry misallocations 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, such that 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. Hence, in the allocative efficient scenario, it has a comparative advantage in sector 1.[31] 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}, respectively, for factor l, 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 feedback 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) I 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, thereby 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, 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 RCA 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, third and fourth graph, respectively), following equation (B.14) 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, thereby capturing the net effect of the considered allocative inefficiency. The fifth graph illustrates the

[31] 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.

implicit HWA of the prevalent distortions, following equation (20), to verify the degree of inter-industry misallocation. The 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,[32] and costly trade, represented by continuous lines. The values for the whole set of parameters used in each simulation are displayed in Table 3.

Introducing only intra-industry misallocation of any factor used in sector 1 reduces its comparative advantage. The effect increases as the variance of the (log) wedges becomes larger and, for the same value of the variance, if the misallocation impacts the factor used intensively by the industry. The total effect is also marginally larger under free trade for the range of variances considered in the graph. It is worth noting that under free trade, there is a threshold for larger variances wherein the system falls in a regime of complete specialization, thereby shutting down the production of sector 1. 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, similar to those 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, although not zero.[33]

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 misallo-

[32]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.

[33]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).

cation. 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 of 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 an 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 inputs at their new cost must exit. Here, there is no TFP loss due to intra-industry misallocation, because all firms in the industry face the same factor prices; hence, average TFP depends only on the physical productivities of the incumbents. Instead, there is selection of more productive firms, thereby raising the average TFP . 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 in 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, although small. Adding the negative effect on the extensive margin due to the exit of firms, which is not much 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 the 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. Reductions in industries' TFP are 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 exercises of removing both jointly 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, method to measure the dispersion in the MRP of the factors under overhead costs, and baseline results. Finally, I conduct some robustness checks and compare the baseline results with those obtained for the single-sector and the closed economies.

4.1 Counterfactual exercises

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 expenditure of country j in sector s. Denote the share of factor l allocated to sector s in country i as $\tilde{Z}_{ils}$. Thus, 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 well-known "exact hat algebra" methodology pioneered by Dekle, Eaton and Kortum (2008) to obtain counterfactual equilibria by expressing the model in relative changes. This approach 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- 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.

For any variable x in the initial equilibrium, x' denotes 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 misallocations (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} and Γ_{is} for Colombia to derive the corresponding proportional changes.

Once $\hat{R}_{is}$ and $\hat{w}_{il}$ are obtained, the computation of relative changes in aggregate expenditure and trade shares, $\hat{E}_i$ and $\hat{\pi}_{ijs}$ respectively, is straightforward. 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:

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

Note 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 2.3. Unlike that section, here I focus only on a particular year (1995). See Appendix A.1 for a detailed description of the data used. 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 the US cost shares reflect actual differences in technology across sectors instead of inter-industry misallocation. 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.[34] 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 the countries different from RoW, trade deficits as a proportion of gross output remain constant in the counterfactual equilibria; the trade deficit of RoW adjusts to ensure global trade balance.

The counterfactual exercises involve misallocation removal of: i) both types; ii) only intra-industry; and iii) only inter-industry for the primary production factors: capital, skilled, and unskilled labor.[35] 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. Under log-normality, it is possible to relate the ex-post HWA wedges $\left(1+\bar{\theta}_{ils}\right)$ to the vector of location parameters and to the variance-covariance matrix of the ex-ante joint distribution of the distortions V_{is} (see equation (B.15) in Appendix B.4, a generalization of equation (B.4) for the multi-factor case with arbitrary covariances). Therefore, I only need measures of the HWA of wedges, which are observed from sectoral data using (17), and estimates of V_{is}, to obtain the latent location parameters and, consequently, both v_{ils} and Γ_{is}. For the estimates of V_{is} I use the observed variance-covariance matrix of the average revenue products of factors, corrected for measurement error as explained in section 2.3. Notice that overhead factors are analogous to an unobservable additive term in measured inputs, and thus the correction for additive measurement error also deals with the possible biases due to their omission (Bartelsman, Haltiwanger and Scarpetta, 2013). Table 4 displays for each industry the employed values for the HWA wedges and the corresponding corrected variances and covariances of factors' average revenue products, along with factor intensities.

Finally, the model is constituted by $N \times (S+L) = 1344$ equations. The multiplicity of non-linearity in the model implies that the common optimization routines find multiple local solutions. To obtain the global solution, I employ both an algorithm to choose a set of ideal

[34]These values are averages of those used by Melitz and Redding (2015) ($\kappa = 4.25$ and $\sigma = 4$) and those estimated by Eaton, Kortum and Kramarz (2011) ($\kappa = 4.87$ and $\sigma = 2.98$).

[35]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.

initial conditions and a state-of-the-art solver for large-scale non-linear systems. Appendix A.4 presents the details on 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-industry, 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 counterfactuals in the vertical one. Panels A and B show the case of removing both types of misallocation. In Panel A, the markers' size represents the observed 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 seems surprisingly large at first glance, these results are in line with the findings in much of the literature that assesses the gains of similar reforms.[36] Table 5 displays a decomposition of the aggregate results. The boost in exports is due to the increase in the dispersion of the 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 confirms 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,[37] which explains

[36] 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 (2021) 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).

[37] The feasibility of non-interior solutions in multi-sector Pareto-Melitz type of models is recently evaluated by Kucheryavyy, Lyn and Rodríguez-Clare (2020). It is possible to show that in a simplified setting than the presented here (a single-factor model without resource misallocation), with the same elasticities of substitution for domestic and foreign varieties and fixed costs of exporting paid in terms of factors of the destination country,

in part the longer left tail in the counterfactual world distribution.[38] The larger dispersion on the frictionless comparative advantage leads to higher degrees of industrial specialization in the frictionless equilibrium, which is evident by comparing the export shares from panel A with 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, accounts for 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 intra-industry dispersion on the factors returns. 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. Note that in both cases, the dispersion of comparative advantage is lower than in Panels A and B, but larger than the original one. Table 5 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 of 13 p.p. in the exports to GDP ratio and a 56% rise in welfare, whereas removing only inter-industry misallocation causes smaller increases (7 pp. and 8% in each variable, respectively).

The direction and magnitude of 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 intra-industry misallocation effect depends on the interaction between the factor intensities and relative variances of distortions, whereas the impact of inter-industry misallocation depends on whether the HWA wedges are lesser 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 wherein firms' TFPR are 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 while others lose export capability when inter-industry misallocation is removed.

the equilibrium is unique, but not necessarily an interior solution.

[38] 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.

A further exploration of the latter results sheds light on the direction and extent of the general equilibrium effects that are present in the model. Similar to section 3.3, I use the decomposition (B.14) in Appendix B.3 to disentangle the effect of each type of misallocation on the three sources of export capability in the model: average TFP, cost of inputs, and the number of varieties produced in each sector. The left panel of Figure 8 (Panel A) displays the effect of removing both misallocations (top graph), only intra- (middle graph) and only inter-industry types (bottom graph), on each sector's RCA. The numbers displayed correspond to the log-differences between the counterfactual values and initial values of the RCA measures, and the lengths of the bars represent the strength of each element in the decomposition, implying that they add up exactly to the number shown.

The most striking result is that the changes in the RCA measure are mainly due to the adjustments in the extensive margin, that is, the number of firms in each sector. This is particularly evident when removing only intra-industry misallocation, where the contribution of each component of the intensive margin (average TFP and cost of inputs) is almost null. To understand this result better, Panel B of Figure 8 shows the same decomposition when the changes in the three sources of export capability are not compared across industries, but 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 as the "absolute advantage" index.[39] In the case of removing 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 factor prices, on average by 0.74 log points (a rise in relative factor prices is shown as a negative contribution). Thus, the net effect of the intensive margin is on average small, but positive. However, since it has a very low dispersion across sectors, it translates to an almost null impact on the RCA measure, contrary to what happens with the impact on the number of varieties.

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

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

where AA denotes the "absolute advantage" index.

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

Importantly, changes in κ or in σ do not alter the ranking of RCA in the counterfactual equilibria and have a small effect, if any, on its dispersion. Figure 9 displays in 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 5). However, for a given distribution of distortions, the extent of factor reallocations across sectors is increasing in κ and decreasing in σ. This is because in each industry, a fraction $\frac{\rho}{\kappa}$ of the sectoral factor demand –the fraction that is allocated to entry– is not affected by firm-level misallocation. Consequently, Table 5 shows that the rise in total exports and export-GDP ratio 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 types of misallocation –together and separately– on welfare gains (Panel A) and exports growth (Panel B). The values of lines in the extreme right –removing 100% misallocation– coincide with the numbers in Table 5. Even the smallest reform that reduces the extent of both types of misallocation by 10%, has a sizable impact on both welfare and exports (6.7% and 11%, respectively).[40] Moreover, 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 re-compute the corresponding US cost

[40] 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.

shares and intra-industry variances of firm's wedges, values displayed in the last row of Table 4. The increase in welfare is similar to the baseline case (70%), but the increase in nominal exports is only 43%, leading to a decrease of 5 p.p. in the export-GDP ratio (see the last row in Table 5).

Closed vs. open economy

Since revenue shares in the closed economy are constant and equal to the expenditure shares in the demand system, there is no change in the industrial composition under the CD demand. However, it is possible to quantify the cost of the same measures of misallocation in terms of welfare. For this, note that in the closed economy, I have $\pi_{iis} = \hat{\pi}_{iis} = 1$ and $\hat{R}_{is} = \hat{E}_{is} = \hat{E}_i$, so I 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 5 shows the increase in welfare in the case wherein 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 welfare gains due to removal of 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.[41]

[41] For the inter-industry case, the results are in line with Święcki (2017*b*), 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 study offers a framework for a country to compute 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 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 causes an increase in the ratio of exports to manufacturing GDP by 18 p.p. and in welfare of 75% in the case of an extreme reform wherein 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 an omitted mechanism in the workhorse models of firm-level resource misallocation in closed economies.

The impact of removing resource misallocation on comparative advantage depends, importantly, on the adjustments in the extensive margin, mainly in the case of removing allocative inefficiencies within industries. In this case, the increases in comparative advantage are larger for those sectors in which the returns of the factors used intensively are relatively more dispersed. The gains in terms of relative unit costs are mainly the result of an increase in the relative number of varieties produced, because at the intensive margin, the increases in sectoral TFP are, in an important proportion, countered by the responses in relative factor prices, and there is not enough variation in the residual effect across industries.

These results suggest that the design of mechanisms that smoothens 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 factor 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." National Bureau of Economic Research Working Paper 23039.

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.

Bai, Yan, Keyu Jin, and Dan Lu. 2019. "Misallocation under trade liberalization." National Bureau of Economic Research Working Paper 26188.

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.

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, John Jong-Hyun Chung, Kalina Manova, and Charlotte Sandoz-Dit-Bragard. 2020. "Trade, productivity and (mis)allocation." Banque de France Working Paper 749.

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2020. "Misallocation or mismeasurement?" National Bureau of Economic Research Working Paper 26711.

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, 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. 2020. "Firm heterogeneity, misallocation, and trade." Auburn Economics Working Paper Series 2020-01.

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, 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. 2019. “The sources of capital misallocation.” *American Economic Review*, 109(7): 2531–2567.

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.

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 C Haltiwanger. 2020. “The size and life-cycle growth of plants: The role of productivity, demand and wedges.” National Bureau of Economic Research Working Paper 27184.

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 nonlinear 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. 2021. "Heterogeneity, measurement error and misallocation: Evidence from African agriculture." *Journal of Political Economy*, 129(1): 1–79.

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.

Haltiwanger, John, Robert Kulick, and Chad Syverson. 2018. "Misallocation measures: The distortion that ate the residual." National Bureau of Economic Research Working Paper 24199.

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. 2020. "Grounded by gravity: A well-behaved trade model with external economies." Unpublished.

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. 2019. "Manufacturing revolutions: Industry policy and industrialization in South Korea." 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.

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.

Rotemberg, Martin, and T. Kirk White. 2019. “Measuring cross-country differences in misallocation.” Unpublished.

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*b*. “Intersectoral distortions and the welfare gains from trade.” *Journal of International Economics*, 104: 138 – 156.

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

White, T. Kirk, Jerome P. Reiter, and Amil Petrin. 2018. “Imputation in U.S. manufacturing data and its implications for productivity dispersion.” *The Review of Economics and Statistics*, 100(3): 502–509.

Yang, Mu-Jeung. 2021. “Micro-level misallocation and selection.” *American Economic Journal: Macroeconomics*, Forthcoming.

Tables

Table 1 – 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	0.442***		0.447***	0.339***
(AEM_{is})	(0.154)		(0.153)	(0.084)
Intra-ind. variance of TFPR		-0.103*		-0.104*
($\sigma^2_{TFPR_{is}}$)		(0.062)		(0.062)
Inter-industry wedges	-0.522***	-0.268**	-0.542***	-0.285**
($\prod_l^L (1+\bar{\theta}_{ils})^{\alpha_{ls}}$)	(0.170)	(0.124)	(0.165)	(0.113)
Efficient TFP	0.786***	0.588***	0.812***	0.611***
(A^e_{is})	(0.189)	(0.139)	(0.186)	(0.133)
Factor prices	-0.428***	-0.169**	-0.450***	-0.187**
($\prod_l^L w_{il}^{\alpha_{ls}}$)	(0.093)	(0.071)	(0.095)	(0.076)
Observations first stage	1158912	1158912	201600	201600
Observations second stage	500	500	500	500

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-year 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. Columns (1) and (2) show the results for the baseline set of countries (see Table A.2), and (3) and (4) for Colombia's 20 main trade partners. Standardized coefficients and bootstrapped standard errors based on 1000 replications.

Table 2 – 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	(6)	$N \times N \times S$
Aggregate stability	(8)	$N \times N \times S$
Free entry	(7)	$N \times S$
Industry price	(9)	$N \times S$
Industry demand	$Q^d_{is} = (\sum_k^N \sum_m^{M_{kis}} q^{\rho}_{kim})^{\frac{1}{\rho}}$	$N \times S$
Trade balance	(14)	N
Aggregate price	$P^d_i = \prod_s^S (\frac{P^d_{is}}{\beta_s})^{\beta_s}$	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 3 – 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 4 – 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)				Intra-industry variances of log-wedges*			Intra-industry covariances of log-wedges*		
		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}
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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

Notes: *Values correspond to the variances of the observed average revenue products of factors, after removing outliers, trimming 1% tails and implementing the method of Bils, Klenow and Ruane (2020) to account for additive measurement error in both revenues and inputs.

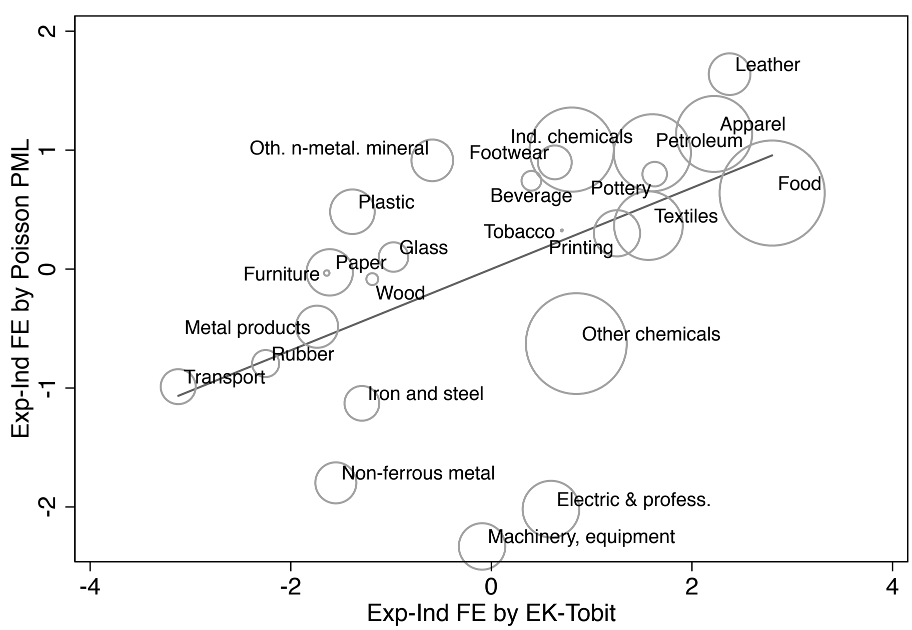
Table 5 – Counterfactuals

	Change in each variable after removing factor misallocation in Colombia						
Variable	Revenue	Value added	Exports	Exports /GDP*	RCA s.d.*	Welfare	Welfare - autarky
Counterfactual	$\hat{R}_{Col}$	$\hat{GDP}_{Col}$	$\hat{X}_{Col}$	$\Delta(\frac{X}{GDP})_{Col}$	$\Delta\sigma_{RCA_{Col}}$	$\frac{\hat{E}_{Col}}{\hat{P}_{Col}}$	$\left[\frac{\hat{E}_{Col}}{\hat{P}_{Col}}\right]^{closed}$
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

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

Figures

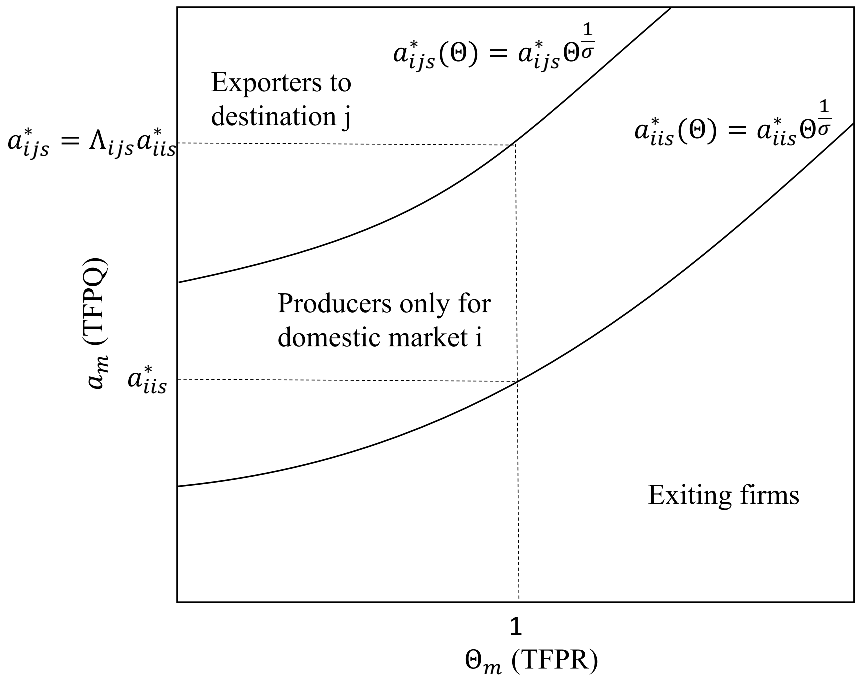
Figure 1 – Revealed comparative advantage (RCA) measures for Colombia



Notes: Figure shows the measures of RCA for Colombian industries in 1995 using PPML and EK-Tobit as estimation methods of equation (1). Markers' sizes represent export shares, and the line the best linear fitting.

Figure 2 – Cutoff functions and selection effects of distortions

Panel A: Cutoff functions for country *i* sector *s**



*For the domestic market and the destination j with lowest Λ_{ijs}

Panel B: Selection effects of distortions

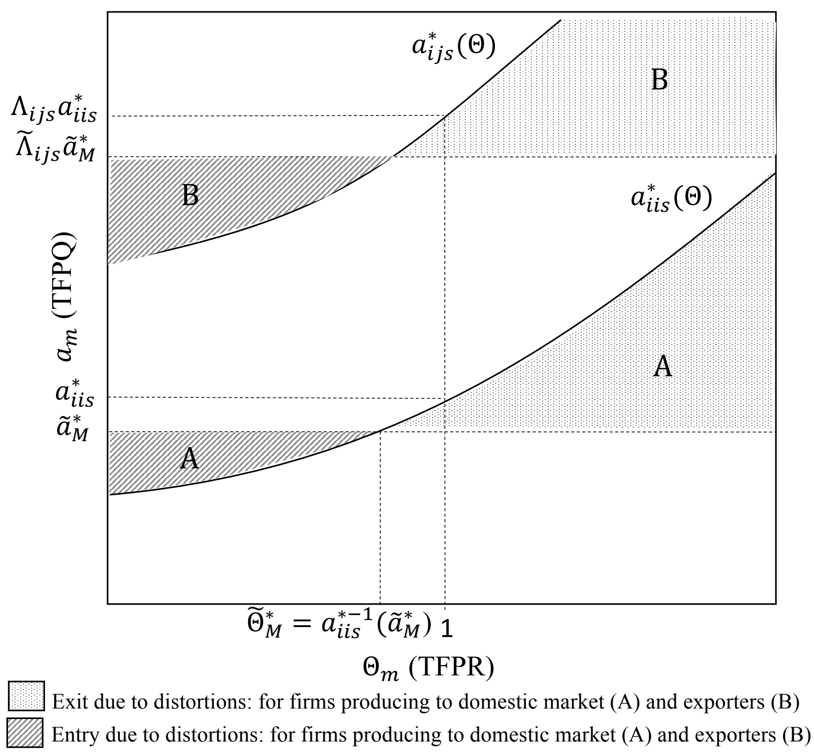
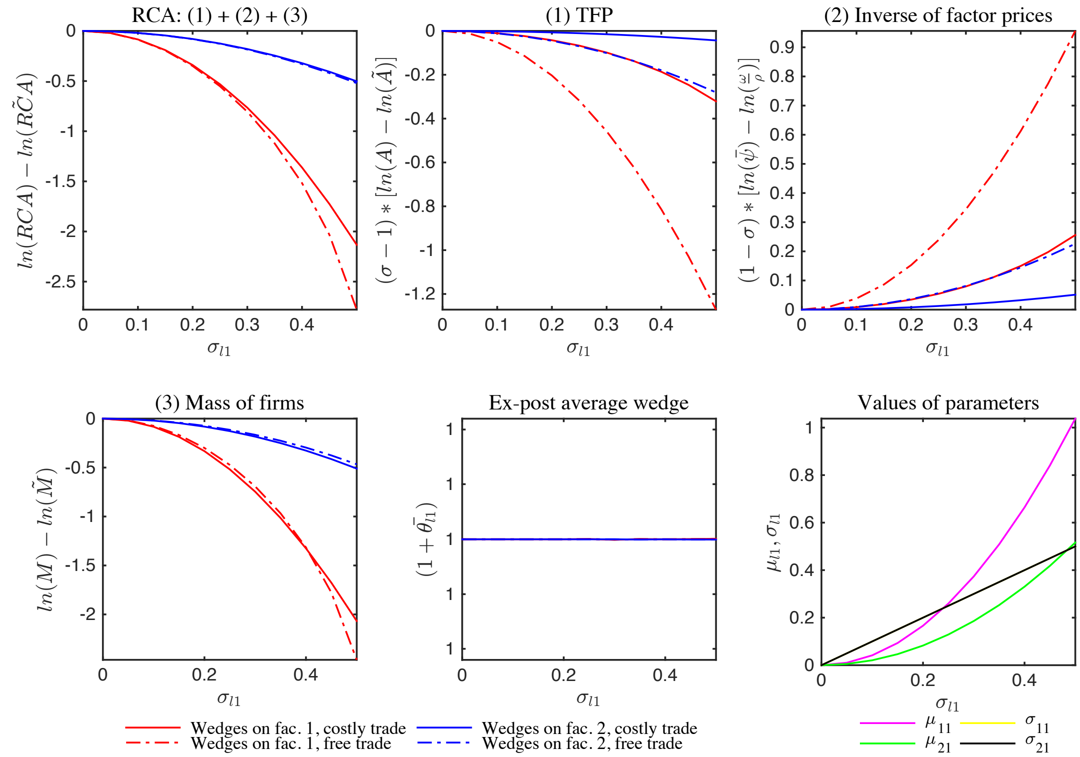
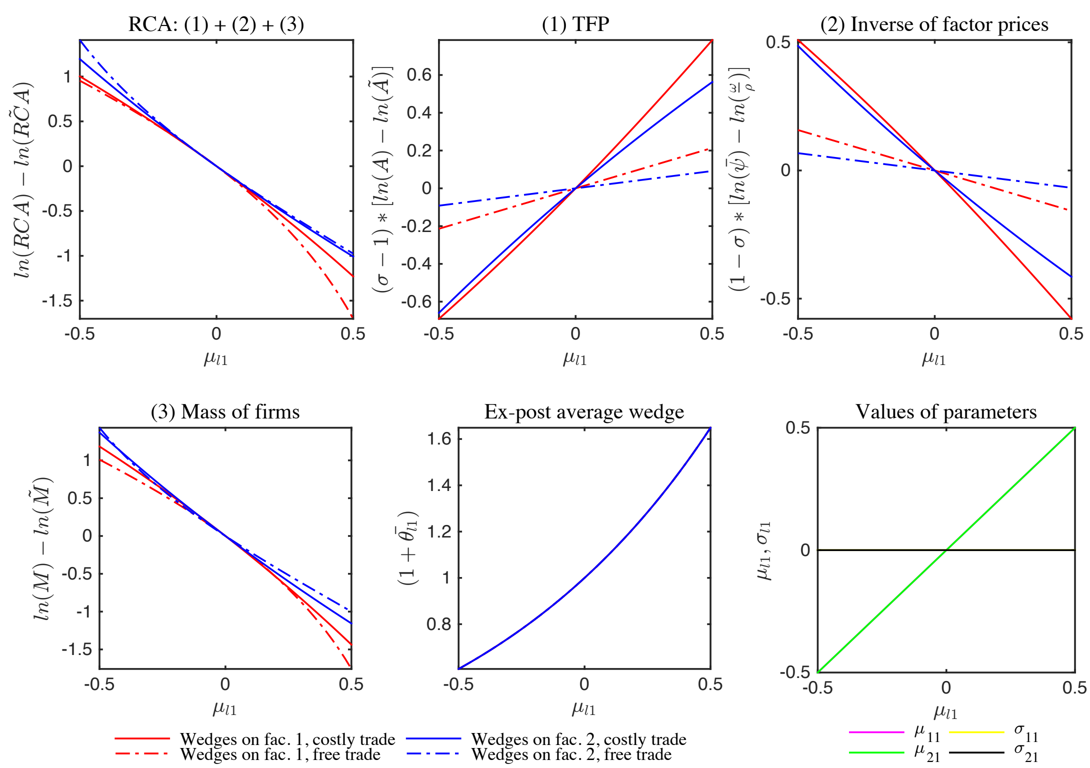


Figure 3 – Effects of intra-industry factor misallocation on RCA and its determinants



Notes: Figure is based on a model simulation in which intra-industry misallocation varies according to the depicted parameters' values. Graphs show the effects on industry's RCA and its determinants under free and costly trade. See section 3.3 for details on the specification.

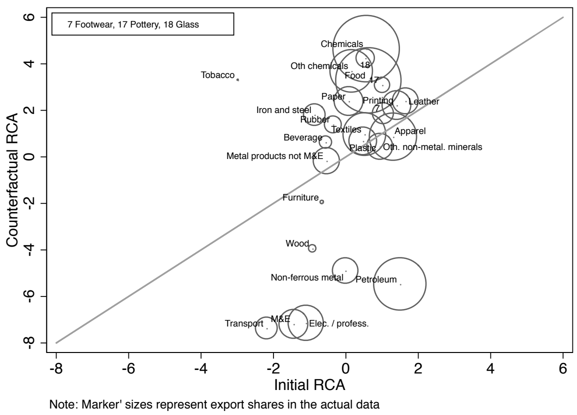
Figure 4 – Effects of inter-industry factor misallocation on RCA and its determinants



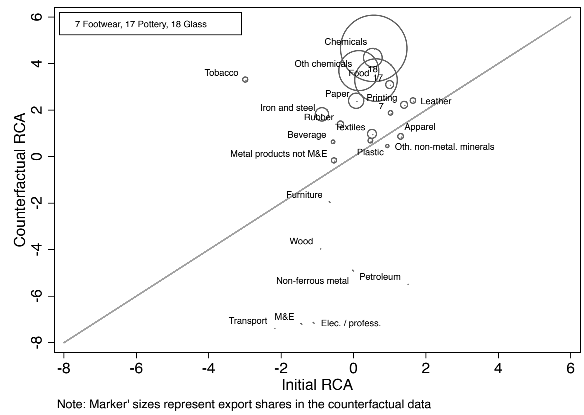
Notes: Figure is based on a model simulation in which inter-industry misallocation varies according to the depicted parameters' values. Graphs show the effects on industry's RCA and its determinants under free and costly trade. See section 3.3 for details on the specification.

Figure 5 – Allocative efficient RCA and observed RCA for Colombia

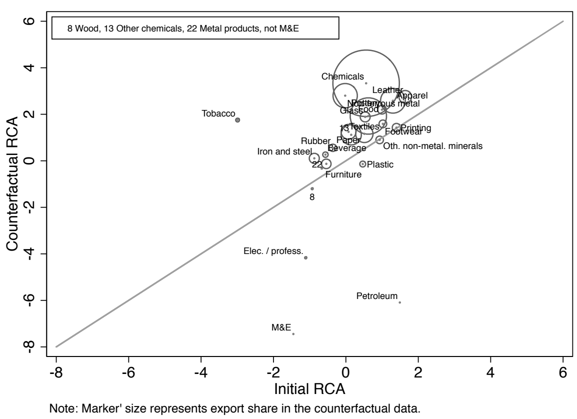
Panel A: Intra- and inter-industry allocative efficient RCA and observed RCA (observed export shares)



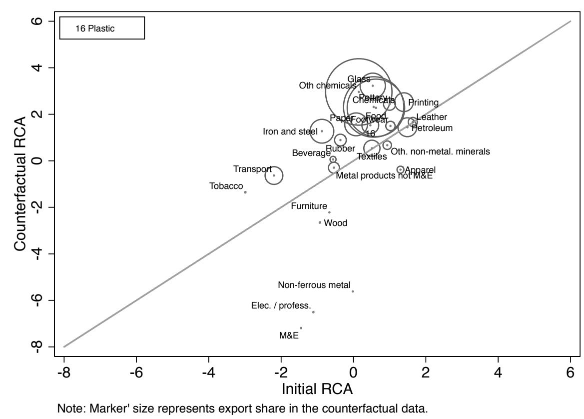
Panel B: Intra- and inter-industry allocative efficient RCA and observed RCA (counterfactual export shares)



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)

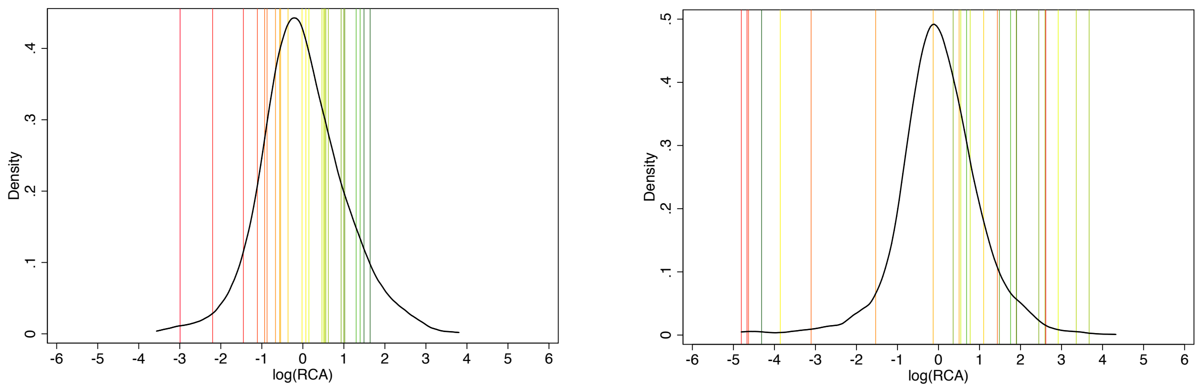


Notes: Each panel compares the RCA measures in the corresponding counterfactuals to the observed RCA measures. Markers' sizes represent the indicated export shares.

Figure 6 – Colombian industries in the world distribution of RCA

Panel A: Distribution under observed data

Panel B: Distribution under Colombia's efficient allocation

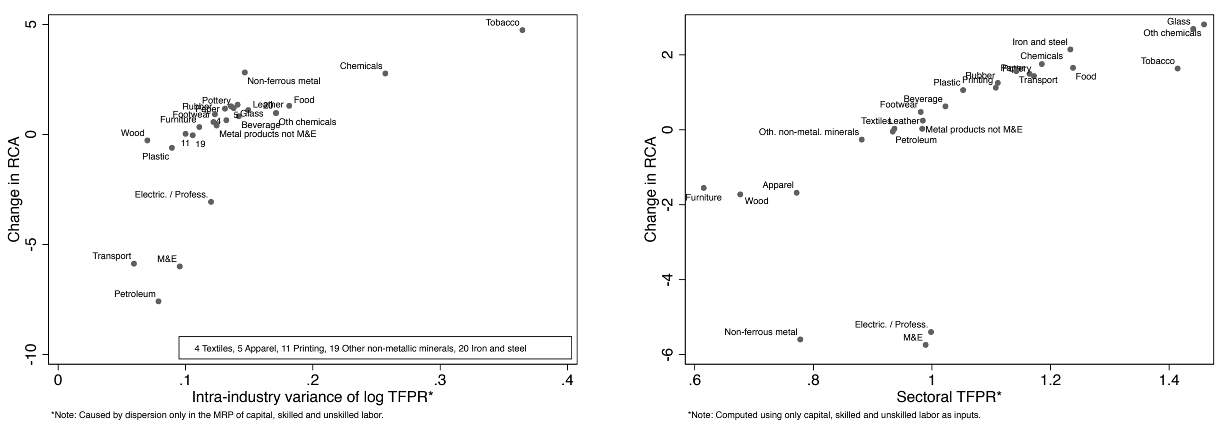


Notes: Each vertical line represents the location of a Colombian industry in the RCA world distribution.

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 B: Change in RCA by removing inter-industry misallocation and sectoral TFPR

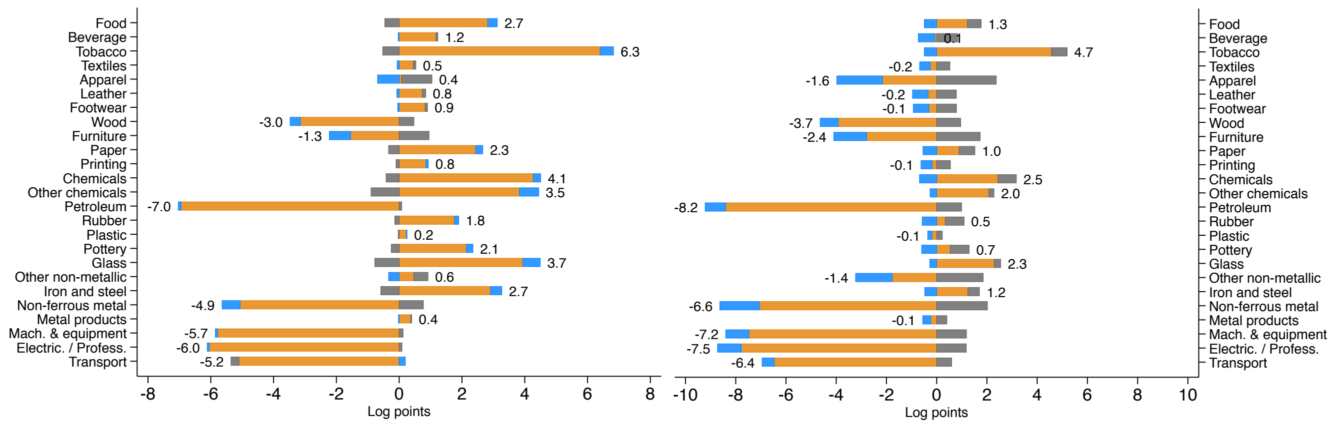


Notes: Each panel shows the change in the RCA measures in the counterfactual exercises compared to the observed amounts of factor misallocation. In Panel A, intra-industry variance of log TFPR is constructed as the weighted average of the within-industry dispersion of the factors' MRP that face misallocation: capital, skilled and unskilled labor. Similarly, in Panel B sectoral TFPR is computed using only capital, skilled and unskilled labor as inputs.

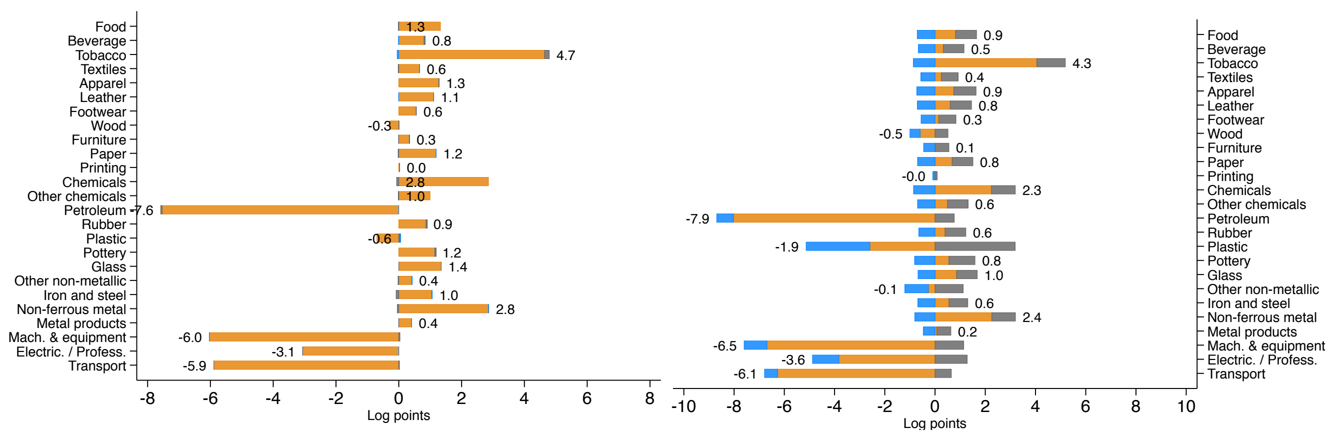
Figure 8 – Changes in determinants of Colombian RCA

Panel A: Determinants of RCA　　　　**Panel B:** Determinants of AA

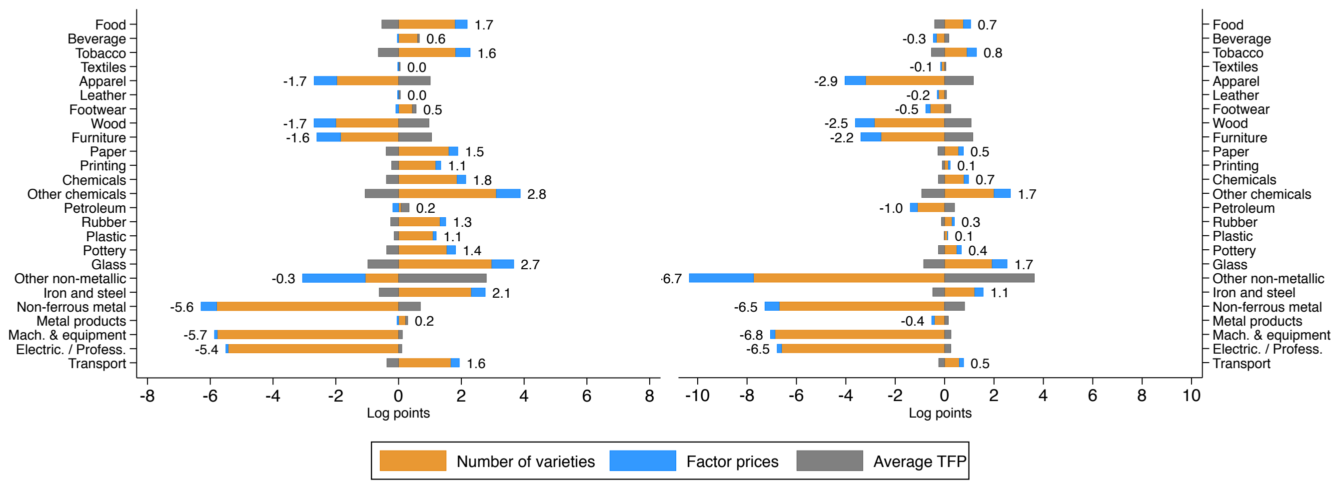
1. Removing intra- and inter-industry misallocation



2. Removing only intra-industry misallocation



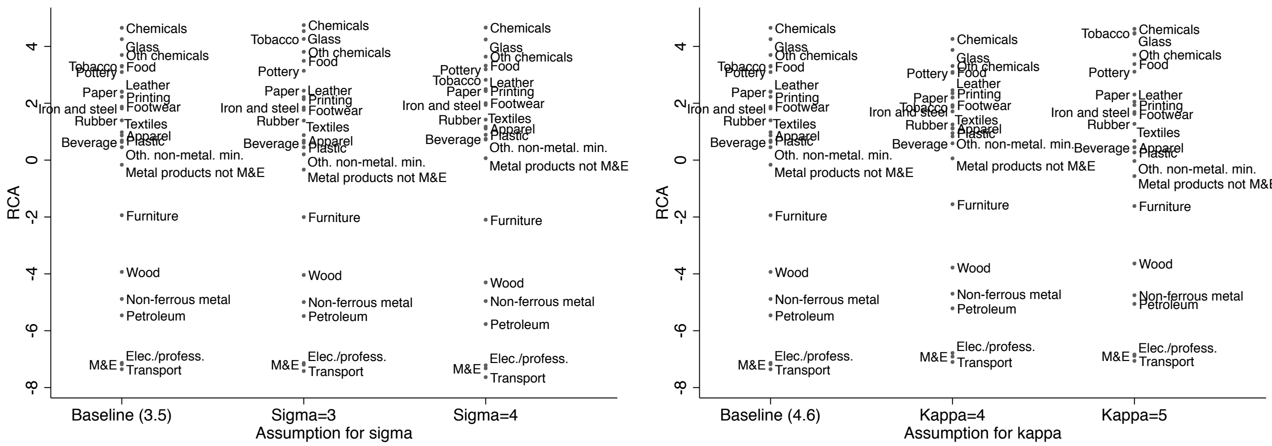
3. Removing only inter-industry misallocation



Notes: Each panel shows the decomposition of the change in the RCA measures on the three sources of export capability in the model, using equation (B.14) in Appendix B.3.

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

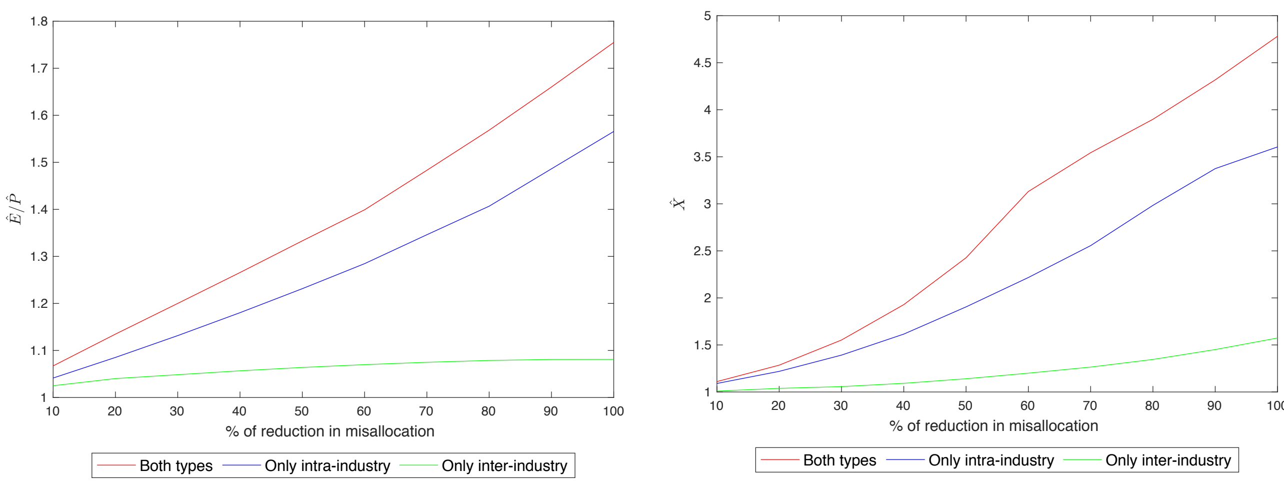
Panel A: Changes in σ **Panel B:** Changes in κ



Notes: Figure shows the ranking of the RCA measures in the counterfactual equilibria derived from removing both types of factor misallocation under different values of σ and κ.

Figure 10 – Welfare gains and export growth from gradual reforms

Panel A: Welfare gains **Panel B:** Export growth



Notes: Figure shows welfare gains and export growth in the counterfactual equilibria derived from removing gradually each type of misallocation.

Appendix For Online Publication

A Data and solution of the model

A.1 Description of the dataset

This paper uses two sets of data: A first group of two "macro" datasets with information at the country-sectoral level, and a second set of two "micro" datasets, with information at the firm level for Colombia.

Regarding the two "macro" datasets, the first one collects sectoral information of gross output and bilateral trade flows for a sample of 48 countries and 25 manufacturing industries (3-digit ISIC rev. 2 level), for the years 1992-2012, and is used in the econometric exercise of section 2.3. Tables A.1 and A.2 at the end of this section display the considered industries and countries respectively. Data for bilateral trade flows merges the CEPII's trade and production database (TradeProd), developed by de Sousa, Mayer and Zignago (2012), which is available from 1980 to 2006, with data from the CEPII's BACI database, which is described in Gaulier and Zignago (2010) and is available from 1995 onwards. While the TradeProd database has estimations for imports from home, BACI only has bilateral trade flows. To construct imports from home in the period when TradeProd is not available, I use gross output from the OECD's Trade in Value Added (TiVA) database (2018's release), which contains several indicators derived from the OECD's Inter-Country Input-Output (ICIO) database.[42] For industries that were not disaggregated enough in TiVA, total output was allocated using the share of sectoral exports in the exports of the available aggregated sector. For two countries that are not available in TiVA (Venezuela and Ecuador, which were included given their relevance as Colombia's trade partners), an imputation scheme was implemented using their total exports and the ratios of total exports to gross output observed in the closest years in TradeProd.

The second macro "dataset" focus only on 1995. It has the same variables for the same sectors and countries of the former one, but collects additional information about factors shares for each country-industry and factor intensities to be used in the counterfactual exercises of section 4. Factors shares were constructed using information from several sources. For materials, I compute the shares using 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, the ICIO database contains information of employment (measured in number of persons engaged) for 42 of the 48 countries considered here. For the remaining countries, data was collected using INDSTAT2 database. Skilled and unskilled labor shares were allocated using GTAP-5 database, which is based on labor force surveys and national censuses where

[42] 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).

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 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 (see Berlemann and Wesselhöft, 2014) on the calculation of the initial stock of capital in the perpetual inventory method 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:

$$\gamma_{iks} = \frac{\frac{GCFC_{is}}{g_{is}+\delta_{is}^{r}}}{\sum_{s}^{S}\frac{GCFC_{is}}{g_{is}+\delta_{is}^{r}}}$$

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_{is}^{r} is an exogenous depreciation rate, which are computed using the NBER-CES Manufacturing Industry database for US.[43] 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 $g_{is} = \delta_{is}^{r} = 0\ \forall i, s$ (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.

Finally for factor intensities, following the misallocation literature, I use average U.S. cost shares at the corresponding aggregation levels from the NBER-CES database. The idea is that under the assumption that the US could be regarded as one of the economies with less resource misallocation in the world, US cost shares would reflect actual differences in technology across sectors instead of inter-industry misallocation.

Regarding the "micro" datasets, for the econometric exercise of section 2.3 I use data from the Colombian Annual Manufacturing Survey (AMS), collected by the Departamento Administrative Nacional de Estadística (DANE), the Colombian national statistical agency, for the period 1991-2012. The data is provided by DANE in his webpage. The AMS is a standard census of plants with 10 or more workers or annual sales above certain limit, which is adjusted over time. For the considered time-frame, I gathered information about sales, intermediate consumption, number of production workers and total workers, total payroll and from production workers, and book values of equipment and structures. TFPQ, TFPR and factor MRP measures were constructed based on the observed factors' average revenue products, using gross-output specifications, the U.S. cost shares as factor intensities, and following the procedure to correct for measurement error explained in Appendix A.2.

[43] 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.

Anonymization procedures implemented by DANE substitute firms' information by industry's averages for firms in which their anonymity was compromised, an issue that occurs mainly in sectors with relatively few firms. I detect those cases by identifying duplicates in the data. For those firms in which its information for a given year was detected as duplicated but it was not in the previous and next year, a linear interpolation was implemented; while the remaining duplicates were dropped from the sample. Further, sectors with less than 20 firms in at least one year were dropped (tobacco and pottery). I also follow HK and remove the 1% tails of TFPR and TFPQ deviations from the industry-averages; and follow BKR and drop observations where TFPR decreases or increases by more than five times relative to the previous year. The final panel contains around 5900 plants in a typical year.

For the counterfactual exercises of section 4, in which I need information for only one year (1995), I work instead with the panel created by Eslava et al. (2004) (EHKK) for the period 1984-1998, using also the AMS. A unique feature of this panel is the availability of plant-level deflators for both total output and materials.[44] In this way, TFPQ can be obtained using its direct definition, reflecting only firm's efficiency. However, since for the counterfactual exercise I take a stance on the distribution of firm's TFPQ and I only need measures based of the factors' average revenue products, those deflators were not directly used. Instead, my preference for this dataset is due to the absence of DANE's anonymization procedures; so I can count with the full census of firms and sectors. Nevertheless, both HK and BKR cleaning procedures were implemented as in the previous "micro" dataset. Statistics about the number of firms are presented in 4.

With the goal to ensure consistency between the macro and the micro datasets used in the counterfactual, 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 macro 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.

[44]EHKK work with information at the product level (comparable to the 6-digit HS) on the value and physical quantities of outputs and inputs. This allows them to obtain prices as unit values for each output and input of every plant, and hence to construct firm-specific prices for total output and materials using Tornqvist indices.

Table A.1 – 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

Table A.2 – 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

A.2 Bils, Klenow and Ruane's (2020) method and results for Colombia

Here I succinctly introduce Bils, Klenow and Ruane's (2020) (BKR hereafter) method to estimate the extent to which the observed TFPR and factors' MRP reflect their true values in the presence of additive measurement error in revenue and inputs, which in the latter case can be also interpreted as overhead factors. BKR show that in the absence of additive measurement error, the elasticity of revenues with respect to inputs should not vary for plants with different levels of TFPR. But with additive measurement error, such elasticity is lower for plants with high observed TFPR. Particularly, the degree of how much such elasticity varies for different levels of TFPR, indicates how much measurement error overstates the dispersion of observed TFPR as a measure of the dispersion of true TFPR, call it τ. Formally, defining 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 $\hat{I}_m = I_m + g_m$, for arbitrary levels of TFPR, call them $TFPR_k$, BKR find that:

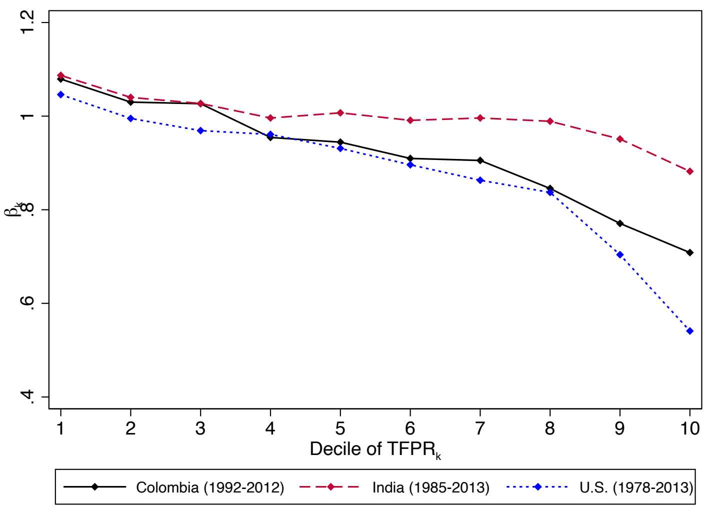
$$Var(\ln \tau_m) = Var(\ln TFPR_m) + Cov\left(\ln TFPR_k, E_k\left(\ln \frac{R_m \hat{I}_m}{I_m \hat{R}_m}\right)\right) \tag{A.1}$$

where the covariance term represents how much the elasticity of revenues with respect to inputs changes for different levels of TFPR. BKR estimate this term in the following way. First, growth rates of observed revenues $\Delta\hat{R}_m$ and inputs $\Delta\hat{I}_m$ plus observed TFPR are computed in a panel data as deviations from the sector-year average for each plant. Next, deviations of TFPR are grouped by deciles, and for plants in each decile k the following regression is computed:

$$\Delta\hat{R}_m = \hat{\lambda}_k + \hat{\beta}_k \Delta\hat{I}_m + \varepsilon_m \tag{A.2}$$

where plants are weighted within each decile by their share of total input costs. So a larger influence of measurement error implies that the estimates of $\hat{\beta}_k$ decrease faster at higher levels of TFPR, implying that the negative covariance term in (A.1) is larger in absolute terms. Figure A.1 plots the coefficients $\hat{\beta}_k$ that I find for Colombia in the period 1992-2012, and compare them to what BKR obtain for India in 1985-2013 and the U.S. in 1978-2013. Similarly to what is found by BKR for India and US, the Colombian estimated elasticities are a decreasing function of the TFPR decile, hereby suggesting the influence of measurement error. Further, the estimated values are between the values found for India and the U.S.

Figure A.1 – Elasticity of revenues to inputs by deciles of TFPR



Sources: For Colombia, own calculations; for India and the U.S. see Bils, Klenow and Ruane (2020). The figure plots the obtained $\hat{\beta}_k$ coefficients from regression (A.2) for each decile k.

Once those elasticities are obtained, the correction proceeds to use them to substitute TFPR by τ as a measure of distortion. Thus, for each plant m, TFPR is substituted by:

$$\ln \hat{\tau}_m = \ln TFPR_m + \ln \hat{\beta}_k + \varepsilon_m \tag{A.3}$$

where $\hat{\beta}_k$ is the estimated coefficient for the corresponding decile and ε_m is a log-normal draw with variance such that equation (A.1) holds.[45] It is assumed that measurement error is common for all inputs, so the same equation applies for each factor's MRP to find the "true" wedges. Also, since measurement error affects TFPQ in a proportional way to TFPR (to a factor of $\frac{\sigma-1}{\sigma}$), the same equation applies for TFPQ, corrected to such factor. Computing the values of $\hat{\tau}_m$ I find that the contribution of measurement error for the intra-industry variance of TFPR (measured as $1-\frac{Var(\ln \hat{\tau}_m)}{Var(\ln TFPR_m)}$) in Colombia is around 30% on average over the considered years, between what BKR find for India (around 26% on average) and for the US (around 63%).

A.3 External validation of misallocation measures

Here I use external data to check whether the computed measures of intra-industry misallocation are related to possible quantifiable sources of such misallocation. For example, misallocation of capital within an industry could be due to the fact that access to credit and loan conditions can vary among different firms. If this is the case, an industry with greater

[45] This is, $Var(\ln \varepsilon_m) = -Cov\left(\ln TFPR_k, \ln \hat{\beta}_k\right) - Var\left(\ln \hat{\beta}_k\right)$. In my implementation I found that the use of ε_m made the obtained metrics of misallocation for some sectors-years (those with relatively smaller number of observations) sensitive to randomization, an undesirable outcome. So I use instead a factor γ such that $\ln \hat{\tau}_m = \gamma \ln TFPR_m + \ln \hat{\beta}_k$ satisfies equation (A.1). Thus, γ is chosen such that the relation $Var(\ln TFPR_m) + Cov\left(\ln TFPR_k, \ln \hat{\beta}_k\right) = Var\left(\gamma \ln TFPR_m + \ln \hat{\beta}_k\right)$ holds, a relation that delivers a quadratic equation in γ.

capital misallocation should exhibit a larger dispersion in the cost of capital for the firms in the industry. Thus, the intra-industry dispersion of capital wedges as a measure of misallocation of capital should be related to the dispersion of the idiosyncratic cost of capital for firms within the same industry. I test this hypothesis as follows. I 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 consistently available from 2007.[46] 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 standard deviation of the interest rates on loans for each manufacturing industry and year from 2007 to 2013.[47] In Table A.3 I estimate a linear regression of those standard deviations on the intra-industry variances of the log-average revenue products of capital corrected for measurement error (following the procedure described in Appendix A.2), the measure of capital wedges used in this paper. Column (1) shows there is a positive and statistically significant correlation between both measures. The correlation is stronger when the metrics are compared only across industries (i.e. controlling for time fixed effects, second column) than when are compared for the same industry over time (i.e. controlling for sector fixed effects, third column); but in any of the two dimensions the correlation is positive and statistically significant.

Table A.3 – Variances of log wedges of capital and dispersion of loans' real interest rates

	(1) $Var(\ln(1+\theta_k))$	(2) $Var(\ln(1+\theta_k))$	(3) $Var(\ln(1+\theta_k))$
Dispersion of interest rates	0.180**	0.355***	0.126***
	(0.073)	(0.110)	(0.046)
Year FE	No	Yes	No
Sector FE	No	No	Yes
Observations	153	153	153
R-square	0.036	0.065	0.714

Notes: * p<0.10, ** p<0.05, *** p<0.01. Standardized coefficients and robust standard errors.

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.

[46]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.

[47]In the exercise I exclude sector 14 (Petroleum) because of the predominance of large state-owned firms, which arguably have different credit markets compared to firms in the remaining manufacturing sectors.

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.[48] 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 $N \times (S+L) = 1344$.

B Mathematical derivations

B.1 Model solution under Assumptions 1 and 2

Under assumptions 1 and 2, it is possible to write:

$$\begin{aligned}
\sum_{m}^{M_{ijs}} \left(\frac{a_{im}}{\Theta_{im}}\right)^{\sigma-1} &= \frac{H_{is}}{d_{is}} \int_{\theta_i} \cdots \int_{\theta_{iL}} \int_{a^*_{ijs}(\Theta)}^{\infty} \left(\frac{a_{im}}{\Theta_{im}}\right)^{\sigma-1} dG_{is} \\
&= \frac{H_{is}\kappa \bar{a}_{is}^{\kappa}}{d_{is}} \int_{\theta_{i1}} \cdots \int_{\theta_{iL}} \int_{a^*_{ijs}(\Theta)}^{\infty} a_{im}^{\sigma-\kappa-2} \Theta_{im}^{1-\sigma} dG_{is} \qquad \text{(B.1)}
\end{aligned}$$

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

Solving the inner integral and using the formula of the cutoff function in (6), (B.1) can be simplified to:

$$\sum_{m}^{M_{ijs}} \left(\frac{a_{im}}{\Theta_{im}}\right)^{\sigma-1} = \frac{H_{is}}{d_{is}} \frac{\kappa}{1+\kappa-\sigma} \left(\frac{\bar{a}_{is}}{a^*_{ijs}}\right)^{\kappa} a^{*\sigma-1}_{ijs} \Gamma_{is} \tag{B.2}$$

with Γ_{is} defined as in 18. Following a similar reasoning, it is possible to obtain:

$$\sum_{m}^{M_{ijs}} \Theta_{im} = \frac{H_{is}}{d_{is}} \left(\frac{\bar{a}_{is}}{a^*_{ijs}}\right)^{\kappa} \Gamma_{is} \tag{B.3}$$

Now, using the formulas for firm-level profits and revenues in (5), the free entry condition in (7) can be restated as:

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

Inserting (B.2), (B.3) and the definition of the productivity cutoff value for undistorted firms given in (6) in the latter expression, the free entry condition can be simplified to:

$$\sum_{j}^{N} \left(\frac{\bar{a}_{is}}{a^*_{ijs}}\right)^{\kappa} f_{ijs} = \frac{d_{is} f^e_{is}(1+\kappa-\sigma)}{\Gamma_{is}(\sigma-1)} \tag{B.4}$$

On the other hand, applying (B.2) and the definition of the productivity cutoff value, bilateral exports X_{ijs} are given by:

$$X_{ijs} = \sum_{m}^{M_{ijs}} \left(\frac{\tau_{ijs}\Theta_{im}\omega_{is}}{\rho a_{im}}\right)^{1-\sigma} E_{js} P_{js}^{\sigma-1} = \frac{\omega_{is} H_{is}}{d_{is}} \frac{\sigma\kappa}{1+\kappa-\sigma} \left(\frac{\bar{a}_{is}}{a^*_{ijs}}\right)^{\kappa} \Gamma_{is} f_{ijs} \tag{B.5}$$

Hence, from (B.4), sectoral revenues $R_{is} = \sum_{j}^{N} X_{ijs}$ are given by:

$$R_{is} = \frac{\kappa}{\rho} \omega_{is} f^e_{is} H_{is} \tag{B.6}$$

Free entry requires that aggregate sectoral profits, Π_{is}, are equal to 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 (B.3) in the definition of $\mathfrak{F}_{is}$ and using again equation (B.6), it is straightforward to obtain equation (17), the total demand of primary factor l in terms of sector revenue, underlying factor prices and HWA wedges. With the definition of v_{ils} as in the text, equation (21) is evident.

Finally, combining (B.5) with the gravity equation, I obtain:

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

By definition of the cutoff function in (6), 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:

$$\frac{a^*_{ijs}}{a^*_{i'js}} = \left(\frac{\tau_{ijs}}{\tau_{i'js}}\right) \left(\frac{\omega_{is}}{\omega_{i's}}\right)^{\frac{1}{\rho}} \left(\frac{f_{ijs}}{f_{i'js}}\right)^{\frac{1}{\sigma-1}} \tag{B.7}$$

Using (B.7) into the denominator of bilateral exports, I obtain:

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

Using (B.6) to substitute for the mass of entrants in terms of sectoral revenue, it simplifies to:

$$X_{ijs} = \frac{\omega_{is}^{-\frac{\kappa}{\rho}} R_{is} \phi_{ijs} \Gamma_{is}}{\sum_k^N \omega_{ks}^{-\frac{\kappa}{\rho}} R_{ks} \phi_{kjs} \Gamma_{ks}} E_{js} \tag{B.8}$$

where ϕ_{ijs} is as in the text. Hence, trade shares are given by (24). The model is closed combining (B.8) with the definitions of sectoral and aggregate revenues, 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, 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.8); iv) mass of entrants from (B.6); v) bilateral cutoffs values for undistorted firms from (B.5); vi) mass of operating firms from (8).

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:

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

Inserting the expression for r_{ijm} given in (5) , we can write:

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}}=\left(\rho\sum_{j}^{N}\left(\frac{\omega_{is}\tau_{ijs}}{\rho}\right)^{1-\sigma}E_{js}P_{js}^{\sigma-1}\sum_{m}^{M_{ijs}}\frac{1}{(1+\theta_{ilm})}\left(\frac{\Theta_{im}}{a_{im}}\right)^{1-\sigma}+\sum_{j}^{N}\omega_{is}f_{ijs}\sum_{m}^{M_{ijs}}\frac{\Theta_{im}}{(1+\theta_{ilm})}\right)^{-1} \tag{B.9}$$

Following a similar procedure to the used to obtain (B.2), it is possible to show that:

$$\sum_{m}^{M_{ijs}}\frac{1}{(1+\theta_{ilm})}\left(\frac{\Theta_{im}}{a_{im}}\right)^{1-\sigma}=\frac{M_{is}^{e}}{d_{is}}\left(\frac{\bar{a}_{is}}{a_{ijs}^{*}}\right)^{\kappa}a_{ijs}^{*\sigma-1}\frac{\kappa\Gamma_{ils}}{1+\kappa-\sigma} \tag{B.10}$$

$$\sum_{m}^{M_{ijs}}\frac{\Theta_{im}}{(1+\theta_{ilm})}=\frac{M_{is}^{e}\Gamma_{ils}}{d_{is}}\left(\frac{\bar{a}_{is}}{a_{ijs}^{*}}\right)^{\kappa} \tag{B.11}$$

with Γ_{ils} defined as in (18). Inserting (B.10), (B.11) and the definition of the productivity cutoff value for the undistorted firms given in (6) in (B.9), we obtain:

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

Using the free entry condition in (B.4) we get:

$$\frac{(1+\bar{\theta}_{ils})}{\rho R_{is}+\mathfrak{F}_{is}}=\left(\omega_{is}M_{is}^{e}f_{is}^{e}\frac{\Gamma_{ils}}{\Gamma_{is}}\frac{\sigma\kappa+1-\sigma}{(\sigma-1)}\right)^{-1} \tag{B.12}$$

Substituting the expression of $\sum_{m}^{M_{ijs}}\Theta_{im}$ given in (B.3) in (10) and using again equation (B.6) it is possible to derive $\rho R_{is}+\mathfrak{F}_{is}=\omega_{is}M_{i}^{e}f_{i}^{e}\frac{\sigma\kappa+1-\sigma}{(\sigma-1)}$. Inserting the latter equation in (B.12):

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

It is possible to repeat the proof to derive an expression for the HWA wedge of the exporters to destination 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:

$$\left(\frac{P_{ijs}P_{i'js'}}{P_{ijs'}P_{i'js}}\right)^{1-\sigma} = \left(\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}}\right)^{1-\sigma}\left(\frac{M_{ijs}M_{i'js'}}{M_{ijs'}M_{i'js}}\right)\left(\frac{\bar{\psi}_{ijs}\bar{\psi}_{i'js'}}{\bar{\psi}_{ijs'}\bar{\psi}_{i'js}}\right)^{1-\sigma}\left(\frac{A_{ijs}A_{i'js'}}{A_{ijs'}A_{i'js}}\right)^{\sigma-1} \tag{B.13}$$

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.13) 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.13) can be written as:

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

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

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

Under Assumptions 1 and 2 the aggregate stability condition (8) can be solved to obtain $M_{ijs}=\frac{H_{is}\Upsilon_{is}}{\delta_{is}}\left(\frac{\bar{a}_{is}}{a^*_{ijs}}\right)^{\kappa}$ with $\Upsilon_{is}=\int_{\theta_{i1}}\cdots\int_{\theta_{iL}}\Theta_i^{-\frac{k}{\rho}}dG^{\theta}_{is}(\vec{\theta})$, 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.7), I obtain for the RHS:

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

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.13):

$$=\ln\left[\frac{\rho_{is}\rho_{i's'}}{\rho_{is'}\rho_{i's}}\frac{R_{is}R_{i's'}}{R_{is'}R_{i's}}\frac{\Upsilon_{is}\Upsilon_{i's'}}{\Upsilon_{is'}\Upsilon_{i's}}\left(\frac{\omega_{is}}{\omega_{is'}}\frac{\omega_{i's'}}{\omega_{i's}}\right)^{-\frac{\kappa}{\rho}-1}\right]+\ln\left[\frac{\omega_{is}\omega_{i's'}\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\omega_{is'}\omega_{i's}\bar{\Theta}_{is'}\bar{\Theta}_{i's'}}\right]^{1-\sigma} \tag{B.14}$$

$$+\ln\left[\left(\frac{\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\bar{\Theta}_{is'}\bar{\Theta}_{i's'}}\right)^{\sigma-1}\left(\frac{\omega_{is}\omega_{i's'}}{\omega_{is'}\omega_{i's}}\right)^{\sigma}\frac{\Gamma_{is}\Gamma_{i's'}}{\Gamma_{is'}\Gamma_{i's}}\frac{\Upsilon_{is'}\Upsilon_{i's}}{\Upsilon_{is}\Upsilon_{i's'}}\right]+B_{ijs}$$

where $B_{ijs} = \ln\left[\left(\frac{\tau_{ijs}\tau_{i'js'}}{\tau_{ijs'}\tau_{i'js}}\right)^{-\kappa}\left(\frac{f_{ijs}f_{i'js'}}{f_{ijs'}f_{i'js}}\right)^{1-\frac{\kappa}{\sigma-1}}\right]$ and $\rho_{is} = \frac{\bar{a}^{\kappa}_{is}}{d_{is}f^e_{is}}$. Simplifying, it is straightforward to derive the gravity equation in (19). Equation (B.14) offers a decomposition of the exporter-industry fixed effect on the three sources of interest: number of exporters (first term in ln), average factor returns (second term in ln) and TFP (third term in ln).

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:

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

$$+\ln\left[\left(\frac{\bar{\Theta}_{is}\bar{\Theta}_{i's'}}{\bar{\Theta}_{is'}\bar{\Theta}_{i's'}}\right)^{\sigma-1}\left(\frac{\check{\omega}_{is}}{\check{\omega}_{is'}}\frac{\check{\omega}_{i's'}}{\check{\omega}_{i's}}\right)^{\sigma}\frac{\Gamma_{is}\Gamma_{i's'}}{\Gamma_{is'}\Gamma_{i's}}\frac{\Upsilon_{is'}\Upsilon_{i's}}{\Upsilon_{is}\Upsilon_{i's'}}\right]$$

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:

$$\Gamma_{is} = \int_{\theta_i}\cdots\int_{\theta_{iL}}\Theta_i^{1-\frac{\kappa}{\rho}}dG^{\theta}_{is} = E\left[\prod_l^L(1+\theta_{il})^{\left(1-\frac{\kappa}{\rho}\right)\alpha_{ls}}\right]$$

Assume $\vec{\theta}_{is} = \{\theta_{i1s},\theta_{i2s},...\theta_{iLs}\}$ has a multivariate log-normal distribution, such the transformed vector $\vec{\theta}^*_{is} = \{\ln(\theta_{i1s}),\ln(\theta_{i2s}),...\ln(\theta_{iLs})\}$ 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: $\vec{\alpha}_s = \left\{(1-\frac{\kappa}{\rho})\alpha_{1s},(1-\frac{\kappa}{\rho})\alpha_{2s},...,(1-\frac{k}{\rho})\alpha_{Ls}\right\}'$. 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:

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

By the same token, let $\vec{\alpha}_{ls}$ a (column) vector with elements:

$$\vec{\alpha}_{ls} = \left\{\left(1-\frac{\kappa}{\rho}\right)\alpha_{1s}, ..., \left(1-\frac{\kappa}{\rho}\right)\alpha_{ls}-1, ..., \left(1-\frac{\kappa}{\rho}\right)\alpha_{Ls}\right\}'$$

this is, $\vec{\alpha}_{ls}$ has the same elements of $\vec{\alpha}_s$ with exception to the element in position l, which is $\left(1-\frac{k}{\rho}\right)\alpha_{ls}-1$. Thus $(1+\theta_{il})^{(1-\frac{\kappa}{\rho})\alpha_{ls}-1} \prod_{h\neq l}^{L} (1+\theta_{ih})^{(1-\frac{\kappa}{\rho})\alpha_{hs}}$ is log-normal distributed with location and shape parameters $(\vec{\alpha}_{ls})'\vec{\mu}_{is}$ and $(\vec{\alpha}_{ls})'V_{is}\vec{\alpha}_{ls}$. 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 $\left(1+\bar{\theta}_{ils}\right)$ in (18) we obtain:

$$\begin{aligned}\ln\left(1+\bar{\theta}_{ils}\right) &= (\vec{\alpha}_s)'\vec{\mu}_{is} + \frac{1}{2}(\vec{\alpha}_s)'V_{is}\vec{\alpha}_s - (\vec{\alpha}_{ls})'\vec{\mu}_{is} - \frac{1}{2}(\vec{\alpha}_{ls})'V_{is}\vec{\alpha}_{ls} \\ &= \mu_{ils} + \frac{1}{2}\left[(\vec{\alpha}_s)'V_{is}\vec{\alpha}_s - (\vec{\alpha}_{ls})'V_{is}\vec{\alpha}_{ls}\right] \qquad \text{(B.15)}\end{aligned}$$

For the case of two production factors and zero covariances, equation (B.15) reduces to equation (20) in the text.

B.5 Welfare

Combining the formula of the consumer price index in sector s and equation (B.2) we obtain:

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

Inserting the definition of the productivity cutoff value for the undistorted firms in (6) in the term $a_{kis}^{*\sigma-1-\kappa}$, the price index can be written as:

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

Using the country i's share of expenditure on itself within sector s from equation (B.8):

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

where $\varsigma_{ijs} = \left(\frac{\rho\bar{a}_{is}}{\tau_{ijs}}\right)^{\kappa}\frac{1}{d_{is}f_i^e}\left(\frac{1}{f_{iis}}\right)^{1-\frac{\kappa}{\sigma-1}}\left(\frac{\kappa}{1+\kappa-\rho}\right)$ 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:

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

Using the fact that $\hat{P}^d_i = \prod_s (\hat{P}^d_{is})^{\beta_s}$, $\hat{E}_{is} = \hat{E}_i$ 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 $\hat{P}^d_{is} = \hat{w}_i (\frac{\hat{\pi}_{iis}}{\hat{Z}_{is}})^{\frac{1}{k}}$, which leads to the Arkolakis, Costinot and Rodríguez-Clare (2012)'s formula to compute the increase in welfare in response to any exogenous shock.

Online Appendix References

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

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.

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2020. "Misallocation or mismeasurement?" National Bureau of Economic Research Working Paper 26711.

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.

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.

Gaulier, Guillaume, and Soledad Zignago. 2010. "BACI: International trade database at the product-level. The 1994-2007 version." CEPII Working Papers 2010-23.

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.

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.