Misallocation and Intersectoral Linkages*

— Link to the most recent version —

Sophie Osotimehin
University of Québec at Montreal
osotimehin.sophie@uqam.ca

Latchezar Popov
Texas Tech University
latchezar.popov@gmail.com

February 15, 2021

Abstract

We analytically characterize the aggregate productivity loss from distortions in the presence of sectoral production linkages. We find that accounting for low input substitutability reduces the productivity loss and the impact of intermediate-input suppliers. Moreover, with elasticities below one (i.e. below Cobb-Douglas), sectoral linkages do not systematically amplify the productivity loss. We quantify these effects in the context of the distortions caused by market power, using industry-level data for 35 countries. With our benchmark calibration, the median aggregate productivity loss from industry-level markups is 1.2%; assuming Cobb-Douglas elasticities would lead to overestimating the productivity loss by a factor of 1.8.

*We thank Simon Anderson, Susanto Basu, Vasco Carvalho, Kerem Cosar, Pete Klenow, Guy Laroque, Julien Martin, Ellen McGrattan, Mathieu Parenti, Jorge Miranda Pinto, Toshihiko Mukoyama, B. Ravikumar, Pierre Sarte, Venky Venkateswaran, Enghin Atalay; participants at various seminars and conferences. Sophie Osotimehin thanks the Opportunity and Inclusive Growth Institute of the Minneapolis Fed for their support and hospitality. We thank Jaeki Jang, Ilhan Guner, and Diego Legal-Cañisá for excellent research assistance; and Susan Langenkamp and James Holt for editorial assistance.

1 Introduction

The production of goods and services involves a complex web of firms connected through supply chains. For example, most of the items found in grocery stores are produced by food manufacturers using farm products and products from other food manufacturers, as well as various other intermediate inputs such as energy, transportation, and business services. These production linkages across firms and sectors are key to understanding aggregate productivity, especially when the firms' input decisions may be distorted by market frictions. Frictions in one sector may ripple outward through the economy, as the frictions not only affect the distorted firms' input choice but may also indirectly affect the decisions of the other firms along the production chain. In this paper, we study how aggregate productivity is affected by frictions that distort the allocation of inputs across firms, taking into account production chains and the sectoral linkages of production. Findings from recent work, which highlight how sectoral linkages amplify and propagate various sector- or firm-level shocks (e.g., Jones, 2011; Acemoglu et al., 2012; Bigio and La'o, 2020; Atalay, 2017), may suggest that the interconnection between sectors amplifies the consequences of such distortions. Do sectoral linkages always amplify the productivity loss from allocative distortions? Do sectors supplying intermediate inputs have a larger impact than other sectors? Which types of linkages are relevant to understanding the effects of distortions on aggregate productivity?

We show that the answers to all these questions depend crucially on how substitutable inputs are. We provide a theoretical characterization and a quantitative evaluation of how input substitutability shapes the effect of distortions and the role of the linkages between sectors. We find that the complementarities in the production process (i.e., low input substitutability) mitigate the effects of distortions on aggregate productivity, reduce the role of sectoral linkages, and reduce the impact of intermediate-input suppliers. Furthermore, we find that sectoral linkages do not systematically amplify the effects of distortions; we derive the conditions under which the productivity loss from distortions is smaller than in an otherwise identical economy without sectoral linkages. Our results indicate that abstracting from input complementarity leads to overestimating the effects of distortions and the strength of the amplification from sectoral linkages.

We investigate the effects of distortions and the role of input substitutability using a multi-sector model with intersectoral linkages of production. The model builds on

Long and Plosser (1983) and is similar to Jones (2011) but allows for more flexible production functions and a richer input-output structure. Each sector combines primary inputs (labor) with intermediate inputs using a nested constant-elasticity-of-substitution production function, with possibly non-unitary elasticity of substitution between the intermediate inputs and between the primary input and the intermediate-input bundle. We consider distortions in the use of primary or intermediate inputs. These distortions could be the result of frictions in financial, input, or product markets (such as borrowing constraints, hiring and firing costs, monopoly power). Following Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), we do not model the sources of the distortions. Rather, we model distortions, in a reduced-form manner, as wedges that lead firms to deviate from the frictionless input allocation.

We provide an analytical characterization of the effects of distortions on aggregate total factor productivity (TFP) in the presence of sectoral linkages, using second-order approximations to TFP. We separately study two types of distortions, depending on whether they affect only primary inputs (like labor-market frictions) or all inputs uniformly (like markups). Our analytical characterizations highlight the role of the interaction between the firms' distortions and the linkages across firms. How this interaction shapes the TFP loss depends on the degree of input substitutability. We derive three main results concerning how input substitutability determines (i) the TFP loss from distortions, (ii) the amplification from sectoral linkages, and (iii) the impact of each sector.

Our first theoretical result is that input complementarity mitigates the effects of the two types of distortions (primary and all-input) on aggregate productivity. This result may seem counter-intuitive. With less flexibility to substitute less-distorted for distorted inputs, one could expect distortions to have a larger impact on aggregate productivity. The role of the elasticity of substitution is complex and is in general non-monotonic. Different opposing forces are at work. On the one hand, higher complementarity amplifies the effect of a change in the firms' inputs on their output; on the other hand, it reduces the effect of the distortions on the firms' input decision because firms respond less to the (distortion-induced) change in the relative prices when the elasticity of substitution is low. With intermediate inputs, additional effects further complicate the analysis. Using our analytical characterizations, we show that the second force dominates, and higher complementarity leads to a smaller aggregate productivity loss.

Low elasticities' mitigating effect for distortions stands in stark contrast with their amplification of negative productivity shocks. In line with previous work, we show that a lower elasticity amplifies negative sectoral productivity shocks.[1] This opposite effect makes it important to be clear on the nature of the frictions when studying their effects on aggregate productivity. When frictions distort relative prices (like in this paper), accounting for input complementarity reduces the TFP loss, whereas when frictions entail a real resource cost (and hence are akin to lower productivity), accounting for input complementarity increases the TFP loss.

Our second theoretical result is that sectoral linkages do not systematically amplify the effects of distortions. We establish the conditions under which sectoral linkages dampen the effects of distortions and find that these conditions are more likely to be satisfied when the distortions affect only primary inputs. With primary-input distortions, the dampening effect occurs if inputs are less substitutable than Cobb-Douglas (and provided sectoral linkages do not affect too much the size of each sector). Our result implies that distortions that affect all inputs, such as markups, are more likely to be amplified by the sectoral linkages than distortions that bear exclusively or mainly on primary inputs, such as labor-market frictions or financial frictions.

Our third theoretical result delves deeper into the role of sectoral linkages. We show how each sector's impact depends on the sector's production linkages. Again, the value of the elasticity of substitution is crucial to understanding which linkages are the most relevant. When the elasticity of substitution is lower, the aggregate impact of distortions in sectors that supply intermediate inputs is attenuated; as a result, the relative impact of final-output suppliers becomes larger. Moreover, we find that the interconnection between distorted sectors plays a smaller role when the elasticity of substitution in lower.

We quantify the role of input substitutability in the context of the distortions caused by market power. Markups create a wedge between the firms' price and their marginal cost and are hence isomorphic to a distortion that uniformly affects all inputs. We calibrate the model on industry-level data from the World Input-Output Database for 35 countries. We measure industry-level markups using the price-cost margins, and we calibrate the production function parameters for each country separately, accounting for the presence of markups and for the value of the elasticities of substitution. Estimates of

[1]Note that because a lower elasticity amplifies negative sectoral productivity shocks, it would also amplify the effects of distortions that reduce sectoral TFP.

these elasticities point to complementarities between inputs. Following Atalay's (2017) estimates, in our benchmark calibration we set the elasticity of substitution between intermediates to 0.01 and the one between labor and the intermediate-input bundle to 0.7. We compute the aggregate productivity gain from reducing industry-level markups to zero, and we find a productivity gain of 1.2% for the median country and one higher than 4% for Turkey, Indonesia, and Mexico. These numbers are conservative estimates of the TFP loss from markups since they abstract from the *within*-sector misallocation caused by markups.

In addition to providing estimates of the cost of industry-level markups, the quantitative analysis, computed used the exact solution of the model, complements our theoretical results by showing the role of input substitutability for the sizable distortions observed in the data and by quantifying the role of the two elasticities of substitution.

First, the quantitative analysis shows that the TFP gains are lower when the model is calibrated with a lower input substitutability than the benchmark, and the TFP gains are higher when the model is calibrated with a higher input substitutability. In particular, the Cobb-Douglas specification (i.e., unit elasticity of substitution between all inputs) leads to overestimating the TFP gain for the median country by a factor of 1.8 relative to the benchmark calibration. Second, we find that the strength of the amplification from sectoral linkages increases with the degree of input substitutability. The ratio of the TFP gain in the economy with and without sectoral linkages is equal to 5.4 under Cobb-Douglas and to 2.9 with our benchmark calibration. Accounting for complementarities in the production process still leads to a sizable amplification from sectoral linkages, although not as large as suggested by the Cobb-Douglas specification. Furthermore, had the distortions affected only primary inputs, there would have been no amplification effect. In that case, the TFP loss is 30 percent *lower* relative to the otherwise identical economy without sectoral linkages. Finally, we find that the Cobb-Douglas specification leads to the overestimation of the impact of sectors that supply primarily intermediate inputs, such as mining, the non-metallic mineral products industry, and the renting-of-material-and-equipment industry. All in all, a central message of the paper is that the Cobb-Douglas specification, ubiquitous in the literature on sectoral linkages of production, leads to greatly overestimating the effects of distortions on aggregate productivity as well as the role of sectoral linkages.

Our paper contributes to the literature on the consequences of allocative distortions on aggregate productivity. This literature, initiated by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), focuses on the allocation of labor and capital; usually, intermediate inputs are not explicitly modeled. By deriving analytical expressions for the effects of distortions, we complement Osotimehin (2019), who derives similar expressions in a setting without intermediate inputs. Our paper is closely related to Jones (2011, 2013) and several related papers such as Bigio and La'o (2020), Bartelme and Gorodnichenko (2015), Fadinger et al. (2015), Caliendo et al. (2017), Grassi (2017), Liu (2019), Luo (forthcoming), Leal (2017), Wang (2020), and Hang et al. (2020), which investigate the effects of distortions in the presence of intersectoral linkages. Like these papers, we highlight how the linkages between firms can give rise to an additional mechanism through which distortions reduce aggregate TFP. We complement these papers by showing that this mechanism is weaker when input complementarities are accounted for and can even lead to a smaller TFP loss than in an otherwise identical economy without intersectoral linkages.

Contrary to Jones (2011), who highlights that input complementarity *amplifies* the effects of distortions on aggregate TFP, we show that input complementarity actually *mitigates* the effects of distortions. Complementarity has opposite implications for productivity and distortions and jointly analyzing them, as done in Jones (2011), may mask the mitigating effect of complementarity on distortions. Contemporaneous papers by Boehm and Oberfield (2020), Miranda-Pinto and Young (2021), and Baqaee and Farhi (2020) (BF) also study the effects of distortions in the context of a production network with non-unitary elasticities, but their focus is different than ours.[2] Boehm and Oberfield (2020) consider the effects of enforcement frictions in the market for inputs. Miranda-Pinto and Young (2021) study the role of the cross-sectoral heterogeneity in input substitutability in the context of financial frictions. Baqaee and Farhi (2020) (BF) focus on presenting their decomposition of aggregate productivity growth.[3] By contrast with BF, we focus on analyzing the TFP loss from distortions. We prove three novel theoretical results, which provide insights on how input substitutability shapes the role of sectoral linkages and the impact of each sector. None of these results are derived

[2]Other recent papers emphasize non-unitary elasticity of substitution between inputs, but with a focus on the propagation of productivity shocks. See for example, Atalay (2017), Carvalho et al. (2015), Miranda-Pinto (2021), and Baqaee and Farhi (2019).

[3]We discuss in detail how our work relates to Baqaee and Farhi (2020) and how our characterizations differ from their formula in Appendix A.8.

in BF. Moreover, although BF also derive a formula for the TFP loss from distortions, their formula is more difficult to interpret and does not convey the intuition behind our results. For example, BF's formula suggests that sectoral linkages magnify the TFP loss, whereas our characterizations show that sectoral linkages can dampen the TFP loss.

The paper also contributes to the literature on the macroeconomic implications of market power. The interest in the topic, which goes back at least to Harberger (1954), has recently surged with the debate on the decline in competition and the rise in market power in the US and other countries.[4] Several recent papers have computed the welfare costs of markups and the consequences of market power on macroeconomic outcomes in various settings (Basu, 1995; Dhingra and Morrow, forthcoming; Behrens et al., 2020; Edmond et al., 2019; Peters, forthcoming; Epifani and Gancia, 2011). Most of these papers abstract from intermediate inputs. Exceptions are Basu (1995), who shows that a uniform markup reduces aggregate productivity in the presence of intermediate inputs, and Edmond et al. (2019), who highlight that most of the welfare loss from markups is similar to that of a uniform output tax. We contribute to this literature by showing how sectoral linkages and input substitutability jointly determine the TFP loss from markups. Moreover, whereas the estimates found in the literature are typically for the US, we provide estimates for 35 countries of the cost of sectoral distortions caused by market power.

The paper is organized as follows. We present the model and the general characterization of aggregate productivity in Section 2. Then, in Section 3, we characterize the aggregate productivity loss from distortions and study the role of input substitutability. In Section 4, we present the results of the quantitative application, in which we quantify our theoretical results and compute the effects of markups on aggregate productivity in 35 countries. We offer concluding remarks in Section 5.

[4]Harberger (1954) finds removing the dispersion of markups across US manufacturing industries (observed in the late 1920s) would yield a negligible welfare gain. For the recent debate on the rise of market power, see, for example, de Loecker et al. (forthcoming).

2 Aggregate productivity in an economy with sectoral linkages and distortions

We consider a multi-sector model in which firms buy and sell intermediate inputs to each other. The sectoral linkages, which result from the firms' purchase and sale of intermediate inputs, are a key feature of the model. In this section, we describe the model and characterize the aggregate productivity in the presence of distortions in the allocation of inputs.

2.1 Model

The model, which is a generalization of the production side of Long and Plosser (1983), shares similarities with Jones (2011).

Production. The economy consists of n sectors. In each sector, a representative firm produces goods using labor and intermediate goods with the CES production function

$$Q_i = A_i \left[(1-\alpha_i)^{1-\sigma} (B_i L_i)^{\sigma} + \alpha_i^{1-\sigma} X_i^{\sigma} \right]^{\frac{1}{\sigma}}, \tag{1}$$

where B_i is the labor-augmenting productivity, A_i is the Hicks-neutral productivity, and L_i is labor.[5] The intermediate-input bundle is given by

$$X_i = \left(\sum_j v_{ij}^{1-\rho} X_{ij}^{\rho} \right)^{\frac{1}{\rho}}, \tag{2}$$

where X_{ij} is the quantity of intermediate goods from sector j used by sector i. We impose $\alpha_i \in [0, 1)$, $v_{ij} \in [0, 1]$, and $\sum_{j=1}^{n} v_{ij} = 1$ for all $i = 1, ..., n$.[6] With the heterogeneity in α_i and v_{ij}, sectors can differ in both their overall use of intermediate inputs and their mix of intermediate inputs. The key parameters of our analysis are $\rho \in (-\infty, 1)$, which determines the elasticity of substitution across intermediate goods; and $\sigma \in (-\infty, 1)$,

[5] For simplicity, we do not explicitly model capital. The primary input L_i can be thought of as a capital-labor bundle.

[6] Note that in our specification, we have $\alpha_i^{1-\sigma}$ and $(1-\alpha_i)^{1-\sigma}$, rather than α_i and $1-\alpha_i$, and similarly $v_{ij}^{1-\rho}$ instead of v_{ij}. This specification implies that measuring inputs at different levels of aggregation has no artificial effect on productivity. Also, we assume that every sector uses a positive amount of labor. This assumption simplifies the proofs and can be relaxed.

which determines the elasticity of substitution between labor the intermediate-input bundle. The corresponding elasticities are given by $\varepsilon_\rho \equiv 1/(1-\rho)$ and $\varepsilon_\sigma \equiv 1/(1-\sigma)$. The Cobb-Douglas production function corresponds to $\sigma = \rho = 0$ (or equivalently, $\varepsilon_\rho = \varepsilon_\sigma = 1$).[7] As highlighted by Klump et al. (2012), performing comparative statics on the elasticity of substitution requires a careful normalization of the CES production function.[8] We show in Appendix A.2 that our formulation is equivalent to normalizing the production function by the frictionless allocation. This normalization point is a natural choice given the focus of the paper.

Final output (consumption) is an aggregate of the goods from the different sectors,

$$Y = \prod_{i=1}^{n} \beta_i^{-\beta_i} \prod_{i=1}^{n} C_i^{\beta_i}, \tag{3}$$

with $\beta_i \in [0,1]$, $\sum_{i=1}^{n} \beta_i = 1$ and where $\prod_{i=1}^{n} \beta_i^{-\beta_i}$ is a convenient normalization that simplifies the expressions of the model's solution without affecting the results.

In this economy, the sectoral linkages are governed by two sets of parameters, collected in the vector $\alpha = (\alpha_1, ..., \alpha_n)'$ and in the matrix $V = (v_{ij})$. Together with the vector of final-good parameters, $\beta = (\beta_1, ..., \beta_n)'$, α, and V shape the production network.

Resource constraints. The resource constraints are given by

$$\sum_{i=1}^{n} X_{ij} + C_j = Q_j, \forall j = 1, ..., n \tag{4}$$

$$\sum_{i=1}^{n} L_i = \bar{L}. \tag{5}$$

The first equation states that each sector's gross output is equal to the sum of all its uses as an intermediate (including by itself) and its use in final consumption. The second equation is the labor resource constraint, with $\bar{L}$ the total (exogenous) supply of labor.

Distortions. The firms' input choices may be distorted by market frictions. Following Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), we model the distortions as wedges in the firms' first-order conditions. We define the distortions as deviations from the social-planner (first-best) allocation. We denote the Lagrange mul-

[7]The Cobb-Douglas specification is given by: $Q_i = A_i(B_i L_i/(1-\alpha_i))^{1-\alpha_i}(X_i/\alpha_i)^{\alpha_i}$ with $X_i = \prod_{j=1}^{n}(X_{ij}/v_{ij})^{v_{ij}}$.

[8]Equilibrium output should be invariant to the elasticity at the normalization point.

tiplier associated with the resource constraint of product i by λ_i and the one associated with labor by η. The allocation of inputs is characterized by the following distorted first-order conditions:

$$\beta_i \frac{Y}{C_i} = \lambda_i \tag{6}$$

$$\lambda_i A_i^\sigma \left(\frac{\alpha_i Q_i}{X_i}\right)^{1-\sigma} \left(\frac{\upsilon_{ij} X_i}{X_{ij}}\right)^{1-\rho} = \lambda_j (1+\tau_{Xij}) \tag{7}$$

$$\lambda_i A_i^\sigma B_i^\sigma \left(\frac{(1-\alpha_i) Q_i}{L_i}\right)^{1-\sigma} = \eta (1+\tau_{Li}). \tag{8}$$

Any allocation that satisfies the production and resource constraints can be rationalized by some collection of distortions.[9] Production efficiency is achieved when $\tau_{Xij} = 0$ and $\tau_{Li} = \bar{\tau}_L$ for all $i, j = 1, ..., n$. Note that the distortions τ_{Xij}, τ_{Li} are not isomorphic to productivity shocks; they reduce aggregate productivity by modifying the perceived marginal cost of inputs and thereby distorting the allocation of inputs across firms.

The distortions encompass various sources of frictions. We allow the distortions for intermediate goods τ_{Xij} to vary by both purchasing industry i (taxes or subsidies, financial frictions) and supplying industry j (for example, the ease of contract enforcement may vary across commodities, as explored by Nunn 2007 and later by Mukoyama and Popov 2020). While we solve the model with these general distortions, our analysis will focus on two types of distortions: distortions on labor only, τ_{Li}; and distortions that uniformly affect all the inputs, $\tau_{Xij} = \tau_{Li} = \tau_i$, $\forall j = 1, ..., n$. In the quantitative application, we study the consequences of market power and the distortions caused by markups, which are isomorphic to all-input distortions.

Aggregate production function. The aggregate production function relates the maximum final output that can be obtained given the distortions defined in equations (6)–(8) and the production possibilities determined by equations (1)–(5). We call this problem *the distorted planner's problem.*[10]

[9] We show in Appendix A.1 that the absence of distortion in equation (6) is without loss of generality: for any allocation characterized by distortions in equation (6), we can find a transformation of the distortions and multipliers that allows us to rewrite the distorted first-order conditions in the form of equations (6) to (8). We also show that there exists a unique set of distortions (up to an affine transformation of the labor distortions) that can rationalize a feasible allocation.

[10] Solving the distorted planner's problem is equivalent to solving the competitive equilibrium in which $\{\tau_{Li}, \tau_{Xij}\}_{i,j=1,...,n}$ are taxes rebated lump sum to the final consumers. We prefer the distorted planner's formulation because it makes clear that whether distortions reduce profits (like taxes) or increase profits

2.2 Aggregate productivity

We solve the distorted planner's problem and characterize aggregate productivity in both the general case and the case where the two elasticities of substitution (between intermediates, and between labor and the intermediate-input bundle) are equal. The general solution is used in the quantitative analysis, whereas the equal-elasticity case, which yields a closed-form expression for aggregate productivity, is used in the theoretical analysis. All the proofs are in Appendix B.

2.2.1 The general case

The following proposition gives the aggregate production function and the aggregate productivity of the distorted economy.

Proposition 1 *Suppose that a solution to the distorted planner's problem exists. Then final output is given by* $Y = TFP\ \bar{L}$*, aggregate productivity is* $TFP = [d'(I - M')^{-1}c]^{-1}$*, the vector of gross output is given by* $Q = (I - M')^{-1}cY$*, final consumption of good* i *is* $C_i = c_i Y$*, and labor inputs are given by* $L_i = d_i Q_i$*, and intermediate inputs by* $X_{ij} = m_{ij} Q_i$*; where* I *denotes the identity matrix and*

i. d is a column vector with

$$d_i = (1 - \alpha_i)(A_i B_i)^{\frac{\sigma}{1-\sigma}} \left(\frac{\lambda_i}{(1 + \tau_{Li})\eta} \right)^{\frac{1}{1-\sigma}};$$

ii. M is a matrix with elements

$$m_{ij} = A_i^{\frac{\sigma}{1-\sigma}} \alpha_i v_{ij} \left[\sum_j v_{ij} \left(\frac{\lambda_i}{(1 + \tau_{Xij})\lambda_j} \right)^{\frac{\rho}{1-\rho}} \right]^{\frac{\sigma-\rho}{\rho(1-\sigma)}} \left(\frac{\lambda_i}{(1 + \tau_{Xij})\lambda_j} \right)^{\frac{1}{1-\rho}};$$

iii. c is a column vector with $c_i = \beta_i / \lambda_i$*;*

iv. λ *is a column vector that contains the Lagrange multiplier and which is given by* $\lambda = \exp(\log(\lambda/\eta) - \beta' \log(\lambda/\eta))$*; and*

(like markups) is irrelevant for aggregate TFP.

v. the Lagrange multipliers solve

$$(\lambda_i/\eta) = A_i^{-1}\left[(1-\alpha_i)B_i^{\frac{\sigma}{1-\sigma}}(1+\tau_{Li})^{-\frac{\sigma}{1-\sigma}} + \alpha_i\left[\sum_{j=1}^{n} v_{ij}\left((1+\tau_{Xij})\lambda_j/\eta\right)^{-\frac{\rho}{1-\rho}}\right]^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}}\right]^{-\frac{1-\sigma}{\sigma}}. \tag{9}$$

The distortions affect aggregate productivity through several channels. In addition to the well-known effects on the allocation of labor (or any primary inputs) across firms, distortions also modify the allocation of intermediate inputs across firms, both directly and indirectly, by modifying the relative shadow price of goods.

Normalization. The parameters $\alpha_i, v_{ij}, A_i, B_i$ are not invariant to the choice of units (except in the Cobb-Douglas case). The following result shows that we can switch the units in a way to normalize all the productivity terms to one. This normalization is convenient since it implies a frictionless aggregate productivity of one. In Appendix A.2 we show how this productivity normalization relates to the CES normalization proposed by Klump et al. (2012).

Proposition 2 *Suppose that* $\alpha_i A_i^{\frac{\sigma}{1-\sigma}} < 1, \forall i$. *Let* $Q_i' = k_i Q$ *and* $Y' = k_Y Y$ *for some arbitrary vector of positive scaling constants* $\{k_i\}_{i=1,\dots,n}, k_Y$. *Then, there exists a set of positive constants* $\{k_i\}_{i=1,\dots,n}, k_Y$ *such that*

$$Q_i' = \left[(1-\alpha_i')^{1-\sigma}L_i^{\sigma} + \alpha_i'^{1-\sigma}\left(\sum_j {v_{ij}'}^{1-\rho}{X_{ij}'}^{\rho}\right)^{\frac{\sigma}{\rho}}\right]^{\frac{1}{\sigma}},$$

where $X_{ij}' = k_j X_{ij}$ *and* $\alpha_i' \in [0,1], v_{ij}' \in [0,1], \sum_{j=1}^{n} v_{ij}' = 1$. *Moreover,*

$$Y' = \prod_{i=1}^{n} \beta_i^{-\beta_i} C_{ii}'^{\beta}.$$

Moreover, when all distortions are equal zero, we have: $Y' = \bar{L}$, $\lambda_i' = 1\ \forall i, \eta' = 1$ *and* $\alpha_i' = \sum_j \lambda_j X_{ij}/(\lambda_i Q_i')$, *and* $v_{ij}' = \lambda_j X_{ij}'/\sum_k \lambda_k X_{ik}'$.

In the frictionless normalized economy, $\alpha_i v_{ij}$ corresponds to the elasticity of the output of i to intermediate j and is equal also to the cost share of good j in the production of

i. Note that we assumed in Proposition 1 that an equilibrium exists. We now turn to the question of the existence of the equilibrium.

Existence and uniqueness. A solution to the distorted planner's problem does not always exist. First, since the CES production function does not satisfy the Inada conditions an equilibrium may not exist even without distortions. Second, when the elasticities of substitution are smaller than one, sufficiently large positive intermediate-goods distortions would lead to infinitely large relative prices (λ_i/η); on the other hand, sufficiently large negative distortions (in absolute value) can lead the demand for intermediate goods to exceed the feasible gross output.[11] In Proposition 2, we show that if a restrictions on the parameters is satisfied then the equilibrium exists (and the change of units behind the normalization is feasible.) We can hence show the existence and uniqueness of the distorted equilibrium when the productivity terms can be normalized to one and the vector of distortions is sufficiently small. In the quantitative application, we find a solution exists even for the sizable distortions observed in the data.

Proposition 3 *Suppose that $\alpha_i A_i^{\frac{\sigma}{1-\sigma}} < 1, \forall i = 1, ..., n$. Then, a solution to the distorted planner's problem exists and is unique in an open neighborhood of distortions containing the zero-distortions vector.*

2.2.2 The equal-elasticity case

When the two elasticities of subsitution are equal, $\sigma = \rho$, we obtain a closed-form solution for aggregate productivity. We present here the case $\sigma = \rho \neq 0$ and describe the Cobb-Douglas case ($\sigma = \rho = 0$) in Appendix A.3.

Proposition 4 *Suppose that $\sigma = \rho$ and $\sigma \neq 0$. Then, aggregate output is given by*

$$\log Y = \log \bar{L} + \log TFP,$$

[11] This is easiest to see in a simple example with only one sector and Cobb-Douglas production. In this case, $\lambda/\eta = (1+\tau_X)^{\frac{1}{1-\alpha}}$ and $X/Q = \alpha/(1+\tau_X)$. Then, if $\tau_X < -(1-\alpha)$, $X/Q > 1$, demand for intermediate goods exceeds gross output, which is clearly impossible. Note that in this example, the Inada conditions on the production function hold and an equilibrium still may not exist.

aggregate productivity is given by

$$\log TFP = \frac{1-\sigma}{\sigma}\beta' \log(\hat{\lambda}) - \log(\hat{c}'[I - \tilde{\alpha}\hat{A}(\Delta^{Xq} \circ V)]^{-1}\hat{A}\hat{B}\Delta^{Lq}(\mathbf{1}-\alpha)) \tag{10}$$

and the vector of Lagrange multipliers is given by

$$\hat{\lambda} = [I - \tilde{\alpha}\hat{A}(\Delta^{Xp} \circ V)]^{-1}\hat{A}\hat{B}\Delta^{Lp}(\mathbf{1}-\alpha); \tag{11}$$

where $\circ$ *denotes the Hadamard (entrywise) product,* I *is the identity matrix,* $\mathbf{1}$ *is a vector of ones,* $\tilde{\alpha}, \hat{A}, \hat{B}, \Delta^{Xp}, \Delta^{Xq}$*,* Δ^{Lp} *,* Δ^{Lq} *are square matrices and* $\hat{\lambda}, \hat{c}$ *are column vectors defined as follows:* $\tilde{\alpha}_{ii} = \alpha_i$ *and* $\tilde{\alpha}_{ij} = 0$ $\forall i \neq j$*;* $\hat{A}_{ii} = A_i^{\frac{\sigma}{1-\sigma}}$ *and* $\hat{A}_{ij} = 0$ $\forall i \neq j$*,* $\hat{B}$ *is defined similarly;* $\Delta^{Xp}_{ij} = (1+\tau_{Xij})^{-\frac{\sigma}{1-\sigma}}$*,* $\Delta^{Xq}_{ij} = (1+\tau_{Xij})^{-\frac{1}{1-\sigma}}$*;* $\Delta^{Lp}_{ii} = (1+\tau_{Li})^{-\frac{\sigma}{1-\sigma}}$*,* $\Delta^{Lp}_{ij} = 0, \forall i \neq j$*;* $\Delta^{Lq}_{ii} = (1+\tau_{Li})^{-\frac{1}{1-\sigma}}$*,* $\Delta^{Lq}_{ij} = 0, \forall i \neq j$*;* $\hat{\lambda}_i = (\lambda_i/\eta)^{-\frac{\sigma}{1-\sigma}}$*;* $\hat{c}_i = \beta_i \hat{\lambda}_i^{-1}$*.*

As shown by equations (10) and (11), the aggregate TFP is determined by a complex interplay of the distortions (Δ^{Xq}, Δ^{Lq}, Δ^{Xp}, Δ^{Lp}) and the production linkages (α, β, V). In particular, the inverse matrices $[I - \tilde{\alpha}\hat{A}(\Delta^{Xq} \circ V)]^{-1}$ and $[I - \tilde{\alpha}\hat{A}(\Delta^{Xp} \circ V)]^{-1}$ play a key role in the interaction between the distortions and the production linkages. Our theoretical results, which we present in the next section, are based on these two equations.

3 Theoretical analysis

In this section, we analytically characterize the effects of distortions on aggregate productivity. We find that the effects of distortions and the role of sectoral linkages depend crucially on how substitutable inputs are. All the proofs and all the derivations of the expressions given in this section are in Appendix B.

3.1 Definitions

We start by defining several useful variables. These sectoral variables, $(x_{ij}, x_i, s_i, s_{ij}, \gamma_i)$, which are all functions of the production network parameters, (α, β, V), will help analytically characterize the effects of distortions. We also define the β-weighted variance and covariance, which will be useful for studying the role of input substitutability.

1. *Leontief inverse.* The Leontief inverse, which appears commonly in input-output analysis, captures the direct and indirect linkages between sectors. We denote $\Omega \equiv$

$(I-\tilde{\alpha}V)^{-1}$ with elements ω_{ij}, the Leontief inverse of the frictionless economy (with normalized productivity). The element ω_{ki} is equal to the percentage reduction in the total labor input required to produce good k following a percentage increase in productivity A_i, including the effect of sector i on the suppliers of sector k, and the suppliers of its suppliers, and so on.

2. *Connections between sectors and intermediate-input intensity.* We define the connection of sector j to sector i as x_{ij}, with $x_{ij} \equiv \omega_{ij}$ if $j \neq i$, and $x_{ii} \equiv \omega_{ii} - 1$. The sectoral connections satisfy the following recursive relationship: $x_{ij} = \alpha_i v_{ij} + \sum_r \alpha_i v_{ir} x_{rj}$. The variable x_{ij} governs the direct and indirect use of input j by i. We define the intermediate-input intensity of sector i as $x_i \equiv \sum_j x_{ij}$.

3. *Importance of a sector.* We define the importance of a sector as $s_i \equiv \sum_k \beta_k \omega_{ki}$.[12] This variable corresponds to the sales-to-GDP ratio of the undistorted economy (and hence sums to more than one).[13] The sector's importance can be decomposed in two terms as follows: $s_i = \beta_i + \gamma_i$. The first term, β_i, measures the *importance of the sector in final output*, and the second term, $\gamma_i \equiv \sum_k \beta_k x_{ki}$, captures the *importance of the sector as an input supplier.*[14]

4. *Importance of the connection between sectors.* We define the importance of the connection between sector i and j as $s_{ij} \equiv \sum_k \beta_k \omega_{ki} \omega_{kj}$, which can be rewritten as $s_{ii} = \beta_i(1 + 2x_{ii}) + \sum_k \beta_k x_{ki}^2$ and $s_{ij} = \beta_i x_{ij} + \beta_j x_{ji} + \sum_k \beta_k x_{ki} x_{kj}$ for $i \neq j$. This indicator measures the connection of both i to j and j to i, weighted by their final-output importance as well as how similarly concentrated the use of the two sectors' products is.

5. *β-weighted covariance and variance.* For arbitrary vectors p and q, we define $\mathrm{cov}_\beta(p,q) \equiv \sum_k \beta_k p_k q_k - \left(\sum_k \beta_k p_k\right)\left(\sum_k \beta_k q_k\right)$. The β-weighted variance is then $\mathrm{var}_\beta(p) = \mathrm{cov}_\beta(p,p)$. For arbitrary matrices P and Q, the β-weighted covariance of two column vectors is written $\mathrm{cov}_\beta(P_{.i}, Q_{.j}) \equiv \sum_k \beta_k P_{ki} Q_{kj} - \left(\sum_k \beta_k P_{ki}\right)\left(\sum_k \beta_k Q_{kj}\right)$.

[12] For some parameter combinations, there could be sectors with $s_i = 0$—that is, they are neither final-goods producers nor intermediate-goods suppliers (direct or indirect) to final-goods producers. We assume away such sectors (which do not affect the results) and hence assume $s_i > 0$ for all i.

[13] We show the connection with the sales-to-GDP ratio in Proposition A.3 in Appendix A. The sales-to-GDP ratio is sometimes referred to as the *Domar weight* (Hulten, 1978).

[14] The term γ_i is related but not equivalent to upstreamness. The importance of the sector as an input supplier is high if the sector is more upstream *and* if the sector is supplying inputs to sectors important for final output.

3.2 Analytical characterization

We provide a characterization of aggregate TFP for two types of distortions: distortions that affect only labor (or more generally, primary inputs) and distortions that symmetrically affect labor and all intermediate inputs.[15] We highlight the fundamental differences between these two types of distortions and their consequences for aggregate TFP. The derivations are obtained by approximating aggregate TFP to the second order around its non-distorted value, under the assumption that the two elasticities of substitution are equal (that is, $\sigma = \rho$). We use the expressions of Proposition 4, and we normalize productivities to one, following Proposition 2. As explained in section 2.2.1, aggregate TFP is equal to one in the normalized frictionless economy. The expressions of aggregate TFP given in this section can therefore be interpreted as the gap between the TFP and its frictionless level.

3.2.1 Distortions on labor

When distortions affect only labor, the source of misallocation is the dispersion in distortions. As shown by the following proposition, misallocation depends also on the interaction between distortions and sectoral linkages and on the degree of input substitutability.[16] The proposition generalizes the derivation of Osotimehin (2019) to a framework with sectoral linkages.

Proposition 5 *Suppose that $\sigma = \rho$, $\alpha_i \in (0,1)$ for all i, $A_i = B_i = 1$ and $\tau_{Xij} = 0, \forall i, j = 1, ..., n$. For small distortions, aggregate TFP is approximatively equal to*

$$\begin{aligned} \log TFP \;\; \approx \;\; & -\frac{1}{2}\left[\sum_i \sum_j s_{ij}(1-\alpha_i)(1-\alpha_j)\tau_{Li}\tau_{Lj} - \left(\sum_i s_i(1-\alpha_i)\tau_{Li}\right)^2\right] \\ & -\frac{1}{2}\frac{1}{1-\sigma}\left[\sum_i s_i(1-\alpha_i)\tau_{Li}^2 - \sum_i \sum_j s_{ij}(1-\alpha_i)(1-\alpha_j)\tau_{Li}\tau_{Lj}\right], \quad (12) \end{aligned}$$

where $\sum_i s_i(1-\alpha_i) = 1$ and $\sum_i \sum_j s_{ij}(1-\alpha_i)(1-\alpha_j) = 1$.

[15] We have derived a similar approximation for distortions on only intermediate goods. We do not include this case in the paper because it does not yield additional insights.

[16] The approximation is derived by standard second-order Taylor approximation. Formally, $\log TFP(\tau_L) = f(\tau_L) + \sum_{|\upsilon|=2} g_\upsilon(\tau_L)\tau_L^\upsilon$, where we use multi-index notation on υ, τ_L is the vector of labor distortions, $f(\tau_L)$ is the right-hand side in the proposition and the functions $g_\upsilon(\tau) \to 0$ as $\tau \to \mathbf{0}$. The same applies to all second-order approximations we present.

To see the fundamental role of the dispersion, consider the case when distortions are identical across firms, $\tau_{Li} = \bar{\tau}_L$. In that case, Proposition 1 implies that aggregate productivity would be equal to its frictionless level—that is, $\log TFP = 0$ (recall that the sectoral productivities are normalized to one).[17] Cross-sectional dispersion, which is the main focus of the misallocation literature, remains the sole source of misallocation for primary-input distortions, even when production linkages are accounted for.

However, the channels through which a given dispersion of distortions translates into a TFP loss depends on both the sectoral linkages, through the importance of the sector and that of the connection between sectors, s_i and s_{ij}, and the value of the elasticity of substitution. The first line of equation (12) gives the TFP loss when inputs are perfect complements, $\varepsilon = 1/(1-\sigma) = 0$, and the second line gives the additional effect when $\varepsilon > 0$. When inputs are perfect complements, the distortions reduce aggregate productivity only through their effect on final consumption. In that case, both labor and intermediate inputs are chosen in fixed proportion, and the production efficiency of each sector is therefore unaffected by the distortions and their consequences on relative prices. When inputs are not perfect complements, intermediate inputs choices are affected by both the distortions and their consequence on relative prices. In that case, labor distortions also reduce the production efficiency of each sector.

3.2.2 Distortions on all inputs

As shown in the following proposition, when distortions affect all inputs, additional sources of misallocation are present, and the TFP loss does not depend only on the dispersion of distortions.[18]

Proposition 6 *Suppose that* $\sigma = \rho$, $\alpha_i \in (0,1)$ *for all* i, $A_i = B_i = 1$ *and* $\tau_{Xij} = \tau_{Li} = \tau_i, \forall i, j = 1, ..., n$.

[17] This result holds in general and does not hinge on the approximation used in Proposition 5. See Appendix A for the derivation of the result.

[18] Baqaee and Farhi (2020) present similar derivations. In Appendix A, we explain how our expression differs from theirs.

For small distortions, aggregate TFP is approximatively equal to

$$\log TFP \approx -\frac{1}{2}\left[\sum_i\sum_j s_{ij}\tau_i\tau_j - \left(\sum_i s_i\tau_i\right)^2\right] \\ -\frac{1}{2}\frac{1}{1-\sigma}\left[\sum_i s_i\tau_i^2 - \sum_i\sum_j s_{ij}\tau_i\tau_j + 2\sum_i s_i x_i \tau_i \bar{\tau}^i_{suppliers}\right], \quad (13)$$

where $\bar{\tau}^i_{suppliers} = \sum_j (x_{ij}/x_i)\tau_j$.

Here as well, the TFP loss is shaped by the sectoral linkages, and the linkages come into play through the importance of the sectors, s_i, and that of the connection between sectors, s_{ij}. The crucial difference is that now the TFP loss is affected also by the interaction between the sectors' distortions and the average distortion of their suppliers, $\bar{\tau}^i_{suppliers}$, and by the interaction between the sectors' importance s_i and their intermediate-input intensity x_i (which includes direct and indirect linkages). The TFP loss will be larger if highly distorted sectors tend to purchase intermediate inputs from each other, and all the more so if the distorted sectors are important sectors.

Moreover, the dispersion in distortions is no longer the only source of misallocation. Let us consider again the case where distortions are identical across sectors—that is, $\tau_i = \overline{\tau}$ for all sectors. In that case,

$$\log TFP \approx -\frac{1}{2}\overline{\tau}^2\left[\sum_i\sum_j s_{ij} - \left(\sum_i s_i\right)^2 + \frac{1}{1-\sigma}\left(\sum_i s_i - \sum_i\sum_j s_{ij} + 2\sum_i s_i x_i\right)\right].$$

Distortions that are identical across firms also reduce aggregate TFP. First, since sectors differ in their intermediate-good intensity x_i, prices can be dispersed even when distortions are symmetric, which leads to inefficiencies in the production of the final good and of the different sectoral goods. Second, the total intermediate-goods usage is reduced because the user cost of intermediate goods (relative to labor) is higher, which leads to further inefficiencies in the production of the final good. This result, also highlighted in Jones (2011), contrasts with the labor-distortions case for which only the dispersion of distortions matters.[19] The key difference is that all-input distortions affect the total quantity of intermediate inputs used, whereas labor distortions do not affect the total

[19]This result is related to Diamond and Mirrlees's (1971) no-taxation-on-intermediate-inputs result.

quantity of labor used (since aggregate TFP is, by definition, computed for a given quantity of labor).[20]

3.3 How does input substitutability shape the effects of distortions?

As the previous section shows, input substitutability is a central determinant of aggregate productivity. In this section, we use Propositions 5 and 6 to explore the role of input substitutability in more detail. In our three main results, we show how the degree of input substitutability modifies the aggregate TFP loss, the amplification from sectoral linkages, and the impact of each sector.

3.3.1 Aggregate TFP loss

Our first result is that the TFP loss is larger when inputs are more substitutable (and smaller when inputs are more complementary). This result, which we state more formally below, by determining the sign of $d \log TFP/d\sigma$, holds for both labor and all-input distortions.[21]

Result 1 *Suppose that $\sigma = \rho$, $A_i = B_i = 1$ $\forall i = 1, ..., n$ and $\tau_i \neq 0$ for some i. For small distortions (on either primary or all inputs),*

$$\frac{d \log TFP}{d\sigma} \leq 0.$$

The inequality is strict if there are sectoral linkages ($\alpha_i > 0$ for some i) and either the distortions are random or the distortions are symmetric all-input distortions.

At first, the result that distortions have less impact when firms cannot easily substitute between inputs could seem counterintuitive. Why isn't the outcome worse when firms are stuck with distorted inputs? The intuition behind the result is that the firms' input decision is less distorted when inputs are less substitutable, precisely because it

[20] Hence, symmetric labor distortions have no impact on aggregate productivity because their only potential effect is on the total quantity of labor used.

[21] Recall here that the production function is normalized at the frictionless allocation, which implies $TFP(0, \sigma) = 1$ for any σ. Hence, computing $d \log TFP/d\sigma$ is equivalent to computing $d \log \left(TFP(\tau, \sigma)/TFP(0, \sigma)\right)/d\sigma$.

is more difficult for firms to substitute away from the distorted input. When the elasticity of substitution is low, firms respond less to the distorted prices, hence the firms' input decision does not deviate much from the first-best allocation. At the limit, with perfectly complementary inputs, the input decision is unaffected by relative price distortions.[22] Smaller changes in the allocation lead to smaller changes in the quantity of labor directly and indirectly used for the production of each sectoral good and hence to a smaller decline in aggregate productivity. This is not, however, the only mechanism at play. A lower elasticity of substitution tends to amplify the consequences that an inefficient mix of inputs can bring about on production, thus amplifying the TFP loss. We show (for small distortions) that the first effect dominates, and a higher degree of complementary leads to a smaller TFP loss.

A similar effect of the role of the elasticity has been found in simpler settings.[23] Note however that the presence of sectoral linkages makes the role of the elasticity of substitution more complex. In fact, the elasticity of substitution has in general a non monotonic effect on the TFP loss. This non-monotonicity comes in particular from the ambiguous effect of intermediate inputs on final output: a lower intermediate input share leads to a lower gross output, but also means higher consumption for a given gross output. As the share of output devoted to intermediates declines further (and that of consumption increases further) when the elasticity of substitution is higher, the TFP loss may become smaller. Our result shows that this effect does not appear when distortions are small. As we will show in Section 4, the ambiguous effect is virtually absent from the quantitative results as well: The TFP loss increases with the elasticity of substitution for empirically relevant values of distortions and elasticities of substitution.

The way input complementarity modifies the consequences of distortions contrasts sharply with how it modifies the consequences of sectoral productivity shocks. Whereas input complementarity dampens the consequences of distortions, it amplifies the consequences of sectoral productivity shocks. We now state this result more formally.

[22]The absence of distortionary effects under perfect complementarity case has been exploited by Rotemberg and Woodford (1993) in their study of business cycles with imperfect competition. They show that there exists a value-added production function that is independent of the value of the markups only if primary inputs and intermediates are perfect complements.

[23]In their quantitative analysis, Hsieh and Klenow (2009) find that, in a framework without sectoral linkages, a higher elasticity of substitution across final goods amplifies the effects of distortions. The result is also reminiscent of Ramsey's (1927) result on the optimal taxation of commodities.

Result 2 *Suppose that $\sigma = \rho$, $\alpha_i \in (0,1)$ for all i, and set $A = \mathbf{1}$. Let $Y(B,\sigma)$ be the output of an economy without distortions with productivity vector B and parameter σ. Then, $Y(B,\sigma)$ is increasing in σ. Suppose there exist some k, j_1, j_2 such that $B_{j_1} \neq B_{j_2}$ and $\omega_{kj_1} > 0, \omega_{k,j_2} > 0$. Then, $Y(B,\sigma)$ is strictly increasing in σ.*

The intuition is most apparent when we consider the effect of a negative productivity shock in some sector j. When the elasticity of substitution is low, switching to other inputs would disrupt production; therefore, firms do not change much their use of sector j's products. This limited reallocation reinforces the TFP decline, since the economy must devote additional resources to a low-productivity sector. On the other hand, when the elasticity of substitution is high, switching to other inputs is not as disruptive. As a result, the economy does not need to devote as many resources to the low-productivity sector, which attenuates the TFP decline.

Results 1 and 2 highlight an interesting contrast between the effects of different inefficiencies. Frictions that distort the relative prices of goods will be less harmful in economies with stronger input complementarities, whereas frictions that reduce firm-level productivity will be less harmful in economies with stronger input substitutability.

3.3.2 Amplification from sectoral linkages

We now turn to the role of sectoral linkages and focus here on their potential amplification effect. Contrary to intuition, we find that sectoral linkages do not systematically amplify the effects of distortions. We present the conditions under which sectoral linkages dampen the effects of the two types of distortions (labor and all-input distortions). We find the dampening conditions are more likely to hold if the distortions affect only labor.

3.3.2.1 TFP in the absence of sectoral linkages

We study the possibility of a dampening effect by comparing TFP to its level in the absence of sectoral linkages (that is, $\alpha_i = 0$ for all sectors), which is given by

$$\log TFP|_{\alpha=0} \approx -\frac{1}{2}\,\mathrm{var}_\beta(\tau). \tag{14}$$

In the absence of sectoral linkages, TFP is simply the β-weighted variance of distortions. Sectoral linkages amplify the effects of distortions if $\log TFP < \log TFP|_{\alpha=0}$ and dampen them if $\log TFP > \log TFP|_{\alpha=0}$.

3.3.2.2 Distortions on labor

Sectoral linkages can dampen the effects of labor distortions when inputs are complements. A key condition under which the dampening effect occurs is when each sector's weight is the same in the economy with and without sectoral linkages.

Result 3 *Consider an economy with only labor distortions. The TFP in an otherwise identical economy without sectoral linkages,* $\log TFP|_{\alpha=0}$*, is given by equation* (14)*.*

1. *If* $v_{ij} = 0$ *for* $i \neq j$*, then* $\log TFP = \log TFP|_{\alpha=0}$.

2. *If there is a sufficiently large number of sectors, inputs are complements (*$\sigma < 0$*) and the distortions* τ_{Li} *are identically and independently distributed and independent from any other model parameters, then* $\log TFP \geq \log TFP|_{\alpha=0}$*. The inequality is strict if* $\alpha_i v_{ij} > 0$ *for some* $i \neq j$*.*

3. *If inputs are complements (*$\sigma < 0$*) and* $(1-\alpha_i)s_i = \beta_i$*, then* $\log TFP \geq \log TFP|_{\alpha=0}$.

Above we establish three separate conditions under which sectoral linkages do not amplify the effects of distortions. First, we consider a special but instructive case in which the economy consists of "island" sectors: firms purchase intermediate goods only from their own sector. In that economy, where all linkages are *intra*-sectoral, the efficiency of sectoral production is unaffected since the distortions equally affect all labor used (directly or indirectly) in producing each product. Therefore, the TFP loss of that economy is identical to that of the no-sectoral-linkages economy. This special case highlights the crucial role of *inter*-sectoral linkages in the amplification of labor distortions.

Next, we consider the case where distortions are random, in the sense that they are distributed independently from any other sectoral parameter.[24] With random distortions, the connections between sectors do not matter when the production function is Cobb-Douglas. In that case, aggregate TFP is identical to that of the economy without sectoral linkages. The sectoral linkages generate two counteracting effects. On the one hand, intermediate inputs are an additional channel through which distortions reduce aggregate productivity. On the other hand, the use of intermediate inputs reduces the effect of

[24]See Appendix A.6 for details of the construction.

labor distortions, as labor then represents a smaller share of the firms' inputs. When the production function is Cobb-Douglas, the two offsetting effects cancel out exactly. When inputs are complements ($\sigma < 0$), the second effect dominates, and sectoral linkages dampen the effects of labor distortions.

Finally, we consider the general case in which distortions can be correlated with the sectoral characteristics. The intuition for this case is very similar to the random distortions case, but a key condition here is that each sector's value added must be the same in the two economies (with and without sectoral linkages), $s_i(1-\alpha_i) = \beta_i$.[25] Note that the result generalizes beyond the case where $s_i(1-\alpha_i) = \beta_i$: If sectors with higher-than-average labor distortions have a value-added share smaller than their final consumption share, then the sectoral linkages will dampen the effect of the distortion. Furthemore, the case $s_i(1-\alpha_i) = \beta_i$ is interesting in itself because it corresponds to how models that do not explicitly account for intermediate inputs are generally calibrated (the sectors' weights are calibrated on their value added shares). Our finding implies that adjusting the calibration is not sufficient to take care of the sectoral-linkages effect. Models that do not explicitly account for intermediate inputs and sectoral linkages overestimate the TFP loss from labor distortions.

3.3.2.3 Distortions on all inputs

When intermediate inputs also are distorted, the presence of sectoral linkages does not reduce the weight of the distorted inputs in production (as they did in the case of labor distortions), and sectoral linkages are therefore less likely to dampen the effects of distortions. We find that sectoral linkages can still dampen the effect of all-input distortions, but the dampening effect occurs under stronger conditions than when only labor is distorted.

Result 4 *Consider an economy with sectoral linkages ($\alpha_i > 0$ for some i) and with all-input distortions. The TFP in an otherwise identical economy but without sectoral linkages, $\log TFP|_{\alpha=0}$, is given by equation (14).*

If $Cov_\beta(\tau_i, x_i\bar{\tau}^i_{suppliers}) < -(1/2)\,\mathrm{var}_\beta(x_i\bar{\tau}^i_{suppliers})$ and the vector of distortions is not

[25]Note that with Cobb-Douglas, aggregate productivity is $\log TFP \approx -\frac{1}{2}\left[\sum_i s_i(1-\alpha_i)\tau_{Li}^2 - \left(\sum_i s_i(1-\alpha_i)\tau_{Li}\right)^2,\right]$ with $\sum s_i(1-\alpha_i) = 1$, and the TFP is the same as in the no-sectoral-linkages economy when $s_i(1-\alpha_i) = \beta_i$.

too large, then for all σ sufficiently small, $\log TFP > \log TFP|_{\alpha=0}$.[26]

The result states that sectoral linkages dampen the effects of distortions if $x_i\bar{\tau}^i_{suppliers}$ is sufficiently negatively correlated to τ_i—that is, if distortions are high in sectors with a low intermediate-input intensity or in sectors purchasing intermediates from low-distortion sectors. To see the intuition behind the result, let us consider the case of perfect complementarity. In that case, aggregate productivity is given by

$$\log TFP \approx -\frac{1}{2}\,\mathrm{var}_\beta\left(\tau_i + x_i\bar{\tau}^i_{suppliers}\right). \tag{15}$$

With perfect complementarity, sectoral production is efficient, and the TFP loss depends only on the consequences of the price dispersion, caused by distortions, for final consumption. Sectoral linkages can therefore amplify or dampen the TFP loss, depending on how they affect the price dispersion, which depends in turn on the interaction between the distortions and the sectoral linkages. If all the firms purchase inputs from the same sectors and in the same proportion, then the TFP loss would be identical in the economies with and without sectoral linkages.[27] If firms in high-distortion sectors purchase inputs from low-distortion sectors, sectoral linkages will tend to reduce the price dispersion.

3.3.3 Impact of each sector

In this section, we further study the role of sectoral linkages by focusing on the role of each sector. Which sectors have a large aggregate impact when distorted? What type of linkages lead to a larger impact? Again, the answers to these questions depend on the value of the elasticity of substitution. In particular, we find that intermediate-input suppliers have a smaller impact if inputs are less substitutable. To study the determinants of the impact of each sector, we first consider the case in which only one sector is distorted and then the case in which two sectors are distorted.

3.3.3.1 When only one sector is distorted

We find that the aggregate impact of a distortion in only one sector depends primarily on the importance of the sector for final output and as an intermediate-input supplier.

[26]Note that for simplicity, we slightly changed the notation relative to the definition of Section 3.1. Here $\mathrm{var}_\beta(\tau_i + x_i\bar{\tau}^i_{suppliers})$ refers to the variance of the vector whose element is $\tau_i + x_i\bar{\tau}^i_{suppliers}$. We use a similar notation for the covariance.

[27]In that case, $x_i = x$ and $\bar{\tau}^i_{suppliers} = \bar{\tau}_{suppliers}$ $\forall i$; therefore, $\mathrm{var}_\beta(\tau_i + x_i\bar{\tau}^i_{suppliers}) = \mathrm{var}_\beta(\tau_i)$.

Result 5 *The effect of a distortion in only one sector is given by*

$$\frac{d\log TFP}{d\tau_{Li}} \approx -\left[(1-\alpha_i)^2(s_{ii}-s_i^2)+\frac{1}{1-\sigma}[(1-\alpha_i)s_i-(1-\alpha_i)^2 s_{ii}]\right]\tau_{Li},$$

when only labor is distorted; and by

$$\frac{d\log TFP}{d\tau_i} \approx -\left[s_{ii}-s_i^2+\frac{1}{1-\sigma}(s_i-s_{ii}+2s_i x_{ii})\right]\tau_i,$$

when all inputs are distorted.

We derive the result for both kinds of distortions but focus our discussion on the case where all inputs are distorted.[28] Note that even in this simple case, with only one distorted sector, the aggregate impact of the distortion cannot be summarized by a single statistics. Furthermore, Result 5 shows that the relevant sectoral linkages depend on the elasticity of substitution. With perfect input complementarity ($1/(1-\sigma)=0$), the relevant sectoral linkages is given by the first term of the expression, $s_{ii}-s_i^2$. The additional effect—given by the second term, $s_i-s_{ii}+2s_i x_{ii}$—matters only when inputs can be substituted. To give more intuition, we approximate the two components and obtain

$$\frac{d\log TFP}{d\tau_i} \approx -\left(\beta_i+\frac{1}{1-\sigma}\gamma_i\right)\tau_i.$$

The key characteristics that determine the sector's impact are therefore β_i, the importance of the sector for final output, and γ_i, the importance of the sector as an intermediate-input supplier. The distortion has a more detrimental impact if it affects a sector that is important for final output, whatever the value of the elasticity of substitution. By contrast, a distortion on an important intermediate-input supplier leads to a larger TFP loss only if inputs can be substituted; the impact of the intermediate-input suppliers increases with the elasticity of substitution.

To develop intuition further and better understand which sectoral linkages are behind γ_i, let us consider the case of a simple input-output economy in which all the firms use the same intermediate-input bundle (that is, $v_{ij}=v_j\ \forall i=1,...,n$). In this simple economy, $\gamma_i=(\bar{\alpha}_\beta v_i)/(1-\bar{\alpha}_v)$, where $\bar{\alpha}_\beta=\sum_i \beta_i\alpha_i$, and $\bar{\alpha}_v$ is similarly defined. Thus,

[28] The labor-distortions case includes similar terms, with two essential differences: the effect of the distortion is scaled down by $1-\alpha_i$ (the weight of labor in the production function); and the double-marginalization term $2s_i x_{ii}$ is absent.

the importance of a sector as an intermediate-input supplier increases with the direct intensity with which the sector's product is used in the intermediate-input bundle's production, v_i; and with the direct intensity with which the intermediate-input bundle is used, $(\bar{\alpha}_\beta, \bar{\alpha}_v)$.

3.3.3.2 Distortions on multiple sectors

How is the impact of a sectoral distortion modified when the distortion in another sector increases? To study this interaction effect and hence better explain how linkages between distorted sectors shape the TFP loss from distortions, we consider the case in which only two sectors are distorted. The analysis highlights the connection between the sectors as well as how similarly the two sectors' products are used by other sectors.

Result 6 *The effect of distortions in two sectors i and j, with $i \neq j$, is given by*

$$\frac{d^2 \log TFP}{d\tau_{Li} d\tau_{Lj}} \approx -\left[(s_{ij} - s_i s_j)\,(1-\alpha_i)(1-\alpha_j) - \frac{1}{1-\sigma} s_{ij} (1-\alpha_i)(1-\alpha_j) \right],$$

when only labor is distorted; and by

$$\frac{d^2 \log TFP}{d\tau_i d\tau_j} \approx -\left[s_{ij} - s_i s_j + \frac{1}{1-\sigma}\,(s_i x_{ij} + s_j x_{ji} - s_{ij}) \right],$$

when all inputs are distorted.

As for Result 5, we discuss here the case in which all inputs are distorted. The first component, $s_{ij} - s_i s_j$, which corresponds to the distortion's effect on the final consumption allocation, holding the efficiency of sectoral production fixed, can be rewritten as $s_{ij} - s_i s_j = -\beta_i \beta_j + \beta_i (x_{ij} - \gamma_j) + \beta_j (x_{ji} - \gamma_i) + \mathrm{cov}_\beta(x_{.i}, x_{.j})$. The first term of that component, $-\beta_i \beta_j$, which is the direct effect of the interaction on final consumption in the absence of sectoral linkages, reduces the TFP loss: a higher distortion on sector j reduces the impact of sector i's distortion because it makes more similar the prices of sectors i and j, which reduces the final-consumption misallocation. The remaining terms capture the role of sectoral linkages and indicate that sectoral linkages amplify distortions if the two sectors use each other's products more than the average sector and if the sectors using the products from the two distorted sectors are concentrated in the same subset of the economy.

The TFP loss from the inefficient mix of inputs is captured by the second term, which can be rewritten: $s_i x_{ij} + s_j x_{ji} - s_{ij} = [-\gamma_i \gamma_j + \gamma_i x_{ij} + \gamma_j x_{ji} - \text{cov}_\beta(x_{.i}, x_{.j})]$. Similar to the effect on consumption discussed above, the terms $-\gamma_i \gamma_j + \gamma_i x_{ij} + \gamma_j x_{ji}$ measure the interaction's effect on the variance of input prices. Intuitively, markups on multiple sectors make input prices more similar, which reduces misallocation. However, when two sectors are closely linked (high x_{ij} and/or x_{ji}) markups on markups (double marginalization) *increases* the variance of input prices and hence the degree of misallocation. The covariance term, $\text{cov}_\beta(x_{.i}, x_{.j})$ is large when sector i and j tend to supply inputs to the same sectors. A higher covariance hence means that a higher distortion in sector j falls mainly on purchasing sectors with already-high input prices, so the additional increase in the variance of input prices will be smaller (and the variance could even fall), which attenuates the impact of the distortions.

In the end, the result shows that the effects of the two distortions reinforce each other if the two distorted sectors are highly interconnected, whatever the value of the elasticity of substitution. The interaction effect increases with the importance of the two sectors as input suppliers but only if the interconnection's strength and the degree of input substitutability are sufficiently high. Finally, the covariance between the use of the two sectors' product (by other sectors) tends to reinforce the interaction effect, but to a lesser degree if inputs are more substitutable.

4 Quantitative analysis: the TFP loss from market power

In this section, we quantify the role of input substitutability in the context of the distortions caused by market power. We compute the TFP gains from eliminating industry-level markups in 35 countries and show how the gains vary with the degree of input substitutability. This section complements our theoretical results (derived for small distortions) by showing that they hold for the sizable distortions found in the data and by presenting their quantitative implications.

4.1 Markups across industries and countries

We measure the degree of market power in each industry and each country by computing price-cost margins, which, as we will see in Section 4.2, map one-to-one with the firms' markups.[29] Given our focus on input-output linkages, our analysis requires data covering all sectors of the economy. We use industry-level data from the World Input-Output Database's (WIOD) 2013 Socio-Economic Accounts.[30] The data set gives the sales, labor compensation, intermediate input costs, and real capital stock by industry (at roughly two-digit ISIC level) for 40 countries over the period 1995–2011.[31] The comprehensive sectoral coverage, the availability of the capital-stock measure at the industry level, the large selection of countries, and the inclusion of several middle-income countries are key advantages of this dataset. In Appendix C.2, we provide more details on the variables as well as the list of countries and sectors.

For each country-industry pair, we measure the price-cost margin, pcm, as sales minus total costs over sales:

$$\mathsf{pcm} = (\mathsf{sales} - \mathsf{labor\ cost} - \mathsf{intermediate\ input\ cost} - \mathsf{capital\ cost})/\mathsf{sales}.$$

Following the approach of Jorgenson (1967), we construct the capital cost as follows:[32]

$$\mathsf{capital\ cost} = (r + \delta)p_I K,$$

where r is the real interest rate, δ the sector-specific depreciation rate, p_I is the price index for investment goods, and K is the real capital stock. Both p_I and K are obtained from the WIOD. We set $r = 0.04$, and δ is computed for each sector using US data from the BEA Fixed Assets Accounts of 2000.[33]

[29] This measure has been recently used by Karabarbounis and Neiman (2018) and Barkai (2020) to document the evolution of markups in the US.

[30] The data are publicly available at: http://www.wiod.org/home. We use the 2013 release instead of the more recent release of the WIOD (2016) because the latter does not contain capital-stock data.

[31] We drop Italy from the sample because the data suggest there could be some classification issues: in Italy, "real estate" has an unusually low capital–output ratio and "renting of material and equipments" has unusually high capital–output ratio (factor of 0.09 and 21 relative to the industry median).

[32] The general expression proposed by Jorgenson also depends on taxes and on the inflation of investment goods, from which we have abstracted. Since in the SEA p_I and K are both benchmarked to the same year, $p_I K$ is the current year replacement cost of the capital stock, which is the relevant opportunity cost when computing the capital cost.

[33] We set the interest rate to 4% in line with the average Bank of America BBB US Corporate Index

To reduce the influence of outliers, for each country-industry pair, we take the median price-cost margin over the period 1995–2007; pcm hence refers to the median.[34] We use the median price-cost margin in the quantitative exercise of the next section as well.

In Appendix C.3, we report the distribution of the pcm across sectors and countries. The pcm is highly dispersed across both countries and sectors, with the lowest 25th percentile at -40% and the highest 75th percentile at 25%. The price-cost margin is negative in several country-sector pairs. We attribute these negative values to measurement errors and set the lower bound of the pcms to zero in the remaining of our analysis.[35] [36] We also find that price-cost margins tend to be higher in lower-income than in high-income countries and higher in North America than in Europe. Finally, we validate our measure by showing that price-cost margins tend to be higher in countries in which regulations are the least competition friendly.

The pcm implies markups values on the lower range of the values that have been found using other methods. For example, the recent paper by de Loecker et al. (forthcoming) finds the average markup among US publicly listed firms to be 40–50% over 1995-2007, whereas we find an average markup of 16% in the US.[37] The key difference between the two measures, beyond the data coverage, is that our measure gives the ratio of price over *average* costs, whereas de Loecker et al. (forthcoming) aim to capture the ratio of the price over *marginal* costs. Although identical in our theoretical framework, the two measures can in reality be quite different. Because our measure underestimates the

Effective Yield, which is equal to 6.4% over 1997-2011, and an expected inflation rate of 2%. We acknowledge that the real interest rate may vary across countries. However, gathering comparable real interest rate data for our sample of countries (or even a sub-sample) is a daunting task that we believe goes beyond the scope of this paper. Still, one may wonder whether we underestimated the real interest rate in the countries in which pcms are higher. We find that, in these countries, the pcms are not particularly high in capital-intensive sectors which suggests that the high pcms are not due to the underestimation of the real interest rate in these countries. On the contrary, pcms tends to be negative in capital-intensive sectors which suggest that we may have over-estimated the real interest rates. See Figure C.8 in Appendix C.5.

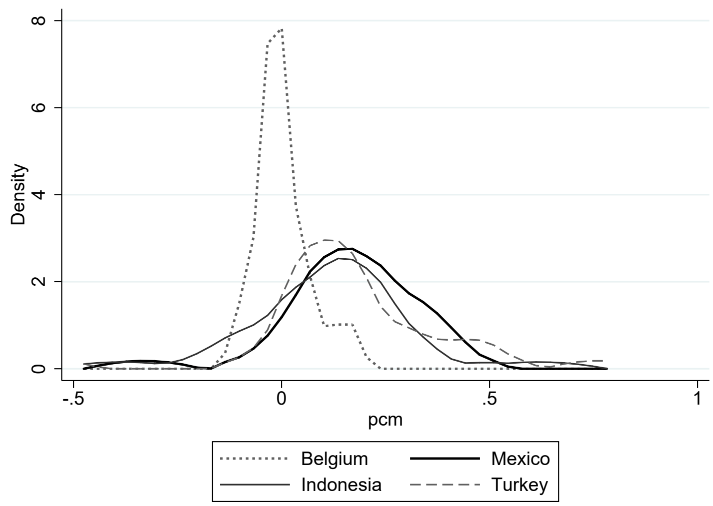
[34]We compute the median over the period 1995–2007 because the real capital stock variable is not at all available in 2010 and 2011 and not available for many countries in 2008 and 2009.

[35]In a recent work, Hall (2018) also attributes the negative markups he finds in some US sectors (using a different estimation method) to measurement errors. He proposes a method to uncover the true distribution of the markups. However, estimating the parameters of the true distribution of markups would not be sufficient for our purpose since the productivity loss does not depend solely on the distribution of the markups but on the joint distribution between markups and sectors' production parameters.

[36]We believe the capital stock is the most likely variable to be incorrectly measured. We assume in our calibration exercise that the intermediate input and labor income share are correctly measured.

[37]We use labor costs as weights and compute our average markup over all the sectors included in our sample. See Section 4.2 for more details on the connection between the markup and the pcm.

Figure 1: The distribution of markups across sectors in Belgium and in a selection of high-pcm countries



Note: The graph shows the density of pcm across industries in Belgium, Mexico, Indonesia and Turkey.

markup over marginal costs, our analysis will underestimate the productivity cost of markups (since the latter depends on the markup over marginal costs).

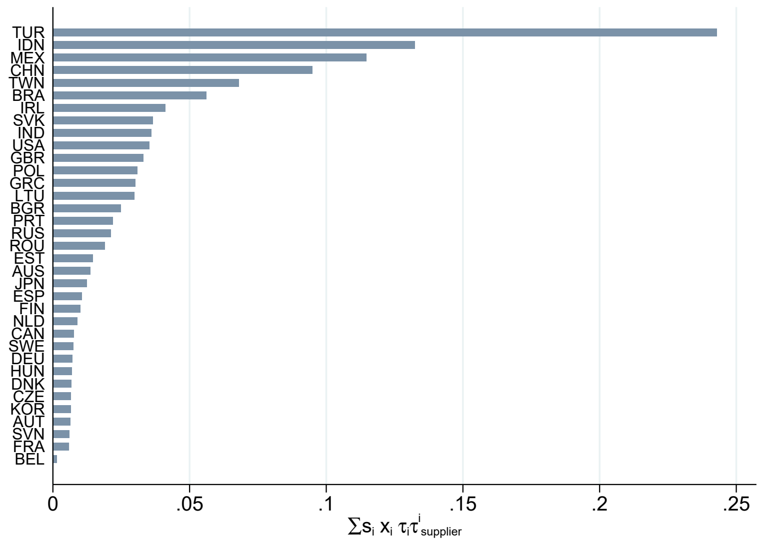
In Figure 1, we report the distribution of pcm in the three countries with the highest median pcm.[38] As shown in the figure, Mexico, Turkey, and Indonesia have not only a higher median pcm but also more dispersion across sectors than countries with lower pcms, such as Belgium. Moreover, Figure 2 shows that the interaction between the sectors' markups and their suppliers' is higher in these three countries. The aggregate-productivity implications of these patterns are investigated in the next section.

4.2 Bringing the model to the data

We now use the model to quantify the productivity losses caused by market power and the role of input substitutability. In Appendix C.1, we describe an economic environment where firms charge a markup on their marginal costs and show that the environment is isomorphic to the distorted planner's problem of Section 2 with distortions common to

[38]In this figure, we did not set the minimum pcm to zero like we do in the quantitative analysis.

Figure 2: Interaction between the sectors' markups and their suppliers'



Note: The graph shows the strength of the interaction between the sectors' markups and their suppliers' for all the countries of the sample.

all inputs. Here, we present the calibration strategy.

The model is calibrated on industry-level data from 35 countries, separately for each country.[39] The data used in the calibration also are obtained from the World Input-Output Database (WIOD). For each country-industry pair, we use median values (over 1995–2007) as targets for the calibration. In Appendix C.2, we provide more details on the data sources and the construction of the variables. In this section, we use price notations, $p_i = \lambda_i$ and $w = \eta$, instead of the multipliers, to simplify the exposition.

Markups. We derive markups from the price-cost margin measure, pcm, which is described in Section 4.1. In Appendix C.1, we show that the markup, μ_i, maps one for one into the price-cost margin:

$$\mu_i = \frac{1}{1 - \mathsf{pcm}_i}.$$

Implicit here is the assumption that other distortions average out at the industry level.

[39]Data for 40 countries are available in the WIOD. We excluded four small countries (Latvia, Malta, Cyprus, and Luxembourg) because too many zeros in the input-output table prevent us from solving the model at the same level of disaggregation as for the other countries. We also excluded Italy because the data patterns make us suspect classification discrepancies with respect to other countries.

Final demand. The final expenditure shares can be used to pin down the parameter β_i. From equation (6), we obtain

$$\beta_i = p_i C_i / (\sum_j p_j C_j),$$

where $p_i C_i$ is the final output of sector i.

Technology. With markups, the parameters of the production function $\{\alpha_i, v_{ij}\}_{i,j=1\ldots n}$ cannot be estimated directly from intermediate-input cost shares because the markups create a wedge between the cost shares and the technology parameters. However, given the values of the industry-level markups and those of the elasticities of substitution, we can derive from the first-order conditions of the firms an estimate for these parameters.[40] From Proposition 1 and $p_i = \mu_i mc_i$, we have

$$n_{ij} \equiv \alpha_i v_{ij} = \frac{p_j X_{ij}}{p_i Q_i / \mu_i} \left(\sum_j v_{ij} [p_i/p_j]^{\frac{\rho}{1-\rho}} \right)^{-\frac{\sigma-\rho}{\rho(1-\sigma)}} \left(\frac{p_i}{p_j} \right)^{-\frac{\rho}{1-\rho}} \mu_i^{\frac{\sigma}{1-\sigma}},$$

where $p_j X_{ij} / (p_i Q_i / \mu_i)$ is the cost share of intermediate good j for sector i.[41] The cost share reflects not only the parameters α_i, v_{ij} but also the entire price vector p, which is itself a function of the parameters v_{ij} and α_i. The calibration of the parameters α_i, v_{ij} hence involves finding a fixed point.[42] A simple iterative procedure quickly finds the fixed point and is robust to different initial guesses. With a solution for $\{n_{ij}\}$, we obtain $\alpha_i = \sum_j n_{ij}$ and then $v_{ij} = n_{ij}/\alpha_i$. Statistics on these parameters are reported in Appendix C.6.

The remaining two parameters—ρ and σ, which govern the elasticity of substitution between intermediate and primary inputs and the one between intermediate inputs—ε_σ and ε_ρ, are not straightforward to calibrate. A few empirical papers have estimated these elasticities of substitution. The elasticity between primary inputs and the intermediate-

[40] In Appendix C.4 we show how to adapt Klump et al. (2012) procedure to our environment and hence calibrate the model directly, without the iterative procedure. This procedure yields equivalent results. With that approach, however, the distribution parameters α_i, v_{ij} and hence s_i are harder to interpret (because the productivity parameters are not equal to one).

[41] Proposition 2 allows us to choose the units in such a way to ensure $A_i = 1$, $B_i = 1$ for all i.

[42] In Figure C.5, we plot the estimated α against the "naive" estimate $\sum_j p_j X_{ij} / p_i Q_i$. As we might expect, for most observations the estimated parameters are higher than the observed shares because of the presence of markups. However the markups cause the price of intermediates in some sectors to increase so much that the observed intermediate share is higher than the underlying technical parameter.

input bundle has been estimated to be between 0.4 and 0.9. Rotemberg and Woodford (1996) find a value of 0.7, Oberfield and Raval (2014) find a range between 0.6 and 0.9, and Atalay (2017) finds estimates ranging between 0.4 and 0.8.[43] Atalay (2017) estimates also the elasticity of substitution between intermediates and finds an estimate close to zero. Based on these estimates, we set our benchmark values of the elasticities at $(\varepsilon_\sigma, \varepsilon_\rho) = (0.7, 0.01)$. Recent estimates by Peter and Ruane (2020) on Indian data point to higher values for the long-run elasticity of substitution between intermediates but smaller values for that between primary factors and intermediate inputs. We present the results using elasticities values in line with their estimates in Appendix C.6.[44]

Note that since the calibration of the technology parameters $\{\alpha_i, v_{ij}\}_{i,j=1\ldots n}$ relies on the value of the elasticities of substitution, these parameters need to be recalibrated whenever the value of the elasticities of substitution is modified.

4.3 Quantitative results

We compute the aggregate productivity gains from removing industry-level markups in each of the 35 countries. In addition to quantifying the cost of industry-level markups, the quantitative analysis allows us to show that the theoretical results of Section 3 hold even when the distortions are sizable.

To quantify the TFP gains from removing markups, we use the general solution described in Proposition 1 and compute each country's TFP with and without markups. The numerical solution of the model can easily be computed by iterating on the Lagrange multipliers. We compute the TFP gain from removing markups for the benchmark values of the elasticities of substitution $(\varepsilon_\sigma, \varepsilon_\rho) = (0.7, 0.01)$, as well as for other combinations of the elasticities' values within the set $(0.01, 0.70, 1.00)$. In line with our normalization approach, we recalibrate the technology parameters $\{\alpha_i, v_{ij}\}_{i,j=1\ldots n}$ whenever the

[43] In a recent paper, Miranda-Pinto and Young (2021) find using US data that this elasticity (between primary inputs and the intermediate-input bundle) varies across sectors, with higher elasticities in downstream than in upstream sectors; The elasticities are equal to or above the Cobb-Douglas case in many sectors. They find that accounting for heterogeneous elasticities does not modify much the TFP loss from financial distortions but increases the amplification relative to an economy with only *intra*-sectoral linkages.

[44] They find a lower value for the elasticity of substitution (EOS) between value added and the intermediate bundle (0.6) but higher values for the EOS between materials, energy and services (0.4) and for the EOS between various types of materials (4.7). We find that this calibration leads to higher or lower TFP costs relative to the Cobb-Douglas specification depending on the country. In the countries with high TFP costs, the result is similar to that obtained with Cobb-Douglas specification.

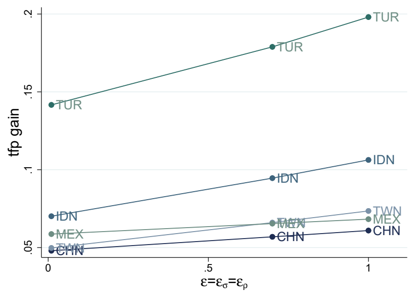
elasticities are let to vary.

For ease of exposition, we present here the results for the median country, as well as for the countries with the largest TFP gains. The complete tables of results, for all 35 countries and for additional values of the elasticities, are reported in Appendix C.6. We view this quantification as a lower bound of the productivity losses caused by markups. Our results capture the consequences of the misallocation only *across* sectors because the nature of the data does not permit measuring *within*-sector misallocation. Moreover, our markup measure, which is based on the firms' average rather than marginal costs, is most likely an underestimate, and our results therefore underestimate the productivity loss.

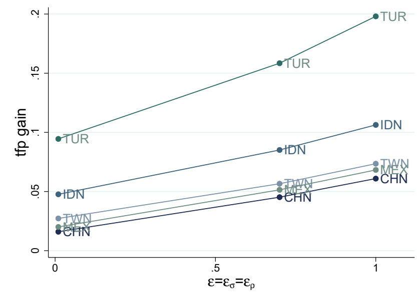
The left panel of Table 1 reports the TFP gains for the benchmark value of the elasticities $(\varepsilon_\sigma, \varepsilon_\rho) = (0.70, 0.01)$—as well as for the two polar cases of (almost) perfect complementarity, $(\varepsilon_\sigma, \varepsilon_\rho) = (0.01, 0.01)$, and Cobb-Douglas, $(\varepsilon_\sigma, \varepsilon_\rho) = (1.00, 1.00)$—for the median country and for the 10 countries with the largest gain. In these 10 countries the gain ranges from 0.6% to 20% depending on the country and on the value of the elasticities of substitution. The quantification of the cost of markups hinges crucially on the value of the elasticities of substitution. With the benchmark value of the elasticities, the TFP gain is equal to 1.2% in the median country. A unit elasticity of substitution between all inputs multiplies the TFP gain for the median country by 1.8 relative to the benchmark calibration, and by 3.5 relative to an elasticity of 0.01 across all inputs.

In Figure 3, we zoom in on the 5 countries with the largest TFP gains. As shown in the first three panels of the figure, the TFP gain typically increases with the two input elasticities (this is also the case for the 30 other countries not shown in the figure) for plausible values of the elasticities. The result that a higher elasticity leads to a larger cost of distortions, derived in Proposition 1, therefore holds empirically as well. As shown in panel (d), the non-monotonocity appears for values of the two elasticities that are well above the empirically-relevant range. The figure also complements the theoretical analysis by showing that the TFP gain increases separately with each elasticity. Furthermore, with the quantitative approach, we can now compare the role of the two elasticities. In the countries of Figure 3, the TFP gain appears more sensitive to the value of the elasticity of substitution between labor and the intermediate-input bundle than to that of the elasticity between intermediates; the same pattern is observed for most (but not all) the other countries (see Appendix C.6).

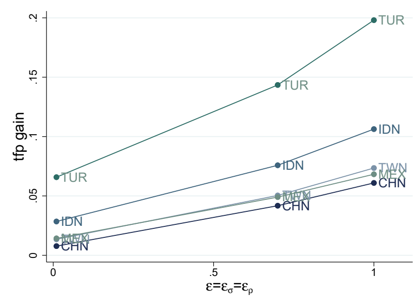
Figure 3: The TFP gain and input substitutability



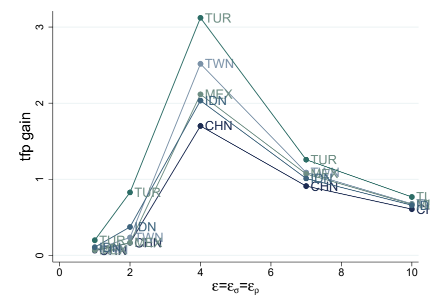
(a) varying the intermediate input elasticity



(b) varying the labor-intermediate input bundle elasticity



(c) varying the (identical) input elasticity



(d) high elasticities

Note: The graphs report the TFP gain from removing markups as a function of the elasticities of substitution ϵ_ρ and ϵ_σ for the 5 countries with the largest gains. In panel (a), we set $\varepsilon_\sigma = 1$ and vary ε_ρ; in panel (b), we set $\varepsilon_\rho = 1$ and vary ε_σ; in panel (c) and (d), we assume $\epsilon_\sigma = \epsilon_\rho$; panel (d) shows the results for higher-than-Cobb-Douglas elasticity values.

The value of the elasticities of substitution also determines the strength of the sectoral linkages' amplification effect. In fact, the larger TFP gain obtained when the elasticities of substitution are higher than the benchmark comes from a stronger amplification effect. The right panel of Table 1 reports the value of the amplification factor—that is, the ratio of the TFP gain with and without sectoral linkages. For these countries, the amplification factor can be as high as 16.2; if China is excluded, it can be as high as 6.4. A higher value of either elasticity of substitution leads typically to a stronger amplification effect. This result holds also for the other countries whose amplification factors are reported in Appendix C.6. The amplification factor of the median country is equal to 2.9 for the benchmark calibration but would be equal to 5.4 if the elasticity of substitution between

Table 1: TFP gain and IO amplification factor

	TFP gain			IO amplification factor		
$(\varepsilon_\sigma, \varepsilon_\rho)$	(0.01,0.01)	(0.70,0.01)	(1.00,1.00)	(0.01,0.01)	(0.70,0.01)	(1.00,1.00)
median	0.006	0.012	0.021	1.3	2.9	5.4
POL	0.006	0.016	0.030	1.1	3.0	5.8
GBR	0.007	0.017	0.032	1.4	3.4	6.4
IND	0.007	0.019	0.036	0.7	1.7	3.2
IRL	0.018	0.026	0.036	1.3	1.8	2.6
BRA	0.010	0.024	0.040	1.4	3.4	5.5
CHN	0.008	0.034	0.061	2.1	9.1	16.2
MEX	0.014	0.043	0.068	1.1	3.5	5.5
TWN	0.013	0.037	0.074	1.1	3.0	6.1
IDN	0.028	0.056	0.106	1.5	2.9	5.6
TUR	0.066	0.114	0.198	1.1	1.9	3.3

Note: This table presents the TFP gain and the Input-Output amplification factor for the median country and for the 10 countries with the largest TFP gain. Results are shown for the benchmark calibration, $(\varepsilon_\sigma, \varepsilon_\rho) = (0.70, 0.01)$ and for two other combinations of the elasticities values. The TFP gain is computed as (TFP without markups)/(TFP with markups) -1. The IO amplification factor is the ratio of the TFP gain in the baseline economy over the TFP gain in the economy with $\alpha_i = 0$, $\forall i = 1, ..., n$. See Appendix C.6 for the full table of results.

all inputs was equal to one. We find also that, in line with Section 3's theoretical results, sectoral linkages do not always amplify the effects of markups on aggregate productivity. This is the case for example in India, where the amplification factor is below one for highly complementary inputs. When inputs are highly complementary, there is no room for inefficiency in the firms' inputs decisions, and the only source of misallocation comes from the effect of the (inefficient) price dispersion on final consumption. The dampening effect then comes from the negative correlation between the sector's markups and either the sector's intermediate-input intensity or the average markup of the sector's suppliers. Our results indicate that the dampening effect is rare in the context of industry-level markups.

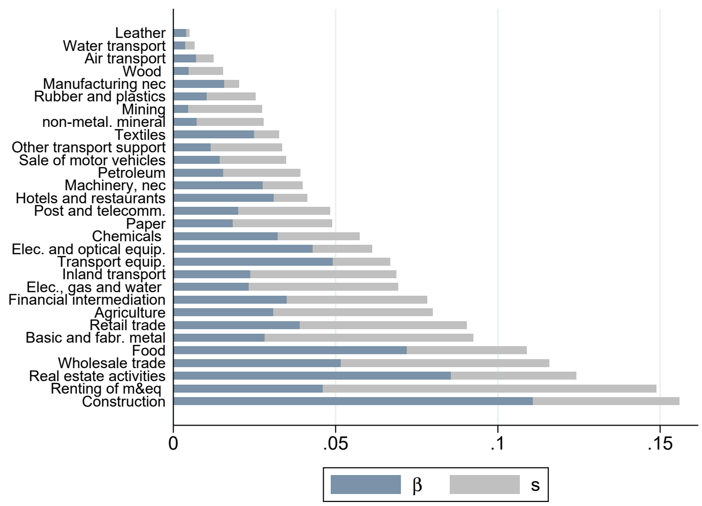
We now turn to the impact of each sector. The degree of input substitutability determines which network statistics is the most relevant to measure the sectors' impact. We report in Figure 4 the median value (across countries) of each sector's overall importance, $s_i = \beta_i + \gamma_i$, together with its importance for final output, β_i. The importance of the sector as an intermediate-input supplier, γ_i, appears on the graph as the difference between s_i and β_i. As explained in Section 3.3.3, the sector's impact is close to s_i when the two input elasticities are equal to one, and it is close to β_i when inputs are perfect complements. The comparison between these two polar cases nicely illustrates how central the value of the elasticity of substitution is. The figure shows that assuming a unit elasticity of substitution leads to overestimating the impact of all the sectors, particularly the sectors that supply primarily intermediate inputs, such as the basic metal, the material-and-equipment rentals, the wood products, the non-metallic mineral, and the mining industries. In half of the sectors, the impact is overestimated by a factor of 2.2 or higher relative to the perfect complement case. The overestimation is not uniform across sectors and hence modifies the sectors' ranking. The basic and fabricated metals', electricity, gas and water, and inland transportation's ranks are the most overestimated, whereas transport equipment's and electrical equipment's ranks are the most underestimated. On the other hand, sectors that are important final-output suppliers—such as construction, real estate, and the food and beverage industry—rank high whatever the value of the elasticities of substitution. For our benchmark value of the elasticities of substitution, the impact of each sector is likely to be in between the two polar cases shown in Figure 4. We believe that assuming a unit elasticity of substitution leads to overestimating the impact of intermediate-input suppliers, albeit to a smaller extent than is suggested by the comparison with the perfect complementary case discussed above.

4.4 Additional results and robustness checks

We undertake additional exercises and counterfactuals, and we verify that the results are robust to outliers. Detailed results are reported in Appendix C.6. Here, we give a summary of the main results.

We first consider the TFP gain that would obtain if the distortions affected only labor (and not all inputs). As highlighted in Section 3, distortions' effects vary depending on whether they affect both labor and intermediates or only labor. In line with the theoretical results, we find that the TFP loss is substantially smaller when the distortions affect only labor, and the sectoral linkages are more likely to dampen the effect of distortions. In fact, sectoral linkages dampen the effect of the distortions in most countries. As shown in Table 2, the median TFP loss is equal to 0.3% and 30 percents smaller relative to an otherwise identical economy without sectoral linkages in the baseline calibration (i.e

Figure 4: The impact of each sector



Note: This graph represents the median value (across countries) of the sectors' overall importance, s_i, and that their importance in final output, β_i. The underlying technology parameters are estimated under the benchmark elasticities values, $(\varepsilon_\sigma, \varepsilon_\rho) = (0.7, 0.01)$.

amplification factor of 0.7). If we further assume that the sectors' sizes are the same in the two economies (that is, $\beta_i|_{\alpha=0} = (1 - \alpha_i)s_i$), the median TFP loss would be even smaller, about 40% smaller than in the economy without sectoral linkages (see Appendix C.6).

We then check the robustness to outliers in the values of the markup. In a first robustness check, we compute the gap between the price-cost margin and the industry's median, and we winsorize the price-cost margin using the 1st and 99th percentile of the gap (computed on pooled data). In the second robustness check, we compute the counterfactual TFP obtained when setting markups to zero in all sectors except real estate.[45] Removing outliers does not affect our main result concerning the role of the elasticity. We find that the TFP gain and the amplification factor typically increase with the two elasticities of substitution.

[45] Given the challenges associated with measuring the capital stock and the large capital-output ratio in that sector, the likelihood of mismeasuring the markup is higher in that sector. As shown in Section 4.1, the real-estate sector displays a tremendous dispersion in pcm across countries; for most countries, the price-cost margin in real estate is negative, but in the few countries in which the price-cost margin is positive, it has very high values.

Table 2: TFP gain and IO amplification factor - markups on labor only

$(\varepsilon_\sigma, \varepsilon_\rho)$	TFP gain (0.01,0.01)	TFP gain (0.70,0.01)	TFP gain (1.00,1.00)	IO amplification factor (0.01,0.01)	IO amplification factor (0.70,0.01)	IO amplification factor (1.00,1.00)
median	0.002	0.003	0.005	0.5	0.7	1.2
BGR	0.003	0.004	0.009	0.6	0.8	1.9
GRC	0.005	0.006	0.010	0.5	0.7	1.2
DNK	0.006	0.007	0.011	0.8	0.9	1.4
IRL	0.008	0.009	0.011	0.6	0.6	0.8
SVN	0.010	0.011	0.014	0.7	0.8	1.0
MEX	0.007	0.010	0.015	0.6	0.8	1.2
IND	0.005	0.008	0.015	0.5	0.7	1.4
TWN	0.008	0.012	0.025	0.7	1.0	2.1
IDN	0.016	0.019	0.037	0.8	1.0	1.9
TUR	0.030	0.032	0.048	0.5	0.5	0.8

Note: This table presents the TFP gain and the Input-Output amplification factor for the median country and for the 10 countries with the largest TFP gain when markups are applied only to labor costs. Results are shown for the benchmark calibration, $(\varepsilon_\sigma, \varepsilon_\rho) = (0.70, 0.01)$ and for two other combinations of the elasticities. The TFP gain is computed as (TFP without markups)/(TFP with markups) -1. The IO amplification factor is the ratio of the TFP gain in the baseline economy over the TFP gain in the economy with $\alpha_i = 0$, $\forall i = 1, ..., n$. See Appendix C.6 for the full table of results.

5 Conclusion

We study how the sectoral linkages of production shape the aggregate TFP loss from distortions and we shed light on the crucial role played by input substitutability.

We show, analytically and quantitatively, that the TFP loss from distortions is smaller when input substitutability is lower. We find that the smaller effect of distortions is related to the smaller role played by sectoral linkages. When input substitutability is lower, the amplification from sectoral linkages is weaker, and sectors that supply intermediate inputs have a smaller impact. Moreover, we find that sectoral linkages do not systematically amplify the effect of distortions. We derive the conditions under which sectoral linkages dampen the effect of distortions. The dampening effect occurs if the elasticity of substitution is smaller than one and if additional conditions, which are more likely to hold when the distortions affect only primary inputs than when the firms' intermediate-input decisions are directly affected, are satisfied.

For our quantitative analysis, we focus on the sectoral distortions caused by market power. Using sectoral-level data from 35 countries, we find that the median TFP gain from removing sectoral-level markups is equal to 1.2%. These estimates as a lower bound of the cost of markups because they do not account for the cost of *firm*-level markups. An important message of the quantitative analysis is that using a unit elasticity of substitution (i.e., the Cobb-Douglas specification), as is commonly done in the literature, would have led to overestimating the cost of industry-level markups by a factor of 1.8. The large quantitative implications documented in our analysis call for caution in the choice of the specification of the production function as well as for more empirical evidence on the values of the elasticities of substitution.

References

Acemoglu, Daron, Vasco M. Carvaho, Asuman Ozdaglar and Alireza Tahbaz-Salehi (2012), 'The Network Originis of Aggregate Fluctuations', *Econometrica* **80**(5), 1977–2016.

Atalay, Enghin (2017), 'How important are sectoral shocks?', *American Economic Journal: Macroeconomics* **9**(4), 254–280.

Baqaee, David and Emmanuel Farhi (2020), 'Productivity and misallocation in general equilibrium', *Quarterly Journal of Economics* .

Baqaee, David Rezza and Emmanuel Farhi (2019), 'The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem', *Econometrica* **87**(4).

Barkai, Simcha (2020), 'Declining labor and capital shares', *The Journal of Finance* **75**(5), 2421–2463.

Bartelme, Dominick and Yuriy Gorodnichenko (2015), Linkages and Economic Development, NBER Working Papers 21251, National Bureau of Economic Research, Inc.

Basu, Susanto (1995), 'Intermediate Goods and Business Cycles: Implications for Productivity and Welfare', *American Economic Review* **85**(3), 512–531.

Behrens, Kristian, Giordano Mion, Yasusada Murata and Jens Suedekum (2020), 'Quantifying the gap between equilibrium and optimum under monopolistic competition', *Quarterly Journal of Economics* .

Bigio, Saki and Jennifer La'o (2020), 'Distortions in production networks', *Quarterly Journal of Economics* .

Boehm, Johannes and Ezra Oberfield (2020), 'Misallocation in the market for inputs: Enforcement and the organization of production', *Quarterly Journal of Economics* .

Caliendo, Lorenzo, Fernando Parro and Aleh Tsyvinski (2017), Distortions and the structure of the world economy. Working paper.

Carvalho, Vasco M., Makoto Nirei, Yukiko U. Saito and Alireza Tahbaz-Salehi (2015), Supply chain disruptions: evidence from the great east japan earthquake. Working paper.

de Loecker, Jan, Jan Eeckhout and Gabriel Unger (forthcoming), 'The rise of market power and the macroeconomic implications', *Quarterly Journal of Economics* .

Dhingra, Swati and John Morrow (forthcoming), 'Monopolistic competition and optimum product diversity under firm heterogeneity', *Journal of Polical Economy* .

Diamond, Peter A. and James A. Mirrlees (1971), 'Optimal Taxation and Public Production I: Production Efficiency', *American Economic Review* **61**(1).

Edmond, Chris, Virgiliu Midrigan and Daniel Yi Xu (2019), How costly are markups?, Technical report.

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

Fadinger, Harald, Christian Ghiglino and Mariya Teteryatnikova (2015), Income Differences and Input-Output Structure, Vienna Economics Papers 1510, University of Vienna, Department of Economics.

Grassi, Basile (2017), Io in i-o: Size, industrial organization, and the input-output network make a firm structurally important.

Hall, Robert E. (2018), New evidence on the markup of prices over marginal costs and the role of mega-firms in the us economy. NBER working paper.

Hang, Jing, Pravin Krishna and Heiwai Tang (2020), Input-output networks and misallocation. .

Harberger, Arnold (1954), 'Monopoly and Resource Allocation', *American Economic Review Papers and Proceedings* **44**(2), 77–87.

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

Hulten, Charles R. (1978), 'Growth Accounting with Intermediate Inputs', *Review of Economic Studies* **45**(3), 511–518.

Jones, Charles I. (2011), 'Intermediate Goods and Weak Links in the Theory of Economic Development', *American Economic Journal: Macroeconomics* **3**(2), 1–28.

Jones, Charles I. (2013), *Advances in Economics and Econometrics,Tenth World Congress*, Vol. II, Cambridge University Press, chapter Misallocation, Economic Growth, and Input-Output Economics.

Jorgenson, Dale (1967), The theory of investment behavior, *in* R.Ferber, ed., 'The Determinants of Investment Behavior', nber edn, pp. 129–156.

Karabarbounis, Loukas and Brent Neiman (2018), Accounting for factorless income, *in* 'NBER Macroeconmoics Annual'.

Klump, Rainer, Peter McAdam and Alpo Willman (2012), 'The normalized ces production function: Theory and empirics', *Journal of Economic Surveys* **26**(5), 769–799.

Leal, Julio (2017), Key sectors in mexico's economic development: a persepective from input-output linkages with sector-specific distortions.

Liu, Ernest (2019), 'Industrial policies and economic development', *Quarterly Journal of Economics* **134**(4).

Long, John B, Jr and Charles I Plosser (1983), 'Real Business Cycles', *Journal of Political Economy* **91**(1), 39–69.

Luo, Shaowen (forthcoming), 'Propagation of Financial Shocks in an Input-Output Economy with Trade and Financial Linkages of Firms', *Review of Economic Dynamics* .

Miranda-Pinto, Jorge (2021), 'Production network structure, service share, and aggregate volatility', *Review of Economic Dynamics* (39), 146–173.

Miranda-Pinto, Jorge and Eric Young (2021), Flexibility and frictions in multisector models. Working paper.

Mukoyama, Toshihiko and Latchezar Popov (2020), 'Industrialization and the evolution of enforcement institutions', *Economic Theory* **69**, 745–788.

Nunn, Nathan (2007), ' Relationship-Specificity, Incomplete Contracts and the Pattern of Trade', *Quarterly Journal of Economics* **122**(2), 569–600.

Oberfield, Ezra and Devesh Raval (2014), Micro Data and Macro Technology, NBER Working Papers 20452, National Bureau of Economic Research, Inc.

Osotimehin, Sophie (2019), 'Aggregate productivity and the allocation of resources over the business cycle', *Review of Economic Dynamics* **32**, 180–205.

Peter, Alessandra and Cian Ruane (2020), The aggregate importance of intermediate input substitutability.

Peters, Michael (forthcoming), 'Heterogeneous mark-ups, growth and endogenous misallocation', *Econometrica* .

Ramsey, Frank (1927), 'A contribution to the theory of taxation', *The Economic Journal* **37**(145).

Restuccia, Diego and Richard Rogerson (2008), 'Policy Distortions and Aggregate Productivity with Heterogeneous Plants', *Review of Economic Dynamics* **11**(4), 707–720.

Rotemberg, Julio J. and Michael Woodford (1993), Dynamic General Equilibrium Models with Imperfectly Competitive Product Markets, NBER Working Papers 4502.

Rotemberg, Julio J and Michael Woodford (1996), 'Imperfect Competition and the Effects of Energy Price Increases on Economic Activity', *Journal of Money, Credit and Banking* **28**(4), 550–577.

Wang, Wenya (2020), Financing intermediate inputs and misallocation: Evidence from china's firm-level data. .