Inflation in Distorted Economies: Evidence from India

Juan Herreño
UC San Diego

Noémie Pinardon-Touati
Columbia University

Malte Thie
Paris Dauphine

March 20, 2025

Abstract

We estimate the direct and indirect effects of demand expansions on inflation in India. Firms face supply and demand distortions that shape their prices, markups, and scale decisions. Using disaggregated data for manufacturing firms and a variety of research designs across firms, industries, and regions, we provide a portable method for the estimation of the slope of the Phillips curve, and the magnitude of the reallocative effects of monetary expansions which are driven by the distribution of demand and supply distortions. The slope of the Phillips curve in India is one order of magnitude steeper than in the United States.

JEL Classification: E31

We thank seminar participants at Columbia University, UC San Diego, the Minneapolis Fed for valuable comments. We thank Fauzi Estiko for excellent research assistance.

1. Introduction

What role do Keynesian economics play in developing economies? In particular, how do aggregate demand expansions translate into inflationary pressures or increases in real output? A large body of empirical evidence in advanced economies shows strong evidence for Keynesian mechanisms. However, developing economies differ in fundamental ways. The macroeconomic environment is characterized by high and volatile inflation, which may affect the sluggishness of the price level to changes in demand. Production may be less scalable, hindering real output expansions. Finally, developing economies are subject to large allocative distortions, which may interact with the inefficient price dispersion caused by inflation.

In this paper, we quantify the effects of domestic demand expansions on inflation in India, a large developing economy. To guide our empirical investigation, we develop a New Keynesian model with two key ingredients: firms face input wedges in the tradition of the misallocation literature, and have variable demand elasticities giving rise to variable markups. In this framework, demand shocks affect inflation not only through the output gap (the standard slope of the Phillips curve) but also through potential changes in allocative efficiency. We fully characterize the response of inflation as a function of a small number of sufficient statistics, and prove how these statistics can be identified in firm-level data on output and input prices using a number of portable causal research designs. Our main result is that the Phillips curve is steep, with a slope roughly 8 times larger than estimates for the United States. This is due to a combination of less sticky prices and steeper marginal cost curves. Changes in allocative efficiency play only a small role.

To reach our conclusions, we introduce a theoretical framework that then guides our empirical analysis. We start from the benchmark infinite-horizon New Keynesian model, with firms facing frictions to reset their prices as in Calvo (1983). We allow for decreasing returns to scale and roundabout production. We incorporate two non-standard ingredients. First, we depart from the CES assumption and allow for non-isoelastic demand curves generated by a Kimball (1995) demand system. Second, we allow for steady-state distortions in firms' input decisions, which have been demonstrated to be a key feature of the macro-economy of developing countries.

In this model, demand expansions affect inflation through two distinct channels. First,

we can define the slope of the Phillips curve as the effect on inflation of an increase in the output gap—holding aggregate allocative efficiency constant. This slope is shaped by three key forces: the extent of price rigidities, the desired passthrough of costs into prices (which reflects micro-level strategic complementarities), and the elasticity of aggregate marginal costs to changes in aggregate quantities.

Second, demand expansions endogenously affect allocative efficiency. In the steady-state of the model, the allocation of resources across firms is inefficient. Hence, demand expansions have a first-order effect on allocative efficiency, via a reallocation of production to more or less distorted firms. The first of these reallocative effects happens due to firms facing different demand elasticities, and hence charging different markups and having different pass-throughs. This channel is the same as in Baqaee, Farhi, and Sangani (2024), which we generalize to non-constant returns to scale. The second channel is new. Heterogeneous markups and input wedges generate a distribution of ex-ante combined distortions, and firms with different distortions have different pass-throughs. Demand expansions improve (worsen) allocative efficiency if they reallocate quantities to firms with ex-ante larger (smaller) combined output markups and input wedges. The allocative effects of demand shocks on inflation enter the Phillips curve as an endogenous cost-push shock: positive (negative) allocative efficiency effects work to flatten (steepen) the Phillips curve.

We fully characterize the responses of inflation and allocative efficiency to demand shocks as a function of a small number of sufficient statistics related to the response of firms to cost and demand shocks. We prove how these statistics can be identified in firm-level data on output and input prices using a number of portable causal research designs. Our approach has three key advantages. First, our methodology circumvents the identification challenges that are ubiquitous in time series identification of the relationship between demand expansions and inflation. Second, by estimating a small number of sufficient statistics, as opposed to a full parametrization of the demand system or of the supply side, our approach gains in robustness to model misspecification. Third, our approach allows to break down the response of the economy into distinct and easily-interpretable mechanisms.

We apply these insights to estimate the slope of the Phillips curve and the allocative effects of monetary policy in India. We exploit the Indian Annual Survey of Industries

(ASI), a large-scale representative survey of manufacturing establishments, over the period 1998-2017. Crucially for our research designs, the ASI contains rich information on the quantity and prices for the products that Indian firms sell, and the inputs they use. This allows us to construct price and quantity sold changes at the firm$\times$product-level across 1,200 highly disaggregated products. As evidence for the quality of our data, we show that aggregating the firm$\times$product-level changes in prices allows to closely match the dynamics of PPI inflation over our sample period. Similarly, on the input side, we construct cost and quantity purchased changes at the firm$\times$input-level.

We start by estimating the response of inflation to a change in the output gap, holding allocative efficiency constant. We show that a combination of three data moments allow us to recover the slope of the Phillips curve: the average pass-through of input cost shocks into prices, the elasticity of marginal costs to demand shocks, and the frequency of price adjustment.

The first input to estimate the slope of the Phillips curve is the pass-through of input cost shocks to product prices. A simple OLS regression of output prices on input costs would yield a biased estimate if unobserved demand shocks cause a firm to increase its price and lead to higher input costs. To circumvent this concern, we instrument the realized change in input costs by a shift-share instrument exploiting variation in firms' pre-determined exposure to intermediate inputs with different price dynamics, similar to Amiti, Itskhoki, and Konings (2019). We find a pass-through equal to 0.22 at the annual frequency, for cost shocks that are well-approximated by an AR(1) with persistence equal to 0.75. Models with constant markups and constant returns to scale would imply estimates four times as large as ours. Because of our inclusion of a battery of fixed effects, these estimates effectively hold competitors' prices constant.

Our second research design estimates the elasticity of marginal costs to changes in firm-level quantities. An obvious identification threat is the presence of firm-specific supply shocks that increase marginal costs and reduce quantities even though they are not informative about the shape of the marginal cost curve that firms face. The ideal source of variation is a demand shifter that affects quantities, and through them affects marginal costs. We introduce a demand shifter that exploits changes in product demand in the market place, and heterogeneity in the exposure of firms to different products. The identifying assumption is that firms more exposed to products that experience demand shocks, are

not, at the same time, more prone to experience supply shocks. We find that at the firm level, a one percent increase in quantities causes an increase in marginal costs of 0.2 percent, consistent with a returns to scale parameter of 0.83 in the short run.

The slope of the Phillips curve depends on the slope of the aggregate cost curves, which reflect the effect of demand shocks on input prices. We then use the same design across industries and regions to identify the elasticity of the marginal cost curve at those levels of aggregation. These IV estimates allow us to measure the shape of the marginal cost curve without imposing additional structure on input suppliers (e.g., the Frisch elasticity of labor supply), firm-level returns to scale, and without the need of using indirect inference via matching the shape of the aggregate response of the economy to monetary policy shocks, which are the two most common approaches in the literature. Our approach is transparent and robust to model mispecification in other blocks of the model. Our estimates imply that after a one percent increase in industry or regional output, marginal costs at the corresponding level of aggregation rise by 0.6-0.7 percent. Our industry-level estimates are three times larger than those estimated by Boehm and Pandalai-Nayar (2022) in the United States, suggesting meaningfully steeper marginal cost curves in India.

The final input is the frequency of price changes which we measure directly in the data. We measure a frequency of price changes of 91 percent at an annual frequency, consistent under our model assumption of random price changes with a monthly frequency of price changes of 0.18. This is higher than the number reported by Nakamura and Steinsson (2008), and comparable to references in the literature for countries and time periods with similar inflation rates as India in our period of study (Gagnon 2009; Nakamura et al. 2018; Alvarez et al. 2019).

Putting these estimates together, we find that the slope of the Phillips curve in India is 0.066 at the quarterly frequency, 8 times larger than in the United States. A key advantage of our approach is our ability to decompose the magnitude of the slope into its three subcomponents, the frequency of price changes, the pass-through of marginal costs into prices, and the elasticity of marginal costs with respect to production. The steeper marginal cost curve in India accounts for most of the difference: if the elasticity of costs to changes in output was equal to that reported by Boehm and Pandalai-Nayar (2022) for the United States, then the slope would be three times flatter. Less sticky prices matter too: if India had the same frequency of price changes observed in developed countries, our esti-

mates would be half as large. The rest of the difference is explained by differences in price pass-through estimates.

Turning to the allocative effects of demand shocks, they depend on the average and differential effect of input cost shocks on prices and quantities across firms with different ex-ante combined distortions. Importantly, these effects cannot be, in general, summarized by differential price pass-through estimates across the size distribution, since size is not a sufficient statistic for ex-ante wedges in our setting. In addition, we present evidence that the differential pass-through of prices across the size distribution, and the differential pass-through across the ex-ante distortion distribution are in fact different in our data, indicative of the empirical relevance of including in the model information on the joint distribution of size and distortions simultaneously.

The allocative effects of changes in aggregate demand are quantitatively small. The data moments that inform these small elasticities are the compressed pass-throughs for firms with different demand elasticities, and that quantity pass-throughs are not quantitatively dispersed enough across firms with different ex-ante combined distortions. These two results imply that quantitatively, assuming that the aggregate supply curve is stable after a temporary aggregate demand expansion is a good assumption for the context we study, even when the economy suffers from substantial inefficiencies.

Finally, we simulate a set of impulse response functions to illustrate the quantitative effect of a monetary policy shock on inflation and output in our setting compared to two leading studies in developed economies (Hazell et al. 2022; Gagliardone et al. 2023). We document that after an exogenous shift in monetary policy of the same magnitude, and the same parameterization of the monetary policy rule and the demand block of the economy, India's output response is half of the response in the United States and Belgium. Inflation is more than twice as responsive.

Related literature. First, this work most closely relates to the large body of work on the slope of the Phillips curve. The overwhelming majority of contributions study the United States or other OECD countries. Most papers rely on time series methods, despite ubiquitous identification and statistical issues (Mavroeidis, Plagborg-Møller, and Stock 2014). A relatively recent series of papers exploits cross-sectional variation: Beraja, Hurst, and Ospina (2019); Hazell et al. (2022); Cerrato and Gitti (2022); Gagliardone et al. (2023).

Closest to the approach implemented here, Gagliardone et al. (2023) also exploit micro-level data on prices and quantities from Belgium to estimate the slope of the Phillips curve. Compared to their work, we introduce two novel contributions. First, the analysis in Gagliardone et al. (2023) focuses on identifying the slope of the marginal cost-based Phillips curve. We develop causal research designs to identify the slope of marginal cost curves in the cross-section of firms, industries, and local labor markets, to recover the slope of the output-based Phillips curve. Second, we characterize and quantify how the slope of the Phillips curve is modified when distortions interact with pricing frictions.

We also complement the few studies that study the slope of the Phillips curve in the context of emerging and developing economies, all using time series methods (Mohanty, Klau et al. 2001; Filardo and Lombardi 2014; Ball, Chari, and Mishra 2016).

Second, we relate to a recent literature studying the effects of monetary shocks on allocative efficiency (Reinelt and Meier 2020; Mongey 2021; Baqaee, Farhi, and Sangani 2024). Our framework builds on Baqaee, Farhi, and Sangani (2024), which we augment to account for potentially non-constant returns to scale and the presence of input wedges. These new ingredients qualitatively affect predictions regarding the effect of monetary shocks for allocative efficiency. We show non-parametric identification of the sufficient statistics that determine the allocative efficiency effect. Finally, we provide a quantitative analysis of this channel in a large developing country—where this type of effects are likely to be strongest.

Third, we complement the literature on pricing decisions in emerging and developing countries (Gagnon 2009; Alvarez et al. 2019; Drenik and Perez 2020). While existing studies focus on hyperinflation episodes, we study a setting with moderately high and persistent inflation, which is representative of "normal times" in these countries. Second, we focus on estimating the effect of domestic demand expansions on inflation.

Finally, our work contributes to the literature on the passthrough of cost shocks into prices (Gopinath and Rigobon 2008; Gopinath and Itskhoki 2011; Amiti, Itskhoki, and Konings 2019) and the literature on micro- and macro-level cost curves (Shea 1993; Bresnahan and Ramey 1994; Boehm and Pandalai-Nayar 2022). We use these estimates as the building blocks to our estimation of the slope of the Phillips curve.

2. Motivating facts

2.1. Inflation in developing countries is high and variable

In this subsection we document a simple fact: average inflation rates decline along the development path. We use data on CPI inflation and GDP per capita in constant 2005 dollars coming from the World Development Indicators.

Figure 1 illustrates using a binned scatter plot, the cross-country relation between GDP per capita in logarithms and CPI inflation for the period 1980-2023. Each observation corresponds to one country and the values are the median value of each variable across years. There is a strong negative association between the two variables. Low income countries have, on average, higher average inflation rates. Our choice of using the median has the intention of filtering the effect of influential observations driven by periods of hyperinflations.

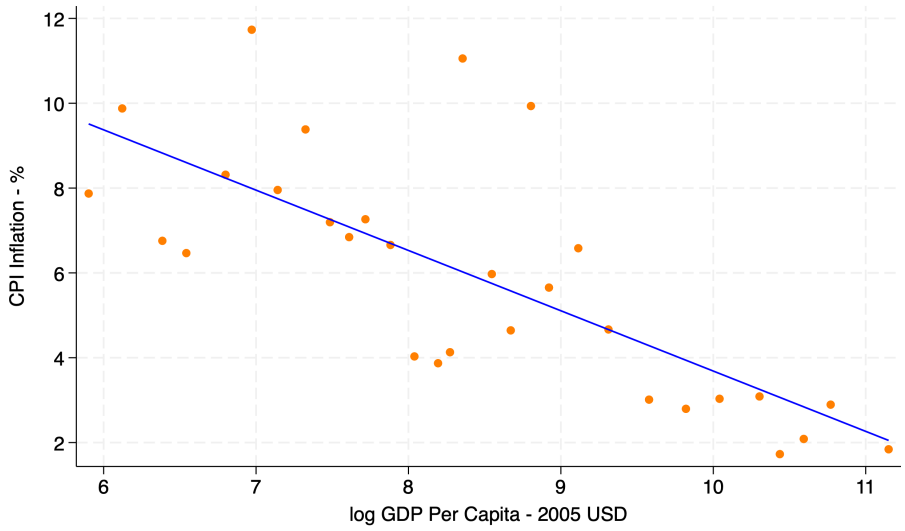


FIGURE 1. Inflation in the cross-section of countries

Note: The Figure shows the association between the median log GDP per capita in a country and its median CPI inflation rate from 1980-2023. Source: World Development Indicators.

2.2. Developing countries are subject to large allocative distortions

A large literature in macroeconomics and economic growth has documented large allocative distortions in developing countries. Intuitively, cost minimization implies that firms

equalize the marginal revenue product of their inputs to their corresponding input costs. If there is dispersion in actual or implicit taxes and subsidies to the input costs firms face, they will translate into dispersion in the marginal product revenue of different inputs. Policies or shocks that induce reductions in the dispersion of marginal products across firms will therefore increase allocative efficiency.

Although we will document the dispersion of input wedges across firms in the sample of firms we will use, in this section we will offer some references that is indicative of the large allocative distortions that developing economies face.

The classical reference is the seminal work of Hsieh and Klenow (2009), that studies the effect of allocative distortions in China, India, and the United States. The ratios of 90th to 10th percentiles of TFPR are 5.0 in India, 4.9 in China, and 3.3 in the United States, consistent with the idea that developing economies face substantially more acute distortions than developed countries.

The Hsieh and Klenow (2009) analysis has been extended to other countries, refined in their methods and sources of variation (Oberfield 2013; Gopinath et al. 2017; Bau and Matray 2023), but the main message persists, developing economies are highly distorted.

2.3. Unobserved shocks make difficult the estimation of the Phillips curve using aggregate data

Our approach in this paper avoids conducting an estimation of the shape of the Phillips curve using time-series techniques for two reasons. First, inference is difficult due to weak instruments problems associated with difficulty of forecasting future inflation (Mavroeidis, Plagborg-Møller, and Stock 2014) , and due to the presence of inflation shifters that are correlated with economic activity, for example changes in beliefs about the future conduct of monetary policy (Hazell et al. 2022). Second, the various layers of variation we will use will equip us to decompose the slope of the Phillips curve in India in different elasticities that reflect different economic mechanisms.

It is useful to inspect two episodes of recent economic history in India to understand the magnitude of the problem of using time series analysis to uncover the slope of the Phillips curve. We use what Stock and Watson (2020) call the *Phillips Correlation* which is an association between inflation and quantities in equilibrium, as opposed to the structural interpretation of the Phillips curve which plays the role of the aggregate supply schedule.

Our measure of inflation is the percent change in the manufacturing Wholesale Price Index, and for quantities we use the cyclical component of the manufacturing quantity index that we recover using the Hamilton (2018) filter.

Figure 2 highlights two key episodes of recent Indian monetary history. Both panels show the same data, but highlight different historical episodes. The x-axis in each figure is the Output gap, and the y-axis is the year-over-year inflation rate. Each dot corresponds to a particular date, which we caption. The top panel highlights the period after the Great Financial Crisis, from 2010 to 2012, which was characterized by a strong deceleration of the economy together with a stable inflation rate of around 7%. Interestingly, this parallels discussion of a "missing disinflation" in the United States at the same time. The second period we highlight is the plummeting of the inflation rate around 2014 after the newly-appointed governor of the Reserve Bank of India Raghuram Rajan announced the adoption of inflation targeting, with potential effects for long-run inflation expectations and consequently for the inflation rate, as well as the end of adverse supply shocks. This vertical shift in the Phillips correlation reminisces the United States during the early 1980s.

Of course, it is a possibility that the Phillips correlation is reflecting a time-varying slope of the Phillips curve that changed from being almost perfectly flat to almost perfectly vertical in a matter of two years. A more plausible possibility however, is that the Phillips correlation is reflecting the occurrence of shifters to aggregate supply, let them be triggered by changes in monetary policy, financial shocks, or changes in relative prices. We develop a method for estimating the structural Phillips curve that is immune to these concerns.

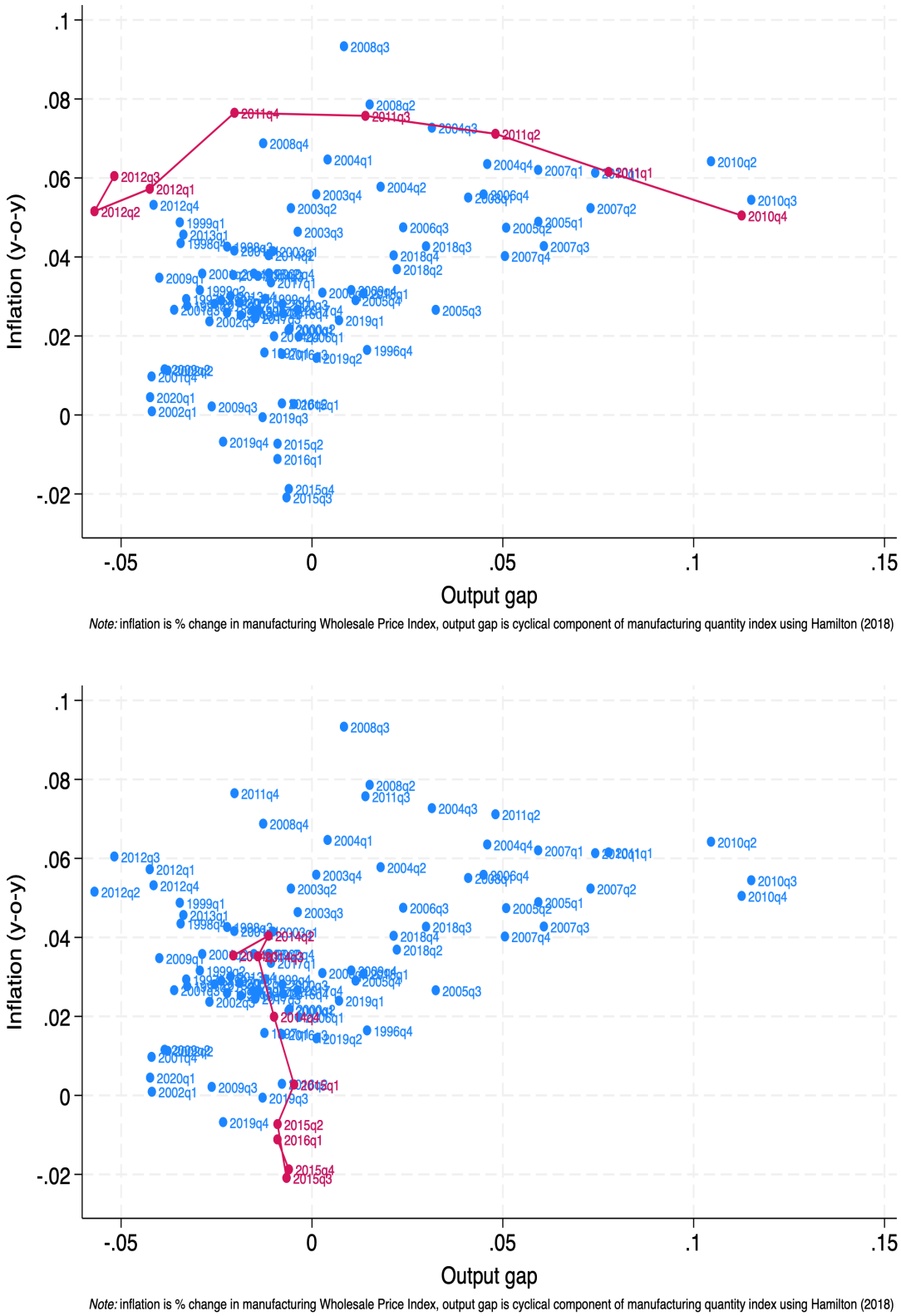


FIGURE 2. Phillips correlation in India

Note: The x-axis is the transitory component of manufacturing output using the Hamilton (2018) index. The y-axis corresponds for the y-o-y inflation rate for the WPI manufacturing index in India. The solid lines connect the points for two subperiods.

3. Model

This section describes the theoretical framework that guide our empirical investigation. The main text provides a succinct description and we leave all details and proofs to Appendix A.

3.1. Environment

The economy is composed of five sectors. Households consume the final good, save, and supply labor. A final good producer produces the final good using differentiated varieties indexed by $i \in [0,1]$. Producers of each differentiated variety i produce using labor and materials and have sticky prices. Material producers use the final good as input and produce materials used by the differentiated varieties producers. A monetary authority sets the nominal interest rate.

Households. Households choose consumption C and labor L to maximize discounted future utility $\mathbb{E}_0\left[\sum_{t=0}^{+\infty} \beta^t u(C_t, L_t)\right]$ subject to a per-period budget constraint $P_t^Y C_t + Q_t B_t = B_{t-1} + w_t^l L_t + T_t$ where P_t^Y is the price of the consumption bundle, B_t is holdings of one-period risk-free nominal bonds with price Q_t, w_t^l is the wage, and T_t denotes any profits rebated to households as lump-sum.

Final good producers. Let Y_t denote aggregate production of the final good. Y_t can be used for consumption C_t or as an input to produce materials Y_t^X so that $Y_t = C_t + Y_t^X$.

The final good Y_t is produced by a perfectly competitive firm using a bundle of differentiated intermediate inputs y_{it} for $i \in [0,1]$. Intermediate input varieties are assembled into the final good using the Kimball aggregator:

$$\int_0^1 \Upsilon\left(\frac{y_{it}}{Y_t}\right) di = 1$$

where the function $\Upsilon(.)$ is strictly increasing, strictly concave, and satisfies $\Upsilon(1) = 1$. The CES aggregator is the special case $\Upsilon(q) = q^{\frac{\theta-1}{\theta}}$ for $\theta > 1$.

Taking the prices p_{it} of the inputs as given and denoting the price of the final good P_t^Y, the final good producer chooses y_{it} to maximize profits. This gives rise to the demand

function:

$$\frac{y_{it}}{Y_t} = \Upsilon'^{-1}\left(\frac{p_{it}}{\mathcal{P}_t}\right) \tag{1}$$

where $\frac{p_{it}}{\mathcal{P}_t}$ determines substitution across varieties. The price index $\mathcal{P}_t$ is given by $\mathcal{P}_t = \frac{P_t^Y}{D_t}$. $P_t^Y = \int_0^1 p_{it}\frac{y_{it}}{Y_t}di$ is the ideal price index. $D_t = \int_0^1 \Upsilon'(\frac{y_{it}}{Y_t})\frac{y_{it}}{Y_t}di$ is a "demand" index. When demand is CES, D_t is a constant equal to $\frac{\theta}{\theta-1}$. Away from the CES case, D_t is not a constant and is increasing in the dispersion of quantity shares.

The price elasticity of demand is only a function of firm relative size and is given by:

$$\theta_{it} = \theta\left(\frac{y_{it}}{Y_t}\right) = -\frac{\partial \log y_{it}}{\partial \log p_{it}} = \frac{\Upsilon'\left(\frac{y_{it}}{Y_t}\right)}{-\frac{y_{it}}{Y_t}\Upsilon''\left(\frac{y_{it}}{Y_t}\right)} \tag{2}$$

Differentiated varieties producers. Each variety i is produced by a single firm. Firms produce with the following technology:

$$y_{it} = e^{z_i}v_{it}^a \quad \text{with} \quad v_{it} = l_{it}^{\phi}x_{it}^{1-\phi} \tag{3}$$

v_{it} is a constant returns-to-scale composite of the variable inputs, labor l_{it} and materials x_{it}. The unit cost of inputs are w_t^l and w_t^x, respectively. From the cost minimization problem, we obtain $w_t^v = (w_t^l/\phi)^{\phi}(w_t^x/(1-\phi))^{1-\phi}$ the unit price of the composite variable input. We introduce input price distortions as wedges in the tradition of Hsieh and Klenow (2009). Each firm i pays $w_t^v(1+\tau_i)$ per unit of variable input. τ_i is a mean-zero steady-state distortions in firm size. This gives rise to the following expression for marginal costs: $mc_{it} = (1+\tau_i)w_t^v\frac{1}{a}e^{-\frac{1}{a}z_i}y_{it}^{\frac{1-a}{a}}$.

A firm has a probability $1-\alpha$ of being able to reset its price in each period. A firm that can reset its price chooses the price that maximizes:

$$\max_{p_{it}} \mathbb{E}_t\left[\sum_{s=0}^{+\infty}\alpha^s\Lambda_{t,t+s}\left[p_{it}y_{it+s} - \mathcal{C}_{it+s}(y_{it+s}, w_{t+s}^v)\right]\right]$$

subject to the demand curve $y_{it+s} = \Upsilon'^{-1}\left(\frac{p_{it}}{\mathcal{P}_{t+s}}\right)Y_{t+s}$ and the cost function $\mathcal{C}_{it+s}(y_{it+s}, w_{t+s}^v) = (1+\tau_i)w_{t+s}^v e^{-\frac{1}{a}z_i}y_{it+s}^{\frac{1}{a}}$. $\Lambda_{t,t+s}$ is the stochastic discount factor of the representative house-

hold.

It will be convenient to define the following quantities. $\mu^f_{it} = \frac{\theta_{it}}{\theta_{it}-1}$ is the desired markup that the firm would choose in a flexible price environment. $\Gamma_{it} = \frac{\partial \log \mu^f_{it}}{\partial \log \frac{y_{it}}{Y_t}}$ is the elasticity of the flexible price markup with respect to relative size. $\rho_{it} = \frac{1}{1+\Gamma_{it}\theta_{it}}$ is the partial equilibrium pass-through of a marginal cost shock into the firm's price in a flexible price environment. In the CES case, all firms face the same demand elasticities, $\Gamma_{it} = 0 \; \forall i$, and $\rho_{it} = 1 \; \forall i$. Away from the CES case, ρ_{it} can be below or above 1 depending on the sign of Γ_{it}. Note that μ^f_{it}, Γ_{it}, and ρ_{it} are only a function of a firm's relative size $\frac{y_{it}}{Y_t}$. Finally, in a sticky price environment, the actual markup of the firm may differ from the flexible price desired markup. We denote the actual markup of the firm: $\mu_{it} = \frac{p_{it}}{mc_{it}}$.

Materials producers. Materials are produced by a perfectly competitive representative firm with the following production function: $X_t = (Y^X_t)^{\frac{\eta}{1+\eta}}$ where Y^X_t are units of the final good used for intermediate inputs production and $\eta \geq 0$. This implies that the price of the material good is given by: $w^x_t = P^Y_t(1+\eta^{-1})X_t^{\eta^{-1}}$.

Monetary authority. Monetary policy sets the nominal interest rate according to a Taylor rule.

Equilibrium. Equilibrium is defined by the following conditions: (i) consumers choose consumption and labor to maximize utility taking prices and wage as given; (ii) firms with flexible prices set prices to maximize their value taking the price index and their residual demand curves as given; firms with sticky prices meet demand at fixed prices; (iii) monetary policy sets the nominal interest rate; (iv) all resource constraints are satisfied.

We solve the model by log-linearization around the zero-inflation steady state. The steady-state distribution of firm size depends on joint distribution of (z_i, τ_i). We take a first-order expansion for small monetary policy shocks. Quantities without a t subscript refer to the steady-state. We denote $\lambda_i = \frac{p_i y_i}{P^Y Y}$ sales share in steady-state. Let $\mathbb{E}_\lambda[X_{it}] = \int_0^1 \lambda_i X_{it} di$.

3.2. Characterization

Marginal cost-based Phillips curve. A firm that can reset its price at time t will choose:

$$\hat{p}_{it|t}^{new} = (1-\beta\alpha)\mathbb{E}_t\left[\sum_{s=0}^{+\infty}(\beta\alpha)^s\left(\zeta_i\rho_i\hat{mc}_{t+s} + (1-\zeta_i\rho_i)\hat{\mathcal{P}}_{t+s}\right)\right] \tag{4}$$

where $\hat{mc}_t \equiv \mathbb{E}_\lambda[\hat{mc}_{it}]$ is the change in the aggregate nominal marginal cost and $\zeta_i = \frac{1}{1+\frac{1-a}{a}\theta_i\rho_i}$. Aggregating across firms, we obtain the marginal cost-based Phillips curve:

$$\hat{\pi}_t = \varphi\mathbb{E}_\lambda[\zeta_i\rho_i](\hat{mc}_t - \hat{P}_t^Y) - \varphi(1-\mathbb{E}_\lambda[\zeta_i\rho_i])\hat{D}_t + \beta\mathbb{E}_t[\hat{\pi}_{t+1}] \tag{5}$$

$\hat{mc}_t - \hat{P}_t^Y$ is the change in the aggregate real marginal cost. $\varphi = \frac{(1-\alpha)(1-\beta\alpha)}{\alpha}$ is the slope of the marginal cost-based Phillips curve in the case of constant returns to scale and CES demand. ρ_i is the flexible price partial equilibrium pass-through of marginal cost shocks into prices and reflects strategic complementarities in price-setting. ζ_i captures the fact that when returns to scale are below 1, a cost shock induces an adjustment in size, which dampens the first-round effect on marginal cost. Together, $\mathbb{E}_\lambda[\zeta_i\rho_i]$ is the average flexible price passthrough of an input cost shock into prices. A lower $\mathbb{E}_\lambda[\zeta_i\rho_i]$, reflecting stronger (micro-level) real rigidities, flattens the Phillips curve.

$\hat{D}_t$ is the change in the demand index. All else equal, a higher $\hat{D}_t$ implies that individual firms compete with a more aggressive price index, which works towards lowering inflation.

Aggregate marginal costs. Let us define aggregate productivity Z_t as satisfying

$$Y_t \equiv Z_t(L_t^\phi X_t^{1-\phi})^a \tag{6}$$

From the market clearing conditions, we obtain the solution for aggregate marginal cost:

$$\hat{mc}_t = \underbrace{\left[\frac{1-a+\psi}{a} + \upsilon\left(1-(1-\lambda_c)\frac{(1+\psi)}{a}\right)\right]}_{\Omega = \text{Elasticity of mc wrt output}}\hat{Y}_t - \underbrace{\left[\frac{\psi}{a} - \upsilon(1-\lambda_c)(1+\psi)\right]}_{\Xi = \text{Elasticity of mc wrt TFP}}\hat{Z}_t + \hat{P}_t^Y \tag{7}$$

$\psi = \frac{\phi\nu^{-1}+(1-\phi)\eta^{-1}+\nu^{-1}\eta^{-1}}{1+\nu^{-1}(1-\phi)+\phi\eta^{-1}}$, where $\nu^{-1} = \frac{u_{ll}L}{u_l} - \frac{u_{cl}L}{u_c}$ is the inverse Frisch elasticity of labor supply. ψ is the slope of the variable input supply curve, reflecting the slope of the labor supply curve ν^{-1} and the slope of the materials supply curve η^{-1}. λ_c is the steady-state share of consumption in total output. υ depends on the chosen assumption on consumption-labor complementarities in the utility function. Under GHH preferences, $\upsilon = 0$. With separable preferences, $\upsilon > 0$, and in the case with only labor in the production function, $\upsilon = \sigma^{-1}$ where $\sigma^{-1} = -\frac{u_{cc}C}{u_c}$ is the inverse elasticity of intertemporal substitution.

Output-based Phillips curve. Combining (5) and (7), we obtain the output-based New Keynesian Phillips curve:

$$\hat{\pi}_t = \varphi\mathbb{E}_\lambda[\zeta_i\rho_i]\left(\Omega\hat{Y}_t - \Xi\hat{Z}_t\right) - \varphi(1-\mathbb{E}_\lambda[\zeta_i\rho_i])\hat{D}_t + \beta\mathbb{E}_t[\hat{\pi}_{t+1}] \tag{8}$$

Allocative efficiency. Let us define the combined allocative distortion as $m_{it} \equiv \mu_{it}(1+\tau_i)$. Let us define the aggregate steady-state distortion as: $\mathcal{M} \equiv a\frac{P^Y Y}{w^v V} = \mathbb{E}_\lambda[m_i^{-1}]^{-1}$. The change in the demand index and aggregate productivity are characterized by:

$$\hat{D}_t = -\text{Cov}_\lambda\left[\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}, \hat{p}_{it}\right] \tag{9}$$

$$\hat{Z}_t = -\text{Cov}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}, \hat{y}_{it}\right] \tag{10}$$

In the simple case where there is no dispersion in input wedges, one can show that:

$$\hat{Z}_t = -\mathcal{M}\hat{D}_t$$

The demand index captures misallocation stemming from heterogeneous demand elasticities. When $\text{Cov}_\lambda[\theta_i, \hat{p}_{it}] > 0$, the relative price of firms with initially low markups rises relative to other firms, reallocating resources away from those firms and towards high markup firms. This increases allocative efficiency. When all $\tau_i = 0$, the only cause of allocative inefficiencies is the dispersion in markups due to imperfect competition. In the more general case with non-constant input wedges τ_i, the change in allocative efficiency depends on whether quantities are reallocated towards high or low combined distortion firms.

Using the solution of the firm's pricing problem, we obtain the law of motions for $\hat{D}_t$ and $\hat{Z}_t$ in terms of model parameters and steady-state values:

$$\hat{D}_t = -\frac{\varphi\kappa_D\left(\Omega\,\hat{Y}_t - \Xi\hat{Z}_t\right)}{1+\beta+\varphi(1+\kappa_D)} + \frac{\hat{D}_{t-1}+\beta\hat{D}_{t+1}}{1+\beta+\varphi(1+\kappa_D)} \tag{11}$$

$$\hat{Z}_t = \frac{\varphi\kappa_Z\left(\Omega\,\hat{Y}_t + \hat{D}_t\right)}{1+\beta+\varphi(1+\Xi\kappa_Z)} + \frac{\hat{Z}_{t-1}+\beta\hat{Z}_{t+1}}{1+\beta+\varphi(1+\Xi\kappa_Z)} \tag{12}$$

$$\text{with } \kappa_D = \mathbb{E}_\lambda[\zeta_i\rho_i]\left(\mathbb{E}_\lambda\left[\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\frac{\zeta_i\rho_i}{\mathbb{E}_\lambda[\zeta_i\rho_i]}\right]-1\right)$$

$$\kappa_Z = \mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]\left(\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i\zeta_i\rho_i}{\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]}\right]-\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\right]\right)$$

The parameters κ_D and κ_Z are the equivalent of the covariances in (9) and (10), noticing that the changes in relative prices $\hat{p}_{it}$ are proportional to the firm-level passthrough coefficients $\zeta_i\rho_i$, and the changes in relative quantities $\hat{y}_{it}$ are proportional to the firm-level coefficients $\theta_i\zeta_i\rho_i$. When ex-ante allocative distortions covary negatively with pass-throughs, output booms $\hat{Y}_t$ driven by monetary shocks are concomitant with improvements in aggregate productivity.

A five-equations New Keynesian model. Additionally solving for the Euler equation (see Appendix equation A.50), we obtain a five-equation version of the New Keynesian model:

$$\text{(NKPC)} \qquad \hat{\pi}_t = \varphi\mathbb{E}_\lambda[\zeta_i\rho_i]\left(\Omega\hat{Y}_t - \Xi\hat{Z}_t\right) - \varphi(1-\mathbb{E}_\lambda[\zeta_i\rho_i])\hat{D}_t + \beta\mathbb{E}_t[\hat{\pi}_{t+1}]$$

$$\text{(LOM for D)} \qquad \hat{D}_t = -\frac{\varphi\kappa_D\left(\Omega\,\hat{Y}_t - \Xi\hat{Z}_t\right)}{1+\beta+\varphi(1+\kappa_D)} + \frac{\hat{D}_{t-1}+\beta\hat{D}_{t+1}}{1+\beta+\varphi(1+\kappa_D)}$$

$$\text{(LOM for Z)} \qquad \hat{Z}_t = \frac{\varphi\kappa_Z\left(\Omega\,\hat{Y}_t + \hat{D}_t\right)}{1+\beta+\varphi(1+\Xi\kappa_Z)} + \frac{\hat{Z}_{t-1}+\beta\hat{Z}_{t+1}}{1+\beta+\varphi(1+\Xi\kappa_Z)}$$

$$\text{(Euler equation)} \qquad c\hat{Y}_t - \tilde{c}\left(\hat{Y}_t - \hat{Z}_t\right) = \mathbb{E}\left[c\hat{Y}_{t+1} - \tilde{c}\left(\hat{Y}_{t+1} - \hat{Z}_{t+1}\right)\right] - \sigma\left(\hat{i}_t - \mathbb{E}_t\left[\hat{\pi}_{t+1}\right]\right)$$

$$\text{(MP rule)} \qquad \hat{i}_t = \phi_\pi\hat{\pi}_t + \phi_y\hat{Y}_t + \varepsilon_t^{MP}$$

Taking the model to the data. Our objective is to quantify the link between demand expansions and inflation as described by equation (NKPC). We first estimate the elasticity

of inflation with respect to the output gap, holding allocative efficiency constant; i.e., we estimate the standard slope of the NKPC $\kappa_Y = \varphi \mathbb{E}_\lambda[\zeta_i \rho_i]\Omega$. This analysis is performed in Section 5. We then turn to the endogenous response of D_t and Z_t. To do so, we estimate the coefficients κ_Z and κ_D in Section 6.

4. Data

Our main data source is the Indian Annual Survey of Industries (ASI). We use data from 1998 to 2017. The ASI contains information on a representative sample of manufacturing establishments, conditional on them taking part of the organized sector, and either employing more than 20 employees, or employing more than 10 employees and using electricity. In terms of coverage, the value added of establishments in ASI covers 61% of total manufacturing value added in India as reported in the national accounts (the latter includes small establishments that the ASI excludes and the informal sector).

The main advantage of the ASI is that both for the products that manufacturing plants produce and the inputs they buy, we observe information on sales, quantities, and unit values, at the product level. Most surveys of individual firms around the world report information on revenues and expenditures, which on top of being a limiting factor to understand the factors that determine price-setting, create alternative issues on other procedures we undertake, like the estimation of production functions, markups, and input wedges (Bond et al. 2021; Hashemi, Kirov, and Traina 2022). We construct a harmonized classification of products and inputs over the whole sample, allowing us to classify all products into around 1,200 product codes (approximately corresponding to the 5-digit level of the India NPC classification).

We construct a panel of firm-product prices and quantities. We denote by $\Delta \log p_{ijt}$ and $\Delta \log q_{ijt}$ the change in the price p and quantity of product j sold by firm i at time t. Even working with narrowly-defined product categories, unobserved heterogeneity could prevent a meaningful comparison of prices across firms. We always work with within firm$\times$product changes in the price (or quantity), largely alleviating this concern. Similarly for inputs, we observe purchase value, unit price, and quantity purchased for material inputs (classified in the same 1,200 products), energy inputs (electricity, oil, and coal), and labor. We denote by $\Delta \log w_{ikt}$ and $\Delta \log x_{ikt}$ the log change in prices and quantities of

input k used by firm i.

We aggregate changes in prices at the firm-level by using Törnqvist weights. The change in the price index of firm i is $\Delta \log p_{it} = \sum_{j \in \mathcal{J}_i} \bar{s}_{ijt} \Delta p_{ijt}$, where s denotes the weight of product j in the portfolio of firm i, and we use the convention of placing a bar on top of the share to denote that these shares are the mid-point of the shares in $t-1$ and t. Similarly, the change in the price of variable inputs of firm i is $\Delta \log w_{it}^{v} = \sum_{k \in \mathcal{K}_i} \bar{s}_{ikt} \Delta w_{ikt}$.

Sample of analysis. We restrict the sample of analysis to plant-year observations which reported a positive output, as well as at least one day worked by at least one person within the plant. Moreover, we restrict the sample to observations that displayed consistent accounting values, i.e. for which individual input and output components closely summed to their reported aggregate values, and only include firms that report positive sales values for at least one product. Finally, we construct a harmonized version of 3-digit industries in the Indian National Industry Classification (NIC), splitting some highly populated sectors into their 4-digit sub-sectors and aggregating other sectors. Our final sample is an unbalanced panel of over 180,000 individual plants in 81 3-digit NIC sectors.

5. Estimation of the slope of the Phillips curve

In this section, we estimate the elasticity of inflation with respect to the output gap in the Phillips curve (NKPC). That is, we estimate $\kappa_Y = \varphi \mathbb{E}_\lambda[\zeta_i \rho_i] \Omega$, which we refer to as the slope of the Phillips curve.

Notice that the slope of the Phillips curve is the product of three objects. These objects are the aggregate flexible-price passthrough of cost shocks to prices $\mathbb{E}_\lambda[\zeta_i \rho_i]$, the elasticity of marginal costs with respect to quantities Ω, and the adjustment factor that comes from our price rigidity assumption under Calvo pricing φ. In what follows, it will be useful to relabel the slope of the Phillips curve as the product of three terms as our research design will provide information on these three pieces separately. Specifically,

$$\kappa^Y = \frac{1}{\alpha} \times (1-\alpha)(1-\beta\alpha)\mathbb{E}_\lambda[\zeta_i \rho_i] \times \Omega \tag{13}$$

Equation (13) is reminiscent of the Galı and Gertler (1999) decomposition of the slope of the Phillips curve into the slope of the marginal cost-based curve κ_{mc} and the output

elasticity of marginal costs Ω. Our key innovation is in how we take this equation to the data. First, we will directly measure each of the key elasticities from cross-sectional research designs that are robust to the presence of unobserved shocks that drive spurious correlation between prices, marginal costs, and quantities, as well as to measurement error in marginal costs and quantities. Second, our research designs do not require us to take a strong stance on the specific microfoundations that shape those elasticities as long as those microfoundations end up producing a Phillips curve that is composed of the terms in equation (13). Importantly, we do not need to take a stance on the structure of preferences that gives rise to a supply curve for inputs, or the specific shape of demand curves that creates differences in pass-throughs. By contrast, the usual practice in the literature is to calibrate some, or all of the elasticities by imposing more rigid functional forms on preferences, input supply, and final products demand, and oftentimes using impulse responses to aggregate shocks, which increase the potential of mispecification in blocks of the model different than the Phillips curve.

5.1. The elasticity of firm-level prices to firm-level cost shocks

First, the firm-level partial equilibrium passthrough of cost shocks into prices identifies $(1-\alpha)(1-\beta\alpha)\mathbb{E}_\lambda[\zeta_i\rho_i]$. To make this point, in Appendix B.1 we augment our baseline model with firm-specific material cost shocks ϑ_{it}, which we can decompose into an aggregate component ϑ_t and a mean-zero idiosyncratic component $\tilde{\vartheta}_{it}$. We show here the case of a zero-persistence shock, and leave the derivations of the case of a persistent shock to the Appendix. When a firm gets a chance to reset its price, the optimal reset price is now given by:

$$\hat{p}^{new}_{it|t} = (1-\beta\alpha)\mathbb{E}_t\left[\sum_{s=0}^{+\infty}(\beta\alpha)^s\left(\zeta_i\rho_i\left(\hat{mc}_{t+s} + (1-\phi)\tilde{\vartheta}_{it+s}\right) + (1-\zeta_i\rho_i)\hat{\mathcal{P}}_{t+s}\right)\right], \tag{14}$$

The Calvo assumption implies that the price charged by firm i in period t, is equal to the its reset price whenever it gets a chance to reset its price, and equal to its last period's price whenever it does not. Formally, $\hat{p}_{it} = \mathbb{1}^p_{it}\hat{p}^{new}_{it|t} + (1-\mathbb{1}^p_{it})\hat{p}_{it-1}$, where $\mathbb{1}^p_{it}$ is a "Calvo-fairy" dummy that takes the value of 1 if a price adjustment is permissible for firm i in period t. By plugging equation (14) into the Calvo law of motion of individual prices we just stated,

we derive an expression for the price change of firm i in period t,

$$\hat{p}_{it}-\hat{p}_{it-1} = \mathbb{1}^{p}_{it}(1-\beta\alpha)\zeta_i\rho_i(1-\phi)\tilde{\vartheta}_{it}+\mathbb{1}^{p}_{it}(1-\beta\alpha)\mathbb{E}_t\left[\sum_{s=0}^{+\infty}(\beta\alpha)^s\left(\zeta_i\rho_i\hat{mc}_{t+s}+(1-\zeta_i\rho_i)\hat{\mathcal{P}}_{t+s}\right)\right]-\mathbb{1}^{p}_{it}\hat{p}_{it-1}.$$

Let us assume that we observe $\mathcal{Z}^{\vartheta}_{it}$, a proxy for $(1-\phi)\vartheta_{it}$ satisfying $(1-\phi)\vartheta_{it} = k^{\vartheta}\mathcal{Z}^{\vartheta}_{it}$. Our first identification result of this section, proven in Appendix B.2, is that a regression of $\Delta\log p_{it}$ on $\Delta\log w^{v}_{it}$ instrumented by the exogenous zero-persistence cost shifter $\mathcal{Z}^{\vartheta}_{it}$, yields in population an IV estimate equal to:

$$\beta^{IV}_{p,w} = (1-\alpha)(1-\beta\alpha)\mathbb{E}_{\lambda}\left[\rho_i\zeta_i\right] \tag{15}$$

This elasticity corresponds to the second term in equation (13).

Intuitively, equation (15) states that after an exogenous change in input costs of 1%, firms change their price with an elasticity equal to ρ_i, which triggers changes in quantities, and due to potentially upward-sloping marginal cost curves, triggers endogenous changes in marginal costs. The strength of the fixed point between prices and marginal cost changes is captured by ζ_i. These two economic mechanisms would be the end of the story if firms were fully able to adjust their prices, but the presence of price rigidity implies that only a fraction $1-\alpha$ of firms is able to reset their prices, and when they do, they take into account that with some probability they will not be able to reset their prices in the future, giving rise to the full term in the equation.

Empirical strategy. The argument derived in the model lends itself to a research design in a panel of firms. Specifically, we estimate the following regression

$$\Delta\log p_{ijt} = \Theta_{jt} + \beta_{p,w}\Delta\log w^{v}_{it} + \epsilon_{ijt}, \tag{16}$$

where $\Delta\log p_{ijt}$ is the change in the price charged by firm i for a given product j, Θ_{jt} is a set of product$\times$time, fixed effects meant to soak variation driven by product-specific shocks, and $\Delta\log w^{v}_{it}$ are our estimates of the change in firm-level input price index, where we consider all variable inputs: labor and material inputs (including energy).

The main concern of estimating equation (16), is that it is very natural to think that firms experience demand shocks for their products and that firms face upward-sloping

supply curves for some or all of their inputs. If that is the case, then a demand shock will induce a positive correlation between prices and input costs, but for an economic mechanism that is distinct to the pass-through of cost shocks we are aiming to estimate. A second empirical concern is that our estimates of pass-through may be attenuated due to measurement error in input costs.

We address this empirical concern by using an instrumental variable approach. For a set of materials and energy inputs (coal, oil, and electricity) that we call $\underline{\mathcal{K}}$, we construct the instrument

$$\mathcal{Z}_{it}^{\vartheta} = \sum_{k \in \underline{\mathcal{K}}} s_{ikt-1} \Delta \log w_{ikt}, \tag{17}$$

where w_{ikt} is the price of input k paid by firm i in period t, and s_{ikt-1} is the pre-determined share of input k used by firm i. Note that the instrument uses lagged weights (as opposed to mid-point weights in the construction of the input price index) and excludes labor, which is the input where it is most likely that firm-specific supply curves play a role. This instrument is the same as the one used by Amiti, Itskhoki, and Konings (2019), except that they only observe input prices for imported inputs. Our identifying assumption is that conditional on controls, firms with a large share of inputs that experience price increases do not experience differential demand shocks.

Results. Table 1 presents our results. Columns 1-3 present OLS estimates, while columns 4-6 present our IV estimations. Our benchmark IV elasticity is equal to 0.22, and remains stable when we control for lagged sales. It is worth noting that our IV estimates are roughly double than our OLS elasticities, suggesting an important role of the instrument in correcting for measurement error in our measure of input costs. The inclusion of product$\times$time fixed effect is crucial to control for the effect of competitor prices in the pricing decisions of firm i, as well for the presence of product demand and supply shocks. In Appendix B.2 we present an extension of our identification argument in which we allow for persistence of cost disruption according to an AR(1) process. We show that estimating a local projection analog to regression (16) allows us to recover the persistence parameter, which is equal to 0.75.

In the limit of perfect flexible price pass-through ($\rho_i = 1$ for all i) and constant returns

TABLE 1. Elasticity of price changes to input cost changes

	$\Delta \ln p$					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln w$	0.139***	0.111***	0.111***	0.291***	0.223***	0.222***
	(0.008)	(0.008)	(0.008)	(0.019)	(0.017)	(0.017)
Year FE	✓			✓		
Year × Product FE		✓	✓		✓	✓
Controls			✓			✓
Observations	356,162	353,568	353,568	355,530	352,934	352,934
F-stat				5195.8	5272.9	5280.9

Note: This table presents the estimation results behind equation 16. The first three columns present OLS regressions, while the final three columns present IV estimates. Different columns include different sets of fixed effects or controls. Our benchmark specification is that of column (6).

to scale ($\zeta_i = 1$ for all i), our model predicts that $\beta^{IV}_{p,w} = (1-\alpha)(1-\beta\alpha)$, which is a function of how forward-looking firms are and the share of prices that cannot be adjusted for a year since our data is annual. Notice that at the annual frequency that term tends to one as α tends to zero. So our statistic rejects the limit of a standard New Keynesian model with constant returns to scale and full flexible-price pass-throughs unless prices are very rigid at an annual frequency, a possibility that we will reject in the following subsection.

In the online Appendix, we include a battery of robustness exercises and extensions. In particular, we include additional fixed effects controlling for geographical shocks, excluding firm-product pairs that experience shocks larger than a threshold, using different weighting approaches, and aggregating our data at the firm level as opposed to the firm × product level. The results are all consistent.

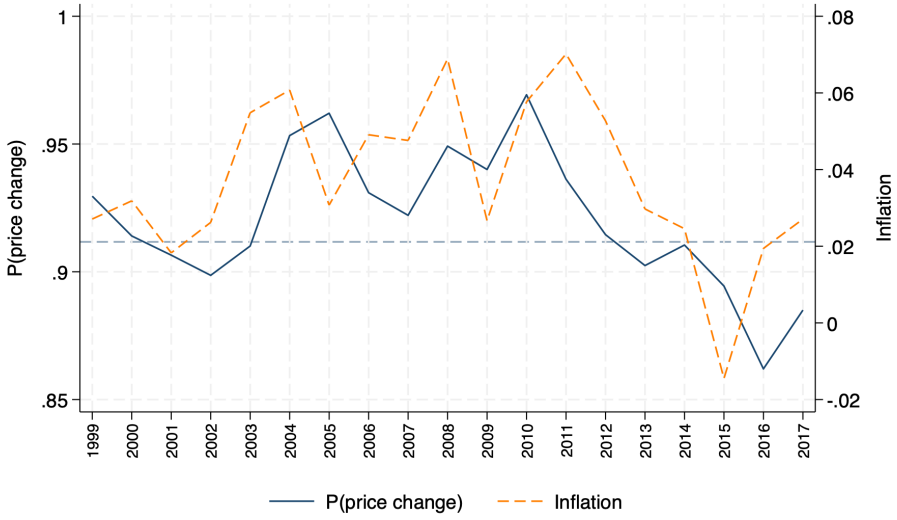
5.2. The frequency of price changes

We estimate the frequency of price changes from the price data directly. We follow standard practice to measure the frequency of price changes and measure an average frequency of price changes at the annual frequency of 0.91, consistent with a frequency of price changes of 0.45 per quarter, or 0.18 per month. These numbers directly map into a coefficient α of 0.55 at the quarterly frequency, or 0.09 at the annual frequency, and are similar to estimates in the literature for Mexico (Gagnon 2009), United States (Nakamura et al. 2018), and Argentina (Alvarez et al. 2019), when those economies faced inflation

rates similar to those of India. Notice that estimates for α directly map into the first term of equation (13).

Figure 3 shows the time-series of our measure of price rigidity. On top of marking the average frequency of price changes in a dotted line, the figure validates our measure by showing that the frequency comoves with the official Wholesale Price Index (WPI) inflation rate in India. This is true despite the two series coming from entirely different sources (the WPI is not based on micro-data from the ASI but on an independent data collection). For simplicity our model abstracts from movements in the frequency of price changes, so we use that variation as a validation of our approach to construct the frequency of price changes, and will use its average value throughout.

FIGURE 3. Frequency of price changes and the inflation rate in India



Note: This figure illustrate the time series behavior of the frequency of price changes along with the behavior of the wholesale price index (WPI) inflation in India.

5.3. The elasticity of marginal costs to output changes

As we shown in Section 3, firm marginal costs are given (in log-deviations) by:

$$\hat{mc}_{it} = \frac{1-a}{a}\hat{y}_{it} + \left(\hat{mc}_t - \frac{1-a}{a}\hat{Y}_t\right) \tag{18}$$

Equation (18) suggests a research design for estimating $\frac{1-a}{a}$ by using a regression of the log marginal costs on the log of quantities. This same equation also suggests what is

the main identification concern. Firm-level supply shifters that are correlated with firm-level quantities will induce a correlation between marginal costs and quantities that are not informative about the slope of the marginal supply curve.

To address this identification concern we leverage the fact that firms in our sample are multiproduct firms with heterogeneous exposure to their product portfolio. As a result, changes in the demand for specific products will induce firm-level demand shocks as a function of their pre-existing product shares.

Formally, we create a firm-level demand shifter of the following form:

$$\mathcal{Z}_{it}^{\xi} = \sum_{j} s_{ijt-1} \Delta \log \mathcal{Q}_{jt}, \tag{19}$$

where $\mathcal{Q}_{jt}$ is our measure of product-level quantities, and s_{ijt-1} is the one-lag share of production of product j in the portfolio of products of firm i. Our identifying assumption is that firms that are differentially exposed to more demand-sensitive products are not differentially exposed to firm-level supply shocks.

Our measure of demand shifters can be easily incorporated in our theoretical framework by extending our demand curves to allow for idiosyncratic demand shifters, as shown in Appendix B.1.

We implement this regression in the data in the following way. We estimate regressions of log changes in marginal costs on log changes in quantities sold at the firm level after controlling for industry $\times$ time fixed effects in order to control for unobserved supply and demand industry shocks and estimate the elasticity by comparing the behavior of firms in the same industry that have differential loadings across products. Formally, we estimate the regression equation

$$\Delta \log mc_{it} = \Phi_{s(i),t} + \beta_{mc,q} \Delta \log q_{it} + \epsilon_{i,t}, \tag{20}$$

and instrument $\Delta \log q_{it}$ using $\mathcal{Z}_{it}^{\xi}$. $\Phi_{s(i),t}$ is a set of industry$\times$time fixed effects.

Table 2 presents our estimates. Our benchmark estimate of 0.2 implies that when quantities increase by 1%, marginal costs increase by 0.2%. Since this coefficient maps to the structural object $\frac{1-a}{a}$, then our results imply a short-run firm-level returns to scale parameter $a = 0.84$. Our inference for a is very stable regardless of whether we use as a dependent

variable a measure of changes in marginal or total costs (column 2 versus column 1). We obtain slightly lower returns to scale when we exclude labor from our cost variables (column 3 versus column 1), consistent with theory. The F-statistic for the first stage is large and stable across specifications.

TABLE 2. The partial equilibrium elasticity of marginal costs to changes in quantities

	Baseline		Excl. labor	
	Δ ln mc	Δ ln C	Δ ln mc	Δ ln C
Δ log q	0.184***	1.106***	0.253***	1.192***
	(0.062)	(0.062)	(0.066)	(0.067)
Year $\times$ Ind. FE	✓	✓	✓	✓
N	258,238	258,238	258,237	258,237
F-Stat	247.18	247.18	247.18	247.18
Returns to scale	0.84	0.90	0.80	0.84

Note: This table presents our estimates to equation 20. The first two columns show results including labor when the dependent variable is our measure of marginal costs in column 1, or changes in total costs in column 2. Columns 3 and 4 are analogous when we exclude labor from our cost variables. The last row in the Table presents our estimates for the returns to scale parameter a.

It is important to note that our estimate $\hat{\beta}_{mc,q}$ is a partial equilibrium short-run supply elasticity that keeps constant prices and quantities at higher levels of aggregation. In particular, as we derived in Section 3, our model implies that Ω, the short-run elasticity of aggregate marginal costs to aggregate quantity, is given by

$$\Omega = \left(\gamma + \frac{\psi + 1 - a}{a}\right), \tag{21}$$

which makes clear that the reaction of input prices, crucially the real wage, to higher demand for inputs will induce a higher elasticity of marginal costs to quantities in general equilibrium than in partial equilibrium. Under the assumption that the equilibrium price adjustment takes place either at the industry or the regional level, we can replicate our research design at a higher level of aggregation.

That we can estimate Ω directly in our data, independent of the specification of other sub-blocks of our model is one of the main advantages of our research design, and is an improvement over variants of the standard approach of inferring Ω from a moment-matching exercise in which an econometrician infers the value of the structural parameters that pin down Ω by minimizing the distance of the impulse response functions to an ag-

TABLE 3. General equilibrium elasticity of marginal costs to changes in quantities

	District		Industry	
	Δ ln mc	Δ ln C	Δ ln mc	Δ ln C
Δ ln q	0.583***	1.487***	0.703**	1.704***
	(0.144)	(0.140)	(0.310)	(0.311)
Year FE	✓	✓	✓	✓
F Stat	103.08	103.08	26.83	26.83
N	7,707	7,707	1,211	1,211

Note: This table presents our estimates to equation 20, when aggregated to the industry or regional level. The first two columns show results when we aggregate firms at the district level, and the final two columns present the results when we aggregate the data at the industry level instead. Estimates in columns two and four use as a dependent variable the change in total costs as opposed to changes in our measure of marginal costs.

gregate shock, usually a monetary policy shock, in the data and those implied by a model. Notably, that approach is subject to criticisms of mispecification in every block of the structural model that influences the whole shape of an impulse response function. By inferring Ω from the data, our approach is less subject to these concerns.

Estimates in Table 3 imply that after an increase in 1% in quantities, marginal costs increase from 0.6% to 0.7% depending on the source of variation, directly mapping to our estimate of Ω, the middle term in equation (13). If a component of the elasticity of marginal costs to quantities is only operative at the national level, then our estimate will omit this component.

5.4. Estimates of the slope of the Phillips curve and comparison to the literature

Now that we have estimated $1/\alpha$, $(1-\alpha)(1-\beta\alpha)\mathbb{E}_\lambda[\zeta_i\rho_i]$, and Ω, we can assemble our estimate for κ_Y, the elasticity of inflation with respect to changes in aggregate quantities.

The Phillips curve in India is steeper than in developed countries. Our estimate for κ_Y is equal to 0.066 at the quarterly frequency, which is one order of magnitude larger than the estimation from Hazell et al. (2022), and six times larger than the estimates for the same object from Gagliardone et al. (2023). To understand that difference it is useful to decompose κ^Y in two components. First, the elasticity of marginal costs into inflation $\kappa^{mc} = \frac{(1-\alpha)(1-\beta\alpha)}{\alpha}\mathbb{E}_\lambda[\zeta_i\rho_i]$, and second, the elasticity of marginal costs to quantities Ω.

The elasticity of inflation to marginal costs κ_{mc} is determined by two distinct economic forces. The first one is the presence of price rigidity, and the second is the extent of micro-

level real rigidities induced by both imperfect pass-through and decreasing returns to scale in production. Our estimates imply $\kappa_{mc} = 0.095$, which is almost twice as large as that documented by Gagliardone et al. (2023). This difference can be fully rationalized by differences in the frequency of price change: if India had the same frequency of price change as developed economies, the slope of the marginal cost based Phillips curve would be similar to that in Gagliardone et al. (2023).

Most of the difference comes from the substantially higher quantity elasticity of marginal costs in India Ω. Our estimates of roughly 0.7 are three times larger than the estimates by Boehm and Pandalai-Nayar (2022) and Gagliardone et al. (2023), and almost four times larger than those in Shea (1993).

6. Allocative efficiency

We turn to the estimation of the sensitivity of the shifters D_t and Z_t with respects to aggregate demand shocks, which we named κ_D and κ_Z, respectively in equations (11), repeated below for convenience:

$$\kappa_D = \mathbb{E}_\lambda[\zeta_i \rho_i] \left(\mathbb{E}_\lambda \left[\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]} \frac{\zeta_i \rho_i}{\mathbb{E}_\lambda[\zeta_i \rho_i]} \right] - 1 \right) \tag{22}$$

$$\kappa_Z = \mathbb{E}_\lambda[\theta_i \zeta_i \rho_i] \left(\mathbb{E}_\lambda \left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]} \frac{\theta_i \zeta_i \rho_i}{\mathbb{E}_\lambda[\theta_i \zeta_i \rho_i]} \right] - \mathbb{E}_\lambda \left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]} \frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]} \right] \right) \tag{23}$$

Identification of κ_D. We estimate (22) as follows. Following the results on the identification of the slope of the Phillips curve, $\mathbb{E}_\lambda[\zeta_i \rho_i]$ is obtained from the regression of $\Delta \log p_{it}$ on $\Delta \log w_{it}^v$, instrumented by $\mathcal{Z}_{it}^\vartheta$, and corrected for the factor $(1-\alpha)(1-\beta\alpha)$:

$$\mathbb{E}_\lambda[\zeta_i \rho_i] = \frac{\beta_{p,w}^{IV}}{(1-\alpha)(1-\beta\alpha)}. \tag{24}$$

Note that estimating $\mathbb{E}_\lambda[\zeta_i \rho_i]$ requires more structure than our previous exercises, since we must take a stance on the extent of forward-looking behavior on the part of firms β, on top on our already imposed structure on the technology of price changes.

Looking at $\mathbb{E}_\lambda \left[\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]} \frac{\zeta_i \rho_i}{\mathbb{E}_\lambda[\zeta_i \rho_i]} \right]$, it is the expectation of the product of the relative price passthrough $\frac{\zeta_i \rho_i}{\mathbb{E}_\lambda[\zeta_i \rho_i]}$ and the relative demand elasticity $\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}$. We use the law of iterated expectations and estimate the price passthrough regression by bins of steady-state relative

demand elasticity.

To obtain demand elasticities, we estimate markups and use the assumption that on average over the whole sample, markups will be equal to desired flexible price markups, which we can invert to obtain demand elasticities. We estimate markups using the production approach, with materials as the flexible input. We rely on the elasticity of marginal costs with respect to quantities estimated above, which identifies the output elasticity of materials, so that our estimation does not suffer from the concern raised by Bond et al. (2021). Our estimation requires that any input wedge on materials is priced. We believe this assumption to be plausible; in particular, Singer (2019) documents that a large fraction of material inputs misallocation in India can be attributed to transportation costs that are reflected in prices recorded in the ASI. We describe the procedure in full details in Appendix B.3. Figure B.1 summarizes the markups and demand elasticities estimated in this way. We find that demand elasticities are decreasing in firm size, consistent with a positive superelasticity of demand.

Identification of κ_Z. $\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]$ is obtained from the regression of $\Delta \log y_{it}$ on $\Delta \log w^v_{it}$, instrumented by $\mathcal{Z}^\vartheta_{it}$, and corrected for the factor $(1-\alpha)(1-\beta\alpha)$. Since an initially exogenous marginal cost shock creates a price adjustment of $\zeta_i\rho_i$, where in general $\zeta_i \neq 1$ under decreasing returns to scale, then this marginal cost adjustment will trigger a $\theta_i\zeta_i\rho_i$ change in relative quantities as consumers substitute more expensive varieties along their demand curves. Therefore, the term $\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]$ is just the sales-weighted average quantity pass-through of a marginal cost shock (divided by $(1-\alpha)(1-\beta\alpha)$).

The research design to estimate this expectation is the same than for estimating price pass-throughs, and only requires substituting the outcome variable for a change in quantities. Formally, using the shape of the demand curves and the equation for reset prices, we find an equation for the change in quantities after an idiosyncratic cost shock. Then, a regression of Δy_{it} on Δw^v_{it} instrumented by the exogenous cost shifter ϑ_{it}, yields in population an IV estimate equal to:

$$\beta^{IV}_{q,w} = (1-\alpha)(1-\beta\alpha)\mathbb{E}_\lambda\left[\theta_i\rho_i\zeta_i\right], \tag{25}$$

$\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i\zeta_i\rho_i}{\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]}\right]$ is the expectation of the product of the relative quantity passthrough

$\frac{\theta_i \zeta_i \rho_i}{\mathbb{E}_\lambda[\theta_i \zeta_i \rho_i]}$ and the relative combined distortion $\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}$. We use the law of iterated expectations and estimate the quantity passthrough regression by bins of steady-state relative distortion.

This requires that we identify firm-level combined distortions m_i^{-1}. We show that m_i^{-1} is proportional to revenue productivity (TFPR). This is the same result as in Hsieh and Klenow (2009) where TFPR captures both input and output wedges. TFPR can be readily estimated in the data by combining estimates of the marginal revenue products of different inputs, which with the Cobb-Douglas assumption are just equal to revenues divided by input quantity. We describe the procedure in full details in Appendix B.3. Figure B.2 shows the obtained marginal revenue products for labor, capital, intermediates, and the resulting TFPR, by deciles of firm market shares.

Finally, $\mathbb{E}_\lambda \left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]} \frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]} \right]$ can be readily estimated from our estimates of distortions and demand elasticities.

Results. We find that $\kappa_D = .00035$. $\kappa_D > 0$ reflects the fact that firms with larger demand elasticities tend to have a larger price passthrough $\rho_i \zeta_i$: firms in the fourth quartile of θ_i (the smallest firms) have a passthrough roughly 1.5 larger compared to firms in the first quartile of θ_i (the largest firms). Lower passthrough for larger firms is consistent with the existing evidence in Amiti, Itskhoki, and Konings (2019).

Regarding Z_t, we find that $\kappa_Z = -.02791$. The key reason is that the inverse distortion $\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}$ is positively correlated with demand elasticities $\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}$. $\kappa_Z < 0$ imply that demand shocks endogenously reallocate quantities towards less distorted firms, reducing allocative efficiency.

7. Monetary non-neutrality in a large developing country

In this section we illustrate the implications of our estimates for the slope of the Phillips curve and our estimated allocative efficiency effects for the behavior of aggregate inflation, and the sensitivity of output to monetary policy interventions.

Fit of the Aggregate Phillips Curve. In our first exercise we use the aggregate measure for the output gap in the manufacturing sector in India that we introduced in section 4 and

underlies Figure C.1, and our estimate for κ^Y to check the time series fit of manufacturing inflation to the realized values of output in the data. The objective of the exercise is to provide an estimate of the importance of local quantity fluctuations in driving inflation holding constant both the effects of exogenous inflation shifters, as well as the endogenous allocative efficiency shifters induced by changes in aggregate demand.

Formally, we fit the following relation

$$\pi_t = \kappa^Y \hat{\omega} \tilde{y}_t, \tag{26}$$

where $\sum_{j=0}^{5} \beta^j \tilde{y}_{t+j} = a_0 + \omega \tilde{y}_t + e_t$. This approach is the same advocated by Hazell et al. (2022), in which they iterate the Phillips curve and estimate the elasticity of inflation with respect to the present value of future expected output gaps.[1]

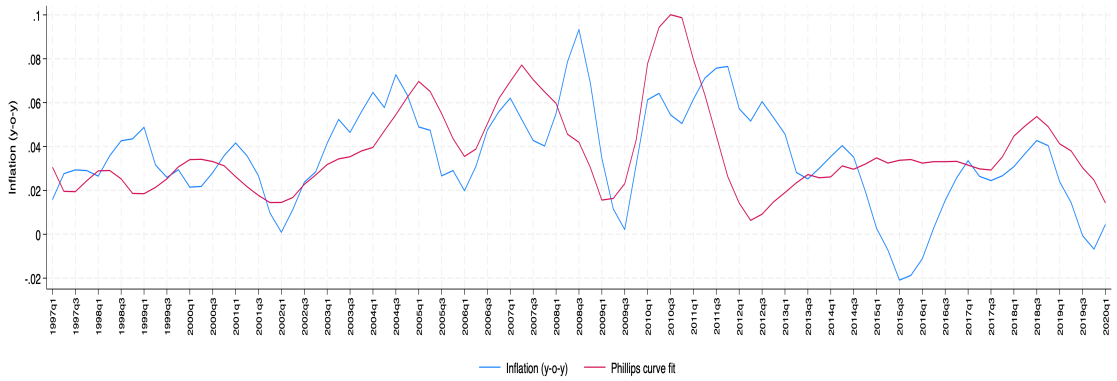


FIGURE 4. Fit of the Aggregate Phillips Curve

Note: This Figure takes the measure of the output gap we introduced in Section 2, and fits the Phillips curve as explained in equation 26 in the main text.

The main result behind Figure 4 is that the volatility of the output gap, plus our estimated slope of the Phillips curve, do a great job in explaining the dynamics of inflation in India during the period we analyze. This result is not guaranteed by construction. We do not use the official WPI quantity or price indexes in our estimate of κ^Y directly.

Conceptually, Figure 4 shows that local changes in quantities are an important driver of local prices, which is different from suggestions that inflation in developing countries

[1]For simplicity, this exercise assumes that the variation in the transitory component of output corresponds to the output gap, and not to transitory changes in the natural rate of output.

is driven exclusively by cost shocks generated in the rest of the world driving inflation in local economies. The two periods for which the fit of the Phillips curve is not good are the two periods we highlighted in figure C.1, in which changes in long-run expectations and time-varying cost shocks affected the dynamics of inflation. In contexts like those, it is natural to expect an imperfect fit.

Impulse response functions implied by our estimates. Our next exercise takes the five-equations New Keynesian model we presented in 3, and simulate the effects of an innovation to the monetary policy shock ε_t^{MP}.

We allow for monetary policy shocks to be persistent. Specifically, we allow for monetary disturbances to follow and AR(1) process with persistence of 0.7 at the quarterly frequency to match the persistence of the quantitative exercises of Baqaee, Farhi, and Sangani (2024).

Our calibration of the model is standard. Given our estimates for α, we infer the value for φ after making an assumption about the discount factor to target a 4% annual interest rate in the flexible price steady state. We calibrate $\sigma = 3$, and then use the expressions and estimated values for Ω under separable preferences to back out a value for ψ. The decreasing returns to scale parameter a, as well as for κ_Z and κ_D, follow our estimation in the main text. Our assumption of separable preferences implies that $\iota_{cl} = 0$. We assume standard values for the monetary policy rule of $\phi_\pi = 1.5$, and $\phi_y = 0.5$.

The purpose of this exercise is to evaluate the effects of the endogenous shifters due to allocative efficiency D and Z on inflation and output, as well as an empirical assessment of the importance of our measure for κ^Y and its components in driving inflation dynamics in India compared to a developed economy.

Figure 5 shows two series, one for our benchmark calibration, and another in which both κ_Z and κ_D are set to zero, effectively shutting down the allocative efficiency effects in the model. That you only see one series for the output and inflation IRFs means that the allocative effects on inflation and output in our benchmark calibration are too small to be detected. These effects are the results of the small κ_Z and κ_D we estimate.

That the allocative effects we estimate are small means that the assumption in which shifts of the aggregate demand curve trace, as opposed to trace and shift, the aggregate supply curve in India is not a bad assumption.

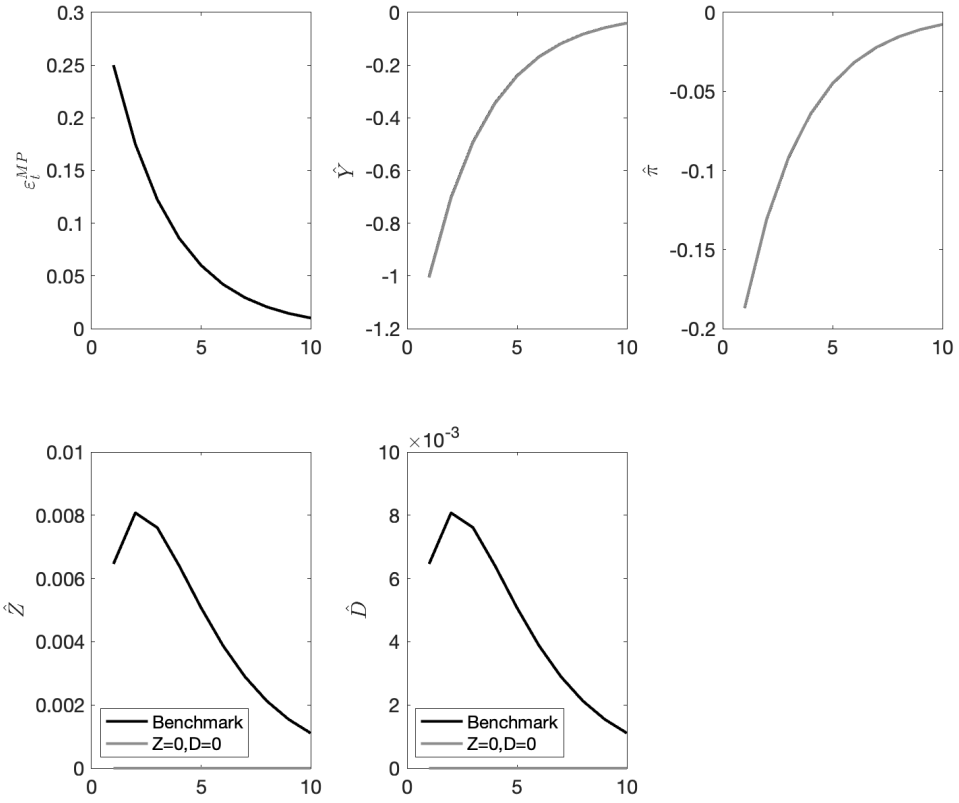


FIGURE 5. Impulse Response Functions after a Monetary Policy Shock

Note: This Figure presents the Impulse Response Function (IRF) in the five-equation New Keynesian model. Refer to the main text for details on the parameterization of the model.

Next, we illustrate the quantitative magnitude on the effects on inflation and output of the difference in the slope of the Phillips curve we estimate versus two leading studies on the slope of the Phillips curve in developed nations, Hazell et al. (2022) for the United States, and Gagliardone et al. (2023) for Belgium.

The main result of Figure 6 is that the extent of monetary non-neutrality of the shocks we consider in India are roughly have of those in developed nations, highlighting the importance of the differences in slopes we estimate. The flip side of reduced monetary non-neutrality is an increase in the elasticity of inflation with respect to a monetary policy shock. Our model predicts that inflation in India reacts substantially by more to a change in nominal interest rates than in developing nations.

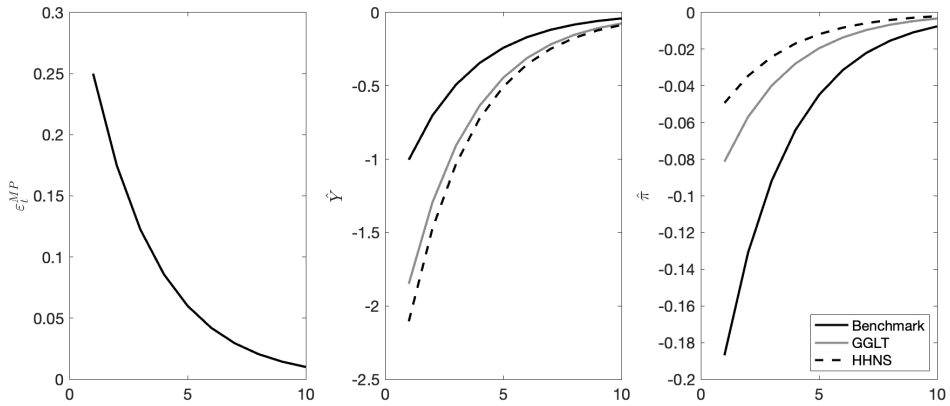


FIGURE 6. Impulse Response Functions after a Monetary Policy Shock

Note: This Figure presents the Impulse Response Function (IRF) in the five-equation New Keynesian model, and in three equation New Keynesian models calibrated to the slopes estimated in the United States and Belgium.

8. Conclusion

We have presented a method to estimate the inflationary effects of an aggregate demand expansion in a distorted economy. Our method is free of indirect inference blocks that use information from the shape of the aggregate demand curve or the monetary policy rule, and does not require to take a strong stance on the particular microfoundations that give rise to residual demand curves, or the structure of the markets for inputs as long as those microfoundations respect a decomposition of the slope of the Phillips curve into three multiplicative factors: the elasticity of prices to marginal costs, the elasticity of marginal costs to changes in quantities, and the frequency of price changes. In distorted economies, aggregate demand expansions have the potential to reallocate production across firms with different markups and across firms with different input wedges. As a result, the covariance of demand elasticities with firm sizes are not sufficient statistics to characterize the allocative efficiency of demand expansions. Instead, the relevant covariances are those of firm-level pass-throughs with respect to firm TFPRs.

We apply to our methodology to India, an economy with vast dispersion in input wedges, and find that the slope of the Phillips curve holding constant the allocation of resources is one order of magnitude larger than in developed countries such as Belgium or the United States. Around 75% of the variation in slopes across these countries is driven by differences in the elasticity of marginal costs with respect to quantities. The remaining

25% is driven by variation in the frequency of price changes across countries. The extent of micro real rigidities is remarkably similar across countries.

References

Alvarez, Fernando, Martin Beraja, Martin Gonzalez-Rozada, and Pablo Andrés Neumeyer. 2019. "From hyperinflation to stable prices: Argentina's evidence on menu cost models." *The Quarterly Journal of Economics* 134 (1): 451–505.

Amiti, Mary, Oleg Itskhoki, and Jozef Konings. 2019. "International shocks, variable markups, and domestic prices." *The Review of Economic Studies* 86 (6): 2356–2402.

Ball, Laurence, Anusha Chari, and Prachi Mishra. 2016. "Understanding inflation in India.", National Bureau of Economic Research.

Baqaee, David R, Emmanuel Farhi, and Kunal Sangani. 2024. "The supply-side effects of monetary policy." *Journal of Political Economy* 132 (4): 1065–1112.

Bau, Natalie, and Adrien Matray. 2023. "Misallocation and capital market integration: Evidence from India." *Econometrica* 91 (1): 67–106.

Beraja, Martin, Erik Hurst, and Juan Ospina. 2019. "The aggregate implications of regional business cycles." *Econometrica* 87 (6): 1789–1833.

Boehm, Christoph E, and Nitya Pandalai-Nayar. 2022. "Convex supply curves." *American Economic Review* 112 (12): 3941–3969.

Boehm, Johannes, Swati Dhingra, and John Morrow. 2022. "The Comparative Advantage of Firms." *Journal of Political Economy* 130 (12): 3025–3100. https://doi.org/10.1086/720630.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. 2021. "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data." *Journal of Monetary Economics* 121: 1–14.

Bresnahan, Timothy F, and Valerie A Ramey. 1994. "Output fluctuations at the plant level." *The Quarterly Journal of Economics* 109 (3): 593–624.

Calvo, Guillermo A. 1983. "Staggered prices in a utility-maximizing framework." *Journal of monetary Economics* 12 (3): 383–398.

Cerrato, Andrea, and Giulia Gitti. 2022. "Inflation since covid: Demand or supply."

Drenik, Andres, and Diego J Perez. 2020. "Price setting under uncertainty about inflation." *Journal of Monetary Economics* 116: 23–38.

Filardo, Andrew J, and Marco J Lombardi. 2014. "Has Asian emerging market monetary policy been too procyclical when responding to swings in commodity prices?" *BIS Paper* (77j).

Gagliardone, Luca, Mark Gertler, Simone Lenzu, and Joris Tielens. 2023. "Anatomy of the Phillips Curve: Micro Evidence and Macro Implications."

Gagnon, Etienne. 2009. "Price setting during low and high inflation: Evidence from Mexico." *The Quarterly Journal of Economics* 124 (3): 1221–1263.

Galı, Jordi, and Mark Gertler. 1999. "Inflation dynamics: A structural econometric analysis." *Journal of monetary Economics* 44 (2): 195–222.

Gopinath, Gita, and Oleg Itskhoki. 2011. "In search of real rigidities." *NBER Macroeconomics Annual* 25 (1): 261–310.

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

Gopinath, Gita, and Roberto Rigobon. 2008. "Sticky borders." *The Quarterly Journal of Economics* 123

(2): 531–575.

Hamilton, James D. 2018. "Why you should never use the Hodrick-Prescott filter." *Review of Economics and Statistics* 100 (5): 831–843.

Hashemi, Arshia, Ivan Kirov, and James Traina. 2022. "The production approach to markup estimation often measures input distortions." *Economics Letters* 217: 110673.

Hazell, Jonathon, Juan Herreño, Emi Nakamura, and Jón Steinsson. 2022. "The slope of the Phillips Curve: evidence from US states." *The Quarterly Journal of Economics* 137 (3): 1299–1344.

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

Kimball, Miles S. 1995. "The Quantitative Analytics of the Basic Neomonetarist Model." *Journal of Money, Credit and Banking* 27 (4): 1241–1277. http://www.jstor.org/stable/2078048.

Martin, Leslie A., Shanthi Nataraj, and Ann E. Harrison. 2017. "In with the Big, Out with the Small: Removing Small-Scale Reservations in India." *American Economic Review* 107 (2): 354–86. https://www.aeaweb.org/articles?id=10.1257/aer.20141335.

Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H Stock. 2014. "Empirical evidence on inflation expectations in the New Keynesian Phillips Curve." *American Economic Journal: Journal of Economic Literature* 52 (1): 124–188.

Mohanty, MS, Marc Klau et al. 2001. "What determines inflation in emerging market economies?" *BIS Papers* 8: 1–38.

Mongey, Simon. 2021. "Market structure and monetary non-neutrality.", National Bureau of Economic Research.

Nakamura, Emi, and Jón Steinsson. 2008. "Five facts about prices: A reevaluation of menu cost models." *The Quarterly Journal of Economics* 123 (4): 1415–1464.

Nakamura, Emi, Jón Steinsson, Patrick Sun, and Daniel Villar. 2018. "The elusive costs of inflation: Price dispersion during the US great inflation." *The Quarterly Journal of Economics* 133 (4): 1933–1980.

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

Reinelt, Timo, and Matthias Meier. 2020. "Monetary policy, markup dispersion, and aggregate TFP."

Rijesh, R. 2022. "Concordance of Manufacturing Industries Using NIC-2008 and NIC-2004: A Disaggregate Level Study.". https://doi.org/10.21203/rs.3.rs-2055633/v1.

Shea, John. 1993. "Do supply curves slope up?" *The Quarterly Journal of Economics* 108 (1): 1–32.

Singer, Gregor. 2019. "Endogenous markups, input misallocation and geographical supplier access."."

Stock, James H, and Mark W Watson. 2020. "Slack and cyclically sensitive inflation." *Journal of Money, Credit and Banking* 52 (S2): 393–428.

Appendix for online publication

Appendix A. Model derivations

A.1. Environment

Households. Households choose consumption C and labor L to maximize discounted future utility $\mathbb{E}_0 \sum_{t=0}^{+\infty} \beta^t u(C_t, L_t)$ subject to a per-period budget constraint $P_t^Y C_t + Q_t B_t = B_{t-1} + w_t L_t + T_t$ where P_t^Y is the price of the consumption bundle, B_t is holdings of one-period risk-free nominal bonds with price Q_t, w_t is the wage, and T_t denotes any profits rebated to households as lump-sum. From the households' optimization problem we obtain the Euler equation:

$$\frac{1}{1+i_t} = \beta \mathbb{E}_t[\frac{u_c(C_{t+1}, L_{t+1})}{u_c(C_t, L_t)} \frac{P_t^Y}{P_{t+1}^Y}] \tag{A.1}$$

and the labor supply function:

$$\frac{u_l(C_t, L_t)}{u_c(C_t, L_t)} = \frac{w_t^l}{P_t^Y} \tag{A.2}$$

Final good producers. Let Y_t denote aggregate production of the final good. This can be used for consumption C_t or as an input to produce materials Y_t^X so that $Y_t = C_t + Y_t^X$.

The final good Y_t is produced by perfectly competitive firms using a bundle of differentiated intermediate inputs y_{it} for $i \in [0, 1]$. Intermediate input varieties are assembled into the final good using the Kimball aggregator:

$$\int_0^1 \Upsilon\left(\frac{y_{it}}{Y_t}\right) di = 1$$

where the function $\Upsilon(.)$ is strictly increasing, strictly concave, and satisfies $\Upsilon(1) = 1$. The CES aggregator is the special case $\Upsilon(q) = q^{\frac{\theta-1}{\theta}}$ for $\theta > 1$.

Taking the prices p_{it} of the inputs as given and denoting the price of the final good P_t^Y, final good producers choose y_{it} to maximize profits. This gives rise to the demand

function:

$$\frac{y_{it}}{Y_t} = \Upsilon'^{-1}\left(\frac{p_{it}}{\mathcal{P}_t}\right) \tag{A.3}$$

where $\frac{p_{it}}{\mathcal{P}_t}$ determines substitution across varieties. The price index $\mathcal{P}_t$ is given by $\mathcal{P}_t = \frac{P_t^Y}{D_t}$. $P_t^Y = \int_0^1 p_{it}\frac{y_{it}}{Y_t}di$ is the ideal price index. $D_t = \int_0^1 \Upsilon'(\frac{y_{it}}{Y_t})\frac{y_{it}}{Y_t}di$ is a "demand" index. When demand is CES, D_t is a constant equal to $\frac{\theta}{\theta-1}$. Away from the CES case, D_t is not a constant and is increasing in the dispersion of quantity shares.

The price elasticity of demand is only a function of firm relative size and is given by:

$$\theta_{it} = \theta(\frac{y_{it}}{Y_t}) = -\frac{\partial \log y_{it}}{\partial \log p_{it}} = \frac{\Upsilon'\left(\frac{y_{it}}{Y_t}\right)}{-\frac{y_{it}}{Y_t}\Upsilon''\left(\frac{y_{it}}{Y_t}\right)} \tag{A.4}$$

Differentiated varieties producers. Each variety i is produced by a single firm. Firms produce with the following technology:

$$y_{it} = e^{z_{it}}v_{it}^a \quad \text{with} \quad v_{it} = l_{it}^{\phi}x_{it}^{1-\phi} \tag{A.5}$$

v_{it} is a constant returns-to-scale composite of the variable inputs, labor l_{it} and materials x_{it}. The unit cost of inputs are w_t^l and w_t^x, respectively.

We first consider the firm's cost-minimization problem, taking output and input prices as given.

$$\min_{l_{it},x_{it}}(1+\tau_i)[w_t^x x_{it} + w_t^l l_{it}] \quad \text{subject to} \quad e^{z_{it}}(l_{it}^{\phi}x_{it}^{1-\phi})^a \geq \bar{y} \tag{A.6}$$

We obtain that for any v_{it}, the choice of labor and material will be given by:

$$l_{it} = \left(\frac{w_t^x}{w_t^l}\frac{\phi}{1-\phi}\right)^{1-\phi} v_{it} \tag{A.7}$$

$$x_{it} = \left(\frac{w_t^x}{w_t^l}\frac{\phi}{1-\phi}\right)^{-\phi} v_{it} \tag{A.8}$$

This defines a price index for variable inputs:

$$w_{it}^v = (w_t^l/\phi)^{\phi}(w_t^x/(1-\phi))^{1-\phi} \tag{A.9}$$

The cost function writes:

$$\mathcal{C}_{it}(y_{it}, w_{it}^v) = (1+\tau_i)w_{it}^v\left(\frac{y_{it}}{e^{z_{it}}}\right)^{\frac{1}{a}} \tag{A.10}$$

The marginal cost for firm i writes:

$$mc_{it} = (1+\tau_i)w_{it}^v\frac{1}{a}e^{-\frac{1}{a}z_{it}}y_{it}^{\frac{1-a}{a}} \tag{A.11}$$

A firm has a probability $1-\alpha$ of being able to reset its price in each period. A firm that can reset its price maximizes choses the price that maximizes:

$$\max_{p_{it}} \mathbb{E}_t\left[\sum_{s=0}^{+\infty}\alpha^s\Lambda_{t,t+s}\left[p_{it}y_{it+s} - \mathcal{C}_{it+s}(y_{it+s}, w_{it+s}^v)\right]\right]$$

subject to the demand curve $y_{it+s} = \Upsilon'^{-1}\left(\frac{p_{it}}{\mathcal{P}_{t+s}}\right)Y_{t+s}$ and the cost function $\mathcal{C}_{it+s}(y_{it+s}, w_{t+s}^v) = (1+\tau_i)w_{t+s}^v e^{-\frac{1}{a}z_i}y_{it+s}^{\frac{1}{a}}$. $\Lambda_{t,t+s}$ is the stochastic discount factor.

It will be convenient to define the following quantities. $\mu_{it}^f = \frac{\theta_{it}}{\theta_{it}-1}$ is the desired markup that the firm would choose in a flexible price environment. $\Gamma_{it} = \frac{\partial \log \mu_{it}^f}{\partial \log \frac{y_{it}}{Y_t}}$ is the elasticity of the flexible price markup with respect to relative size. ρ_{it} is the partial equilibrium pass-through of a marginal cost shock into the firm's price.

$$d\log p_{it} = d\log \mu_{it}^f + d\log mc_{it} = \Gamma_{it}\, d\log\frac{y_{it}}{Y_t} + d\log mc_{it} = -\Gamma_{it}\theta_{it}\,(d\log p_{it} - d\log \mathcal{P}_t) + d\log mc_{it}$$

$$d\log p_{it} = \frac{1}{1+\Gamma_{it}\theta_{it}}d\log mc_{it} + \frac{\Gamma_{it}\theta_{it}}{1+\Gamma_{it}\theta_{it}}d\log \mathcal{P}_t$$

Therefore, $\rho_{it} \equiv \frac{\partial \log p_{it}}{\partial \log mc_{it}} = \frac{1}{1+\Gamma_{it}\theta_{it}}$. In the CES case, all firms face the same demand elasticities, $\Gamma_{it} = 0\ \forall i$, and $\rho_{it} = 1\ \forall i$. Away from the CES case, ρ_{it} can be below or above 1 depending on the sign of Γ_{it}.

Note that μ_{it}^f, Γ_{it}, and ρ_{it} are only a function of a firm's relative size $\frac{y_{it}}{Y_t}$.

Finally, in a sticky price environment, the actual markup of the firm may differ from

the flexible price desired markup. We denote the actual markup of the firm: $\mu_{it} = \frac{p_{it}}{mc_{it}}$.

Materials producers. Materials are produced by a perfectly competitive representative firm with the following production function: $X_t = (Y_t^X)^{\frac{\eta}{1+\eta}}$ where Y_t^X are units of the final good used for intermediate inputs production and $\eta > 0$. Solving

$$\max_{X_t} w_t^x X_t - P_t^Y (X_t)^{1+\eta^{-1}} \tag{A.12}$$

implies that the price of the material good is given by: $w_t^x = P_t^Y (1+\eta^{-1}) X_t^{\eta^{-1}}$.

Monetary authority. Monetary policy sets the nominal interest rate according to a Taylor rule.

Equilibrium. Equilibrium is defined by the following conditions: (i) Consumers choose consumption and labor to maximize utility taking prices and wage as given; (ii) Firms with flexible prices set prices to maximize their value taking the price index and their residual demand curves as given; firms with sticky prices meet demand at fixed prices ; (iii) Monetary policy sets the nominal interest rate; (iv) All resource constraints are satisfied.

We solve the model by log-linearization around the zero-inflation steady state. The steady-state distribution of firm size depends on joint distribution of (z_i, τ_i). We take a first-order expansion for small monetary policy shock. Quantities without a t subscript refer to the steady-state.

We denote $\lambda_i = \frac{p_i y_i}{P^Y Y}$ sales share in steady-state. Let $\mathbb{E}_\lambda[X_{it}] = \int_0^1 \lambda_i X_{it} di$.

A.2. Characterization

It is useful to the following expressions. First, linearizing the definition of the Kimball aggregator:

$$0 = \mathbb{E}_\lambda[\widehat{\frac{y_{it}}{Y_t}}] \quad \Leftrightarrow \quad \hat{Y}_t = \mathbb{E}_\lambda[\hat{y}_{it}] \tag{A.13}$$

Second, linearizing the price index:

$$\hat{P}_t^Y = \mathbb{E}_\lambda[\hat{p}_{it}] \tag{A.14}$$

Third, linearizing $\mathcal{P}_t$:

$$\hat{\mathcal{P}}_t = \hat{P}_t^Y - \hat{D}_t \tag{A.15}$$

Fourth, linearizing the demand index:

$$\hat{D}_t = -\mathbb{E}_\lambda[(\theta_i - 1)(\hat{p}_{it} - \hat{\mathcal{P}}_t)] = -\frac{\mathrm{Cov}_\lambda[\theta_i, \hat{p}_{it}]}{\mathbb{E}_\lambda[\theta_i]} \tag{A.16}$$

Finally, we can write:

$$\hat{\mathcal{P}}_t = \frac{\mathbb{E}_\lambda[\theta_i \hat{p}_{it}]}{\mathbb{E}_\lambda[\theta_i]} \tag{A.17}$$

Marginal cost-based Phillips curve. A firm that can reset its price at time t will choose:

$$\hat{p}_{it|t}^{new} = (1 - \beta\alpha)\mathbb{E}_t \left[\sum_{s=0}^{+\infty} (\beta\alpha)^s (\hat{\mu}_{it+s|t}^f + \hat{mc}_{it+s|t}) \right] \tag{A.18}$$

$\hat{\mu}_{it+s|t}^f$ is the optimal markup at $t+s$ of a firm that could last reset its price at t. It is given by:

$$\hat{\mu}_{it+s|t}^f = -\Gamma_i \theta_i (\hat{p}_{it|t}^{new} - \hat{\mathcal{P}}_{t+s}) \tag{A.19}$$

Let $\hat{mc}_t \equiv \mathbb{E}_\lambda[\hat{mc}_{it}]$ the change in the aggregate nominal marginal cost. We can write:

$$\hat{mc}_{it+s|t} - \hat{mc}_{t+s} = -\frac{1-a}{a} \theta_i (\hat{p}_{it|t}^{new} - \hat{\mathcal{P}}_{t+s}) \tag{A.20}$$

We obtain:

$$\hat{p}_{it|t}^{new} = (1 - \beta\alpha)\mathbb{E}_t \left[\sum_{s=0}^{+\infty} (\beta\alpha)^s \left(\zeta_i \rho_i \hat{mc}_{t+s} + (1 - \zeta_i \rho_i) \hat{\mathcal{P}}_{t+s} \right) \right] \tag{A.21}$$

where ρ_i is the flexible price partial equilibrium pass-through of marginal cost shocks into prices and $\zeta_i = \frac{1}{1+\frac{1-a}{a}\theta_i\rho_i}$. ζ_i captures the fact that when returns to scale are below 1, a cost shock induces an adjustment in size, which dampens the first-round effect on marginal cost. $\zeta_i\rho_i$ combines these two terms and captures the flexible price passthrough of an input cost shock into prices.

We can write this equation recursively as:

$$\hat{p}^{new}_{it|t} = (1-\beta\alpha)\left(\zeta_i\rho_i\hat{mc}_t + (1-\zeta_i\rho_i)\hat{\mathcal{P}}_t\right) + \beta\alpha\mathbb{E}_t[\hat{p}^{new}_{it+1|t+1}] \tag{A.22}$$

Inflation dynamics. The change in the price of the final good is given by:

$$\hat{P}^Y_t = \mathbb{E}_\lambda[\hat{p}_{it}] \tag{A.23}$$

Let $\mathbb{1}^p_{it}$ be a dummy equal to 1 if firm i can reset their price at t.

$$\hat{p}_{it} = (1-\mathbb{1}^p_{it})\,\hat{p}_{it-1} + \mathbb{1}^p_{it}\,\hat{p}^{new}_{it|t} \tag{A.24}$$

Therefore,

$$\hat{P}^Y_t = \mathbb{E}[(1-\mathbb{1}^p_{it})\lambda_i\hat{p}_{it-1}] + \mathbb{E}[\mathbb{1}^p_{it}\lambda_i\hat{p}^{new}_{it|t}] \tag{A.25}$$

The Calvo fairy is orthogonal to the firm's steady state sales share so that

$$\hat{P}^Y_t = \alpha\hat{P}^Y_{t-1} + (1-\alpha)\mathbb{E}_\lambda[\hat{p}^{new}_{it|t}] \tag{A.26}$$

By definition,

$$\hat{\pi}_t = \hat{P}^Y_t - \hat{P}^Y_{t-1} \tag{A.27}$$

$$\hat{\mathcal{P}}_t = \hat{P}^Y_t - \hat{D}_t \tag{A.28}$$

Aggregating across firms, we obtain the marginal cost-based Phillips curve:

$$\hat{\pi}_t = \varphi\mathbb{E}_\lambda[\zeta_i\rho_i](\hat{mc}_t - \hat{P}^Y_t) - \varphi(1-\mathbb{E}_\lambda[\zeta_i\rho_i])\hat{D}_t + \beta\mathbb{E}_t[\hat{\pi}_{t+1}] \tag{A.29}$$

$\hat{mc}_t - \hat{P}_t^Y$ is the change in the aggregate real marginal cost. $\hat{D}_t$ is the change in the demand index. $\varphi = \frac{(1-\alpha)(1-\beta\alpha)}{\alpha}$ is the slope of the marginal cost-based Phillips curve in the case of constant returns to scale and CES demand. $\mathbb{E}_\lambda[\zeta_i \rho_i]$ reflects micro-level real rigidities due to decreasing returns to scale ζ_i and strategic complementarities ρ_i.

Aggregate marginal costs. Let us define aggregate productivity Z_t as satisfying

$$Y_t \equiv Z_t (L_t^{\phi} X_t^{1-\phi})^a \tag{A.30}$$

From the cost-minimization problem of the producers of differentiated varieties, we obtain the log-linearized input demands:

$$\hat{l}_{it} = (1-\phi)(\hat{w}_t^x - \hat{w}_t^l) + \hat{v}_{it} \tag{A.31}$$

$$\hat{x}_{it} = -\phi(\hat{w}_t^x - \hat{w}_t^l) + \hat{v}_{it} \tag{A.32}$$

Using the first-order conditions of the cost-minimization, and aggregating across firms, we can derive aggregate input demand functions as:

$$\hat{L}_t = (1-\phi)(\hat{w}_t^x - \hat{w}_t^l) + \hat{V}_t = (1-\phi)\left(\hat{w}_t^x - \hat{w}_t^l\right) + \frac{1}{a}\left(\hat{Y}_t - \hat{Z}_t\right); \tag{A.33}$$

$$\hat{X}_t = -\phi(\hat{w}_t^x - \hat{w}_t^l) + \hat{V}_t = -\phi\left(\hat{w}_t^x - \hat{w}_t^l\right) + \frac{1}{a}\left(\hat{Y}_t - \hat{Z}_t\right). \tag{A.34}$$

The (log-linearized) labor supply function is derived from the utility maximization problem of the household:

$$\nu^{-1}\hat{L}_t + (\sigma^{-1} - \iota_{lc})\hat{C}_t = \hat{w}_t^l - \hat{P}_t^Y, \tag{A.35}$$

where $\sigma^{-1} = -\frac{u_{cc}C}{u_c}$ and $\iota_{lc} \equiv -\frac{u_{lc}C}{u_l}$. From the output maximization of the materials producer, we obtain the materials supply curve, again in log-linear form:

$$\eta^{-1}\hat{X}_t = \hat{w}_t^x - \hat{P}_t^Y \tag{A.36}$$

From the market clearing conditions, we obtain equilibrium prices for labor and materials as

$$\hat{w}_t^l = \frac{\nu^{-1}(1+\eta^{-1})}{1+\eta^{-1}\phi+\nu^{-1}(1-\phi)}\frac{1}{a}\left(\hat{Y}_t - \hat{Z}_t\right) + \frac{(\sigma^{-1}-\iota_{lc})(1+\eta^{-1}\phi)}{1+\eta^{-1}\phi+\nu^{-1}(1-\phi)}\hat{C}_t + \hat{P}_t^Y; \tag{A.37}$$

$$\hat{w}_t^x = \frac{\eta^{-1}(1+\nu^{-1})}{1+\eta^{-1}\phi+\nu^{-1}(1-\phi)}\frac{1}{a}(\hat{Y}_t - \hat{Z}_t) + \frac{\eta^{-1}\phi(\sigma^{-1}-\iota_{lc})}{1+\eta^{-1}\phi+\nu^{-1}(1-\phi)}\hat{C}_t + \hat{P}_t^Y. \tag{A.38}$$

Aggregate steady-state marginal costs are equal to

$$\hat{mc}_t = \mathbb{E}_\lambda\left[\hat{mc}_{it}\right] = \phi\hat{w}_t^l + (1-\phi)\hat{w}_t^x + \frac{1-a}{a}\hat{Y}_t. \tag{A.39}$$

From the goods market clearing condition we can derive

$$\hat{C}_t = \upsilon_0\hat{Y}_t - \upsilon_0(1-\lambda_c)(1+\psi)\frac{1}{a}(\hat{Y}_t - \hat{Z}_t), \tag{A.40}$$

where λ_c denotes the share of the final good that is used for consumption in steady-state and where we defined $\upsilon_0 \equiv \frac{1+\eta^{-1}\phi+\nu^{-1}(1-\phi)}{\lambda_c(1+\eta^{-1}\phi+\nu^{-1}(1-\phi))+(1-\lambda_c)(1+\eta^{-1})\phi(\sigma^{-1}-\iota_{lc})}$. Together with equilibrium prices from equations (A.37) and (A.38), we can then derive the following solution for aggregate marginal cost:

$$\hat{mc}_t = \underbrace{\left[\frac{1-a+\psi}{a} + \upsilon\left(1-(1-\lambda_c)\frac{(1+\psi)}{a}\right)\right]}_{\Omega = \text{Elasticity of mc wrt output}}\hat{Y}_t - \underbrace{\left[\frac{\psi}{a} - \upsilon(1-\lambda_c)(1+\psi)\right]}_{\Xi = \text{Elasticity of mc wrt TFP}}\hat{Z}_t + \hat{P}_t^Y. \tag{A.41}$$

$\sigma^{-1} = -\frac{u_{cc}C}{u_c}$ is the inverse elasticity of intertemporal substitution, $\nu^{-1} = \frac{u_{ll}L}{u_l} - \frac{u_{cl}L}{u_c}$ is the inverse Frisch elasticity of labor supply, and $\iota_{cl} = \frac{u_{cl}L}{u_c}$. $\psi = \frac{\phi\nu^{-1}+(1-\phi)\eta^{-1}+\nu^{-1}\eta^{-1}}{1+\nu^{-1}(1-\phi)+\phi\eta^{-1}}$ is the slope of the variable input supply curve, reflecting the slope of the labor supply curve ν^{-1} and the slope of the materials supply curve η^{-1}. We define $\upsilon \equiv \frac{(\sigma^{-1}-\iota_{lc})(1+\eta^{-1})\phi}{\lambda_c(1+\eta^{-1}\phi+\nu^{-1}(1-\phi))+(1-\lambda_c)(1+\eta^{-1})\phi(\sigma^{-1}-\iota_{lc})}$.

Output-based Phillips curve. Combining (B.15) and (A.41), we obtain the output-based New Keynesian Phillips curve:

$$\hat{\pi}_t = \varphi\mathbb{E}_\lambda[\zeta_i\rho_i]\left(\Omega\hat{Y}_t - \Xi\hat{Z}_t\right) - \varphi(1-\mathbb{E}_\lambda[\zeta_i\rho_i])\hat{D}_t + \beta\mathbb{E}_t[\hat{\pi}_{t+1}] \tag{A.42}$$

Allocative efficiency. Let us defined the combined allocative distortion as $m_{it} \equiv \mu_{it}(1 + \tau_i)$. Using the definition of the markup and of the marginal cost,

$$m_{it} = a\frac{p_{it}y_{it}}{w_t^v v_{it}} \tag{A.43}$$

We define the aggregate distortion as solving:

$$\mathcal{M}_t \equiv a\frac{P_t^Y Y_t}{w_t^v V_t} \tag{A.44}$$

Step 1: First, we show that:

$$\mathcal{M}_t = \mathbb{E}_\lambda\left[m_{it}^{-1}\right]^{-1} \tag{A.45}$$

Step 2:

$$\hat{Z}_t = a\left[\hat{\mathcal{M}}_t - \mathbb{E}_\lambda[\hat{m}_{it}]\right] \tag{A.46}$$

By definition of the aggregate distortion,

$$\hat{\mathcal{M}}_t = \widehat{P_t^Y Y_t} - \widehat{w_t^v V_t}$$

By definition of the firm level distortion

$$\hat{m_{it}} = \hat{\mu}_{it} = \hat{p}_{it} - \hat{mc}_{it} = \hat{p}_{it} - \hat{w}_t^v - \frac{1-a}{a}\hat{y}_{it}$$
$$\mathbb{E}_\lambda[\hat{m}_{it}] = \hat{P}_t^Y - \hat{w}_t^v - \frac{1-a}{a}\hat{Y}_t$$

Therefore,

$$a(\hat{\mathcal{M}}_t - \mathbb{E}_\lambda[\hat{m}_{it}]) = \hat{Y}_t - a\hat{V}_t = \hat{Z}_t$$

Step 3:

$$\hat{Z}_t = -\mathcal{M}\mathrm{Cov}_\lambda[m_i^{-1}, \hat{y}_{it}] = \mathcal{M}\mathrm{Cov}_\lambda[m_i^{-1}, \theta_i(\hat{p}_{it} - \hat{\mathcal{P}}_t)] \tag{A.47}$$

Log-linearizing the expression for $\mathcal{M}_t$ in (A.45) yields:

$$\hat{\mathcal{M}}_t - \mathbb{E}_\lambda[\hat{m}_{it}] = -\mathcal{M}\mathbb{E}_\lambda[(m_i)^{-1}(\hat{\lambda}_{it} - \hat{m}_{it})] - \mathbb{E}_\lambda[\hat{m}_{it}] = -\mathcal{M}\text{Cov}_\lambda[m_i^{-1}, \hat{\lambda}_{it} - \hat{m}_{it}]$$

In addition,

$$\hat{\lambda}_{it} - \hat{m}_{it} = (\hat{p}_{it} + \hat{y}_{it} - (\hat{P}_t^Y + \hat{Y}_t)) - (\hat{p}_{it} - \hat{mc}_{it}) = -(\hat{P}_t^Y + \hat{Y}_t) + \hat{w}_t^v + \frac{1}{a}\hat{y}_{it}$$

Therefore, $\hat{Z}_t = -\mathcal{M}\text{Cov}_\lambda[m_i^{-1}, \hat{y}_{it}]$.

Step 4: We have two equations that characterize $\hat{D}_t$ and $\hat{Z}_t$ as a function of the change in relative prices:

$$\hat{D}_t = -\frac{\text{Cov}_\lambda[\theta_i, \hat{p}_{it}]}{\mathbb{E}_\lambda[\theta_i]}$$
$$\hat{Z}_t = \mathcal{M}\text{Cov}_\lambda[m_i^{-1}, \theta_i(\hat{p}_{it} - \hat{\mathcal{P}}_t)]$$

where the first equation is (A.16) and the second equation is (A.47).

We know show how to express $\hat{D}_t$ and $\hat{Z}_t$ as a function of model parameters and steady-state objects.

$$\hat{p}_{it} = \mathbb{1}_{it}^p \hat{p}_{it|t}^{new} + (1 - \mathbb{1}_{it}^p)\hat{p}_{it-1}$$

Taking the expectation over the realization of the Calvo fairy, we obtain:

$$(\hat{p}_{it} - \hat{p}_{it-1}) - \beta(\hat{p}_{it+1} - \hat{p}_{it}) = \varphi\left(\zeta_i \rho_i \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) + \hat{\mathcal{P}}_t\right) - \varphi\hat{p}_{it} \tag{A.48}$$

Let us first derive the equation for $\hat{D}_t$. Multiplying by θ_i and applying the $\mathbb{E}_\lambda$ operator yields:

$$(\mathbb{E}_\lambda[\theta_i\hat{p}_{it}] - \mathbb{E}_\lambda[\theta_i\hat{p}_{it-1}]) - \beta(\mathbb{E}_\lambda[\theta_i\hat{p}_{it+1}] - \mathbb{E}_\lambda[\theta_i\hat{p}_{it}]) = \varphi\mathbb{E}_\lambda\left[\theta_i\left(\zeta_i\rho_i\left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) + \hat{\mathcal{P}}_t\right)\right] - \varphi\mathbb{E}_\lambda[\theta_i\hat{p}_{it}]$$
$$(\hat{\mathcal{P}}_t - \hat{\mathcal{P}}_{t-1}) - \beta(\hat{\mathcal{P}}_{t+1} - \hat{\mathcal{P}}_t) = \varphi\frac{\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]}{\mathbb{E}_\lambda[\theta_i]}\left(\hat{mc}_t - \hat{\mathcal{P}}_t\right)$$

where we repeatedly use equation (A.17). Substracting the NKPC equation,

$$[\hat{D}_t - \hat{D}_{t-1} - \beta(\hat{D}_{t+1} - \hat{D}_t)] = -\varphi \frac{\text{Cov}_\lambda[\theta_i, \zeta_i \rho_i]}{\mathbb{E}_\lambda[\theta_i]} \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) - \varphi \hat{D}_t$$

In addition,

$$\hat{mc}_t - \hat{\mathcal{P}}_t = \Omega \hat{Y}_t - \Xi \hat{Z}_t + \hat{D}_t$$

We can rewrite this as equation (11).

$$[\hat{D}_t - \hat{D}_{t-1} - \beta(\hat{D}_{t+1} - \hat{D}_t)] = -\varphi \kappa_D \left(\Omega \hat{Y}_t - \Xi \hat{Z}_t + \hat{D}_t\right) - \varphi \hat{D}_t$$
$$\hat{D}_t(1 + \beta + \varphi(1 + \kappa_D)) = -\varphi \kappa_D \left(\Omega \hat{Y}_t - \Xi \hat{Z}_t\right) + \hat{D}_{t-1} + \beta \hat{D}_{t+1}$$

We use the same logic to derive the expression for $\hat{Z}_t$.

$$(\hat{p}_{it} - \hat{p}_{it-1}) - \beta(\hat{p}_{it+1} - \hat{p}_{it}) - (\hat{\mathcal{P}}_t - \hat{\mathcal{P}}_{t-1}) - \beta(\hat{\mathcal{P}}_{t+1} - \hat{\mathcal{P}}_t) = \varphi \left(\zeta_i \rho_i \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) + \hat{\mathcal{P}}_t\right) - \varphi \hat{p}_{it} - (\hat{\mathcal{P}}_t - \hat{\mathcal{P}}_{t-1}) - \beta(\hat{\mathcal{P}}_{t+1} - \hat{\mathcal{P}}_t)$$
$$- [(\hat{y}_{it} - \hat{y}_{it-1}) - \beta(\hat{y}_{it+1} - \hat{y}_{it}) - (\hat{Y}_t - \hat{Y}_{t-1}) - \beta(\hat{Y}_{t+1} - \hat{Y}_t)] = \varphi \left(\theta_i \zeta_i \rho_i \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) + \theta_i \hat{\mathcal{P}}_t\right) - \varphi \theta_i \hat{p}_{it} - \theta_i(\hat{\mathcal{P}}_t - \hat{\mathcal{P}}_{t-1}) - \beta\theta_i(\hat{\mathcal{P}}_{t+1} - \hat{\mathcal{P}}_t)$$
$$(\hat{Z}_t - \hat{Z}_{t-1}) - \beta(\hat{Z}_{t+1} - \hat{Z}_t) = -\mathcal{M}\text{Cov}_\lambda[m_i^{-1}, \varphi\theta_i\zeta_i\rho_i \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) - \theta_i(\hat{\mathcal{P}}_t - \hat{\mathcal{P}}_{t-1}) - \beta\theta_i(\hat{\mathcal{P}}_{t+1} - \hat{\mathcal{P}}_t)] - \varphi\hat{Z}_t$$
$$(\hat{Z}_t - \hat{Z}_{t-1}) - \beta(\hat{Z}_{t+1} - \hat{Z}_t) = -\varphi\mathcal{M}\text{Cov}_\lambda[m_i^{-1}, \theta_i\zeta_i\rho_i] \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) + \mathcal{M}\text{Cov}_\lambda[m_i^{-1}, \theta_i]((\hat{\mathcal{P}}_t - \hat{\mathcal{P}}_{t-1}) + \beta(\hat{\mathcal{P}}_{t+1} - \hat{\mathcal{P}}_t)) - \varphi\hat{Z}_t$$
$$(\hat{Z}_t - \hat{Z}_{t-1}) - \beta(\hat{Z}_{t+1} - \hat{Z}_t) = -\varphi\mathcal{M} \left(\text{Cov}_\lambda[m_i^{-1}, \theta_i\zeta_i\rho_i] - \text{Cov}_\lambda[m_i^{-1}, \theta_i]\frac{\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]}{\mathbb{E}_\lambda[\theta_i]}\right) \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) - \varphi\hat{Z}_t$$
$$(\hat{Z}_t - \hat{Z}_{t-1}) - \beta(\hat{Z}_{t+1} - \hat{Z}_t) = -\varphi\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i] \left(\mathbb{E}_\lambda[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i\zeta_i\rho_i}{\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]}] - \mathbb{E}_\lambda[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}]\right) \left(\hat{mc}_t - \hat{\mathcal{P}}_t\right) - \varphi\hat{Z}_t$$

Then,

$$(\hat{Z}_t - \hat{Z}_{t-1}) - \beta(\hat{Z}_{t+1} - \hat{Z}_t) = -\varphi\kappa_Z \left(\Omega\hat{Y}_t - \Xi\hat{Z}_t + \hat{D}_t\right) - \varphi\hat{Z}_t$$

from which we can obtain the expression in (11).

Euler equation. From the utility maximization problem of households, we obtain the generic (log-linearized) Euler equation as:

$$\hat{C}_t - \sigma\iota_{cl}\hat{L}_t = \mathbb{E}_t\left[\hat{C}_{t+1} - \sigma\iota_{cl}\hat{L}_{t+1}\right] - \sigma\left(\hat{i}_t - \mathbb{E}_t\left[\hat{\pi}_{t+1}\right]\right) \tag{A.49}$$

Knowing the aggregate demand function for labor from equation (A.33), and equilibrium prices from equations (A.37) and (A.38), and using the expression for $\hat{C}_t$ in (A.40) we can

derive the Euler equation as:

$$c\hat{Y}_t - \tilde{c}\left(\hat{Y}_t - \hat{Z}_t\right) = \mathbb{E}\left[c\hat{Y}_{t+1} - \tilde{c}\left(\hat{Y}_{t+1} - \hat{Z}_{t+1}\right)\right] - \sigma\left(\hat{i}_t - \mathbb{E}_t\left[\hat{\pi}_{t+1}\right]\right) \tag{A.50}$$

where $c \equiv \frac{1+\eta^{-1}\phi+\nu^{-1}(1-\phi)+\sigma\iota_{cl}(1-\phi)(\sigma^{-1}-\iota_{lc})}{\lambda_c(1+\eta^{-1}\phi+\nu^{-1}(1-\phi))+(1-\lambda_c)(1+\eta^{-1})\phi(\sigma^{-1}-\iota_{lc})}$ and $\tilde{c} \equiv \frac{(1+\psi)}{a}\left(c(1-\lambda_c) + \frac{\sigma\iota_{cl}}{(1+\nu^{-1})}\right)$.

Appendix B. Identification results

B.1. Extended model: adding capital, firm-specific input cost and demand shocks

To consider the mapping between the model and the data, we make three modifications to the baseline model. First, we introduce firm-specific shocks to the cost of materials. Second, we introduce firm-specific demand shocks. Third, we introduce capital as a fixed factor in the production function.

B.1.1. Additional assumptions

Firm-specific demand shocks. We use the specification of the demand curve in xx:

$$p_{it} = \xi_{it}\Upsilon'(\frac{y_{it}}{Y_t})\mathcal{P}_t \tag{B.1}$$

where $\mathcal{P}_t$ is defined as before. We can then define the demand function:

$$\frac{y_{it}}{Y_t} = \Psi\left(\frac{p_{it}}{\xi_{it}P_t}\right) \tag{B.2}$$

where $\Psi = \Upsilon'^{-1}$.

Intermediate varieties producers.

$$y_{it} = z_i(k_i^{\gamma}v_{it}^{1-\gamma})^a \tag{B.3}$$

We consider firm-specific shocks to the price of materials, akin to iceberg transportation cost shocks, denoted $e^{\vartheta_{it}}$. ϑ_{it} is an idiosyncratic shock satisfying $\vartheta_{it} \perp \lambda_i$.

B.1.2. Characterization

Households. We again denote the price elasticity of demand by θ_{it}, a function of relative size only.

$$\hat{y}_{it} - \hat{Y}_t = -\theta_i(\hat{p}_{it} - \hat{P}_t) + \frac{\theta_i}{Y}\hat{\xi}_{it} \tag{B.4}$$

Intermediate varieties producers. We first consider the firm's cost-minimization problem, taking output and input prices as given. We can first solve for the optimal choice of labor and materials, taking variable inputs and prices as given:

$$\min_{l_{it},x_{it}}[e^{\vartheta_{it}} w_t^x x_{it} + w_t^l l_{it}] \quad \text{subject to} \quad (l_{it}^{\phi} x_{it}^{1-\phi}) \geq \bar{v} \tag{B.5}$$

We obtain that for any v_{it}, the choice of labor and material will be given by:

$$l_{it} = \left(\frac{w_t^x e^{\vartheta_{it}}}{w_t^l}\frac{\phi}{1-\phi}\right)^{1-\phi} v_{it} \tag{B.6}$$

$$x_{it} = \left(\frac{w_t^x e^{\vartheta_{it}}}{w_t^l}\frac{\phi}{1-\phi}\right)^{-\phi} v_{it} \tag{B.7}$$

This defines a price index for variable inputs, inclusive of the iceberg cost shock (but exclusive of the input wedge):

$$w_{it}^v = (w_t^l/\phi)^{\phi}(w_t^x e^{\vartheta_{it}}/(1-\phi))^{1-\phi} \tag{B.8}$$

We then solve for the optimal choice of variable inputs, for a given level of output:

$$\min_{v_{it}}(1+\tau_i)(w_{it}^v v_{it} + w_t^k k_i) \quad \text{subject to} \quad z_i(k_i^{\gamma} v_{it}^{1-\gamma})^a \geq \bar{y} \tag{B.9}$$

The solution writes:

$$(1+\tau_i)w_{it}^v = \lambda_{it}\frac{\partial y_{it}(k_i, v_{it})}{\partial v_{it}} = \lambda_{it}(1-\gamma)a z_i k_i^{\gamma a} v_{it}^{(1-\gamma)a-1} \tag{B.10}$$

Where λ_{it} is the Lagrange multiplier on the constraint, equal to the marginal cost. Hence,

$$mc_{it} = (1+\tau_i)w_{it}^v \frac{1}{a(1-\gamma)}\frac{v_{it}}{y_{it}} = (1+\tau_i)w_{it}^v \frac{1}{a(1-\gamma)} y_{it}^{\frac{1}{a(1-\gamma)}-1}\left(z_i k_i^{a\gamma}\right)^{-\frac{1}{a(1-\gamma)}}$$

The cost function then writes:

$$C(y_{it}, w_{it}) = (1+\tau_i)w_{it}^v y_{it}^{\frac{1}{a(1-\gamma)}}\left(z_i k_i^{a\gamma}\right)^{-\frac{1}{a(1-\gamma)}} + (1+\tau_i)w_{it}^k k_i \tag{B.11}$$

Marginal cost-based Phillips curve. A firm that can reset its price at time t will choose:

$$\hat{p}_{it|t}^{new} = (1-\beta\alpha)\mathbb{E}_t\left[\sum_{s=0}^{+\infty}(\beta\alpha)^s(\hat{\mu}_{it+s|t}^f + \hat{mc}_{it+s|t})\right] \tag{B.12}$$

$\hat{\mu}_{it+s|t}^f$ is the optimal markup at $t+s$ of a firm that could last reset its price at t. It is given by:

$$\hat{\mu}_{it+s|t}^f = -\Gamma_i\theta_i(\hat{p}_{it|t}^{new} - \hat{\mathcal{P}}_{t+s}) + \Gamma_i\frac{\theta_i}{Y}\hat{\xi}_{it+s|t} \tag{B.13}$$

Let $\hat{mc}_t \equiv \mathbb{E}_\lambda[\hat{mc}_{it}]$ the change in the aggregate nominal marginal cost. We can write:

$$\begin{aligned}\hat{mc}_{it+s|t} - \hat{mc}_{t+s} &= (1-\phi)\tilde{\vartheta}_{it+s} + \frac{1-a(1-\gamma)}{a(1-\gamma)}(\hat{y}_{it+s|t} - \hat{Y}_{t+s}) \\ &= (1-\phi)\tilde{\vartheta}_{it+s} - \frac{1-a(1-\gamma)}{a(1-\gamma)}\theta_i(\hat{p}_{it|t}^{new} - \hat{\mathcal{P}}_{t+s}) + \frac{1-a(1-\gamma)}{a(1-\gamma)}\frac{\theta_i}{Y}\hat{\xi}_{it+s|t}\end{aligned}$$

where $\tilde{\vartheta}_{it}$ is the idiosyncratic component of the cost shock $\vartheta_{it} = \vartheta_t + \tilde{\vartheta}_{it}$, $\vartheta_t = \mathbb{E}_\lambda[\vartheta_{it}]$.

$$\hat{p}_{it|t}^{new} = (1-\beta\alpha)\mathbb{E}_t\left[\sum_{s=0}^{+\infty}(\beta\alpha)^s\left(\rho_i\zeta_i\hat{mc}_{t+s} + \rho_i\zeta_i(1-\phi)\tilde{\vartheta}_{it+s} + (1-\rho_i\zeta_i)\hat{\mathcal{P}}_{t+s} + (1-\rho_i\zeta_i)\frac{1}{Y}\hat{\xi}_{it+s|t}\right)\right]$$

where ρ_i is the flexible price partial equilibrium pass-through of marginal cost shocks into prices and $\zeta_i = \frac{1}{1+\frac{1-a(1-\gamma)}{a(1-\gamma)}\theta_i\rho_i}$.We can write this equation recursively as:

$$\hat{p}_{it|t}^{new} = (1-\beta\alpha)\left(\zeta_i\rho_i\hat{mc}_t + \rho_i\zeta_i(1-\phi)\tilde{\vartheta}_{it} + (1-\zeta_i\rho_i)\hat{\mathcal{P}}_t + (1-\rho_i\zeta_i)\frac{1}{Y}\hat{\xi}_{it}\right) + \beta\alpha\mathbb{E}_t[\hat{p}_{it+1|t+1}^{new}] \tag{B.14}$$

Aggregating across firms, we obtain the marginal cost-based Phillips curve:

$$\hat{\pi}_t = \varphi \mathbb{E}_\lambda[\zeta_i \rho_i](\hat{mc}_t - \hat{P}_t^Y) - \varphi(1 - \mathbb{E}_\lambda[\zeta_i \rho_i])\hat{D}_t + \beta \mathbb{E}_t[\hat{\pi}_{t+1}] \tag{B.15}$$

Aggregate marginal cost is given by:

$$\hat{mc}_t = \Omega \hat{Y}_t - \Xi \hat{Z}_t + \iota^\vartheta \vartheta_t + \hat{P}_t^Y \tag{B.16}$$

$$\hat{\pi}_t = \varphi \mathbb{E}_\lambda[\zeta_i \rho_i] \left(\Omega \hat{Y}_t - \Xi \hat{Z}_t + \iota^\vartheta \vartheta_t \right) - \varphi(1 - \mathbb{E}_\lambda[\zeta_i \rho_i])\hat{D}_t + \beta \mathbb{E}_t[\hat{\pi}_{t+1}] \tag{B.17}$$

Allocative efficiency. We are interested in an aggregate production function of the form

$$Y_t = Z_t \left(K^\gamma V_t^{1-\gamma} \right)^a, \tag{B.18}$$

which implies, after imposing that aggregate capital is time-invariant, that aggregate productivity is given by

$$\hat{Z}_t = \hat{Y}_t - a(1-\gamma)\hat{V}_t. \tag{B.19}$$

In log-deviations, firm-level distortions are given by

$$\hat{m}_{it} = \hat{p}_{it} + \hat{y}_{it} - \hat{w}_{it} - \hat{v}_{it},$$

which after imposing the shape of the production function $\hat{v}_{it} = \frac{1}{(1-\gamma)a}\hat{y}_{it}$, implies that

$$\hat{m}_{it} = \hat{p}_{it} + \frac{(1-\gamma)a - 1}{(1-\gamma)a}\hat{y}_{it} - \hat{w}_{it}.$$

Taking expectations across firms yields:

$$\mathbb{E}_\lambda \hat{m}_{it} = \hat{P}_t^Y + \frac{(1-\gamma)a - 1}{(1-\gamma)a}\hat{Y}_t - \hat{w}_t. \tag{B.20}$$

The aggregate distortion in log-linear terms is given by

$$\hat{\mathcal{M}}_t = \hat{P}_t^Y + \hat{Y}_t - \hat{W}_t - \hat{V}_t \tag{B.21}$$

which with some manipulation yields an equation similar to what we had before

$$\hat{Z}_t = (1-\gamma)a(\hat{\mathcal{M}}_t - \mathbb{E}_\lambda \hat{m}_{it}) \tag{B.22}$$

Now, I will use the expression

$$\mathcal{M}_t = \frac{1}{\mathbb{E}_\lambda m_{it}^{-1}}, \tag{B.23}$$

which in logs is equal to $\log \mathcal{M}_t = -\log \int \lambda_{it} m_{it}^{-1}$ A first-order approximation of this equation yields

$$\hat{\mathcal{M}}_t = -\frac{\int \lambda_{it}^* (m_i^*)^{-1} (\hat{\lambda}_{it} - \hat{m}_{it}}{\int \lambda_{it}^* (m_i^*)^{-1}} \tag{B.24}$$

$$= -\mathcal{M}^* \mathbb{E}_\lambda (m_{it}^{*-1} (\hat{\lambda}_{it} - \hat{m}_{it}). \tag{B.25}$$

Using the properties of the covariance, we get:

$$\hat{\mathcal{M}}_t = \mathbb{E}_\lambda \hat{m}_{it} - \mathcal{M}^* Cov_\lambda (m_i^{*-1}, \hat{\lambda}_{it} - \hat{m}_{it}), \tag{B.26}$$

and using the definition of $\hat{Z}$, we obtain:

$$\hat{Z}_t = -(1-\gamma)a\mathcal{M}^* Cov_\lambda (m_i^{*-1}, \hat{\lambda}_{it} - \hat{m}_{it}) \tag{B.27}$$

Using the definitions for marginal costs and sale shares, we can express the driver term in the last equation as:

$$\hat{\lambda}_{it} - \hat{mc}_{it} = -(P_t^Y + Y_t) + \hat{w}_t^v + \hat{\vartheta}_{it} + \frac{1}{(1-\gamma)a} \hat{y}_{it}. \tag{B.28}$$

Notice that most of the terms in the last line are either aggregate terms, so they are orthogonal to firm-level variables in steady state, or are assumed to be orthogonal, like the presence of input cost shocks. Therefore:

$$\hat{Z}_t = -\mathcal{M}^* Cov_\lambda (m_i^{*-1}, \hat{y_{it}}), \tag{B.29}$$

or alternatively using the shape of demand curves:

$$\hat{Z}_t = \mathcal{M}^* Cov_\lambda(m_i^{*-1}, \theta_i(\hat{p}_{it} - \hat{\mathcal{P}}_t)). \tag{B.30}$$

Note that this result does not depend on the exponent of variable inputs on production.

Another important point is that the expression for reset prices does not depend explicitly on a or γ once we control for marginal costs, which of course, depend on the shape of the production technology.

The law of motion of actual prices follows the second-order difference equation

$$p_{it} - p_{i,t-1} = \varphi\Delta_{it} + \beta(p_{i,t+1} - p_{it}) - \varphi p_{it}, \tag{B.31}$$

for $\Delta_{it} = \zeta_i\rho_i(\hat{mc}_t + (1-\phi)\tilde{\vartheta}_{it}) + (1-\zeta_i\rho_i)\hat{\mathcal{P}}_t$.

Notice that the second-order difference equation for prices is the same than in the main body of the paper. The result that comes from the definition of $\hat{D}$ also holds

$$\mathcal{P}_t\mathbb{E}_\lambda\theta_i = \mathbb{E}_\lambda(\theta_i\hat{p}_{it}). \tag{B.32}$$

Take the law of motion of prices and subtract in each side $\mathcal{P}_t - \mathcal{P}_{t-1} - \beta(\mathcal{P}_{t+1} - \mathcal{P}_t)$.

$$p_{it}-p_{i,t-1}-(\mathcal{P}_t - \mathcal{P}_{t-1} - \beta(\mathcal{P}_{t+1} - \mathcal{P}_t)) = \varphi\Delta_{it}+\beta(p_{i,t+1}-p_{it})-\varphi p_{it}-(\mathcal{P}_t - \mathcal{P}_{t-1} - \beta(\mathcal{P}_{t+1} - \mathcal{P}_t))$$

Collect terms in the LHS

$$(p_{it}-\mathcal{P}_t)-(p_{i,t-1}-\mathcal{P}_{t-1})-\beta\left((p_{it+1} - \mathcal{P}_{t+1}) - (p_{it} - \mathcal{P}_t)\right) = \varphi\Delta_{it}-\varphi p_{it}-(\mathcal{P}_t - \mathcal{P}_{t-1} - \beta(\mathcal{P}_{t+1} - \mathcal{P}_t))$$

Multiply by $m_i^{-1*}\theta_i$.

$$m_i^{*-1}\theta_i(p_{it}-\mathcal{P}_t)-m_i^{*-1}\theta_i(p_{i,t-1}-\mathcal{P}_{t-1})-\beta\left(m_i^{*-1}\theta_i(p_{it+1} - \mathcal{P}_{t+1}) - m_i^{*-1}\theta_i(p_{it} - \mathcal{P}_t)\right) = m_i^{*-1}\theta_i\varphi\Delta_{it}-m_i^{*-1}\theta_i\varphi$$

Now take sales-weighted expectations

$$\mathbb{E}_\lambda m_i^{*-1}\theta_i(p_{it}-\mathcal{P}_t)-\mathbb{E}_\lambda m_i^{*-1}\theta_i(p_{i,t-1}-\mathcal{P}_{t-1})-\beta\left(\mathbb{E}_\lambda m_i^{*-1}\theta_i(p_{it+1} - \mathcal{P}_{t+1}) - \mathbb{E}_\lambda m_i^{*-1}\theta_i(p_{it} - \mathcal{P}_t)\right) = \mathbb{E}_\lambda m_i^{*-1}\theta_i\varphi$$

and realize that $\mathbb{E}_\lambda m_i^{*-1}\theta_i(\hat{p}_{it} - \hat{\mathcal{P}}_t) = \mathcal{M}^{*-1}\hat{Z}_t$.

$$\hat{Z}_t - \hat{Z}_{t-1} - \beta\left(\hat{Z}_{t+1} - \hat{Z}_t\right) = \mathbb{E}_\lambda m_i^{*-1}\theta_i\varphi\Delta_{it} - \mathbb{E}_\lambda m_i^{*-1}\theta_i\varphi p_{it} - \mathbb{E}_\lambda m_i^{*-1}\theta_i\left(\mathcal{P}_t - \mathcal{P}_{t-1} - \beta(\mathcal{P}_{t+1} - \mathcal{P}_t)\right),$$

We will use now the result from the previous section that $(\mathcal{P}_t - \mathcal{P}_{t-1} - \beta(\mathcal{P}_{t+1} - \mathcal{P}_t)) = \varphi\frac{\mathbb{E}_\lambda(\theta_i\zeta_i\rho_i)}{\mathbb{E}_\lambda\theta_i}(\hat{mc}_t - \hat{\mathcal{P}}_t)$. This result comes from the firm-level second order difference equation of firms, after multiplying by θ_i, taking sales-weighted expectations, and using the result that $\hat{\mathcal{P}}_t\mathbb{E}_\lambda\theta_i = \mathbb{E}_\lambda(\theta_i\hat{p}_{it})$, which comes from the definition of $\hat{D}$.

$$\mathcal{M}^{-1}(\hat{Z}_t - \hat{Z}_{t-1} - \beta\left(\hat{Z}_{t+1} - \hat{Z}_t\right)) = \mathbb{E}_\lambda m_i^{*-1}\theta_i\varphi\zeta_i\rho_i(\hat{mc}_t - \hat{\mathcal{P}}_t) - \varphi\mathcal{M}^{-1}\hat{Z}_t - \mathbb{E}_\lambda(m_i^{*-1}\theta_i)\varphi\frac{\mathbb{E}_\lambda(\theta_i\zeta_i\rho_i)}{\mathbb{E}_\lambda\theta_i}(\hat{mc}_t - \hat{\mathcal{P}}_t),$$

$$\hat{Z}_t - \hat{Z}_{t-1} - \beta\left(\hat{Z}_{t+1} - \hat{Z}_t\right) = \mathcal{M}\varphi\mathbb{E}_\lambda\left(m_i^{*-1}\theta_i\zeta_i\rho_i\right)(\hat{mc}_t - \hat{\mathcal{P}}_t) - \varphi\hat{Z}_t - \mathcal{M}\varphi\mathbb{E}_\lambda(m_i^{*-1}\theta_i)\frac{\mathbb{E}_\lambda(\theta_i\zeta_i\rho_i)}{\mathbb{E}_\lambda\theta_i}(\hat{mc}_t - \hat{\mathcal{P}}_t),$$

$$\hat{Z}_t - \hat{Z}_{t-1} - \beta\left(\hat{Z}_{t+1} - \hat{Z}_t\right) = \mathcal{M}\varphi(\hat{mc}_t - \hat{\mathcal{P}}_t)\left(\mathbb{E}_\lambda\left(m_i^{*-1}\theta_i\zeta_i\rho_i\right) - \mathbb{E}_\lambda(m_i^{*-1}\theta_i)\frac{\mathbb{E}_\lambda(\theta_i\zeta_i\rho_i)}{\mathbb{E}_\lambda\theta_i}\right) - \varphi\hat{Z}_t$$

Which again, is the same second order difference equation for Z. The only difference then can come from the determination of marginal costs.

B.2. Identification results: slope

Passthrough of cost shock into prices. Take the case of a shock with zero persistence. First, write $\hat{p}_{it} = \mathbb{1}_{it}^p\hat{p}_{it|t}^{new} + (1-\mathbb{1}_{it}^p)\hat{p}_{it-1}$, where $\mathbb{1}_{it}^p$ is a "Calvo-fairy" dummy that takes the value of 1 if a price adjustment is permissible for firm i in period t. Using the formula for the optimal reset price,

$$\hat{p}_{it} - \hat{p}_{it-1} = \mathbb{1}_{it}^p(1-\beta\alpha)\zeta_i\rho_i(1-\phi)\tilde{\vartheta}_{it} + \mathbb{1}_{it}^p(1-\beta\alpha)(1-\rho_i\zeta_i)\frac{1}{Y}\hat{\xi}_{it} + \mathbb{1}_{it}^p(1-\beta\alpha)\mathbb{E}_t\left[\sum_{s=0}^{+\infty}(\beta\alpha)^s\left(\zeta_i\rho_i\hat{mc}_{t+s} + (1-\zeta_i\rho_i)\hat{\mathcal{P}}_{t+s}\right)\right] - \mathbb{1}_{it}^p\hat{p}_{it-}$$

Proposition 1. *Let us assume that $\mathbb{1}_{it}^p \perp \tilde{\vartheta}_{it}$, $\mathbb{1}_{it}^p \perp \lambda_i$, and $\tilde{\vartheta}_{it} \perp \hat{\xi}_{it}$. In addition, we maintain the assumption outlined above that $\lambda_i \perp \tilde{\vartheta}_{it}$. Finally, assume that we observe $\mathcal{Z}_{it}^{\vartheta}$ a proxy for $(1-\phi)\vartheta_{it}$*

satisfying $(1-\phi)\vartheta_{it} = k^{\vartheta} z_{it}^{\vartheta}$. *Then, the coefficient of the regression*

$$\Delta \log p_{it} = \alpha_t + \beta_0^{RF} z_{it}^{\vartheta} + \varepsilon_{it} \tag{B.33}$$

identifies $\beta_0^{RF} = k^{\vartheta}(1-\beta\alpha)(1-\alpha)\mathbb{E}_\lambda[\rho_i\zeta_i]$. *The coefficient of the regression:*

$$\Delta \log w_{it} = \alpha_t + \beta_0^{FS} z_{it}^{\vartheta} + \varepsilon_{it} \tag{B.34}$$

identifies $\beta_0^{FS} = k^{\vartheta}$. *Consequently, the IV coefficient identifies* $\beta_0^{IV} = (1-\beta\alpha)(1-\alpha)\mathbb{E}_\lambda[\rho_i\zeta_i]$.

What about the case where the input cost shock has some persistence? Let us assume that ϑ_{it} is an AR(1) with persistence ρ_ϑ.

Proposition 2. *Let us denote* β_h *the local projection coefficient at horizon h. Under the same notations and assumptions,* $\beta_0^{RF} = k^{\vartheta}\frac{1-\beta\alpha}{1-\beta\alpha\rho_\vartheta}(1-\alpha)\mathbb{E}_\lambda[\rho_i\zeta_i]$, $\beta_1^{RF} = k^{\vartheta}\frac{1-\beta\alpha}{1-\beta\alpha\rho_\vartheta}\mathbb{E}_\lambda[\rho_i\zeta_i](1-\alpha)\,(\rho_\vartheta+\alpha)$, $\beta_2^{RF} = k^{\vartheta}\frac{1-\beta\alpha}{1-\beta\alpha\rho_\vartheta}\mathbb{E}_\lambda[\rho_i\zeta_i](1-\alpha)(\rho_\vartheta^2+\alpha\rho_\vartheta+\alpha^2)$.

B.3. Identification results: allocative efficiency

B.3.1. Identification strategy

$$\kappa_D = \mathbb{E}_\lambda[\zeta_i\rho_i]\left(\mathbb{E}_\lambda\left[\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\frac{\zeta_i\rho_i}{\mathbb{E}_\lambda[\zeta_i\rho_i]}\right]-1\right)$$
$$\kappa_Z = \mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]\left(\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i\zeta_i\rho_i}{\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]}\right]-\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\right]\right)$$

We first detail how to estimate each objects, assuming m_i^{-1} and θ_i are known. We then detail the measurement of m_i^{-1} and θ_i.

Identification of $\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]$ and $\mathbb{E}_\lambda[\zeta_i\rho_i]$. $\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]$ is obtained from the regression of $\Delta\log y_{it}$ on $\Delta\log w_{it}^v$, instrumented by ϑ_{it}. $\mathbb{E}_\lambda[\zeta_i\rho_i]$ is obtained from the regression of $\Delta\log p_{it}$ on $\Delta\log w_{it}^v$, instrumented by ϑ_{it}.

Identification of $\mathbb{E}_\lambda\left[\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\frac{\zeta_i\rho_i}{\mathbb{E}_\lambda[\zeta_i\rho_i]}\right]$ and $\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i\zeta_i\rho_i}{\mathbb{E}_\lambda[\theta_i\zeta_i\rho_i]}\right]$. We obtain these quantities by estimating the price and quantity regressions by bins of $\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}$ and $\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}$, respectively, and using the law of iterated expectations.

Identification of $\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\right]$. $\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\right]$ is estimated directly from the data.

B.3.2. Measurement of m_i^{-1} and θ_i

Note that because we only need $\frac{\mathbb{E}_\lambda[m_i^{-1}\theta_i]}{\mathbb{E}_\lambda[\theta_i]}$, our procedure is robust to mis-measurement in the level of demand elasticities.

Estimation of the demand elasticities θ_i. We estimate demand elasticities by inverting the formula for desired markups: $\theta_i = \frac{\mu_i^f}{\mu_i^f - 1}$. This strategy is infeasible for individual firms for two reasons: (i) individual markup measures are very noisy, with many values very close to or below 1, making the estimate of demand elasticities extremely sensitive to measurement error ; (ii) with sticky prices, individual firm markups will diverge from ideal markups. To circumvent this problem, we take seriously the prediction of the model that within an industry, variation in the ideal markup only comes from variation in firm-level market shares. Define Q bins of market shares. We assume that $\forall i \in \mathcal{Q}_q,\ \mu_i^f = \mu_q^f = \mathbb{E}_\lambda[\mu_i^f|\mathcal{Q}_q]$. With the assumption that the markup is constant within $\mathcal{Q}_q$, we then obtain $\mathbb{E}_\lambda[\theta_i|\mathcal{Q}_q] = \frac{\mu_q^f}{\mu_q^f - 1} = \theta_q$. Using the law of iterated expectations, we obtain $\mathbb{E}_\lambda[\theta_i]$ and $\mathbb{E}_\lambda\left[\frac{m_i^{-1}}{\mathbb{E}_\lambda[m_i^{-1}]}\frac{\theta_i}{\mathbb{E}_\lambda[\theta_i]}\right]$.

Figure B.1 illustrates the markups and demand elasticities estimated in this way. We find that markups are increasing in firm's market shares, consistent with existing evidence. As a result, demand elasticities decline in firm size.

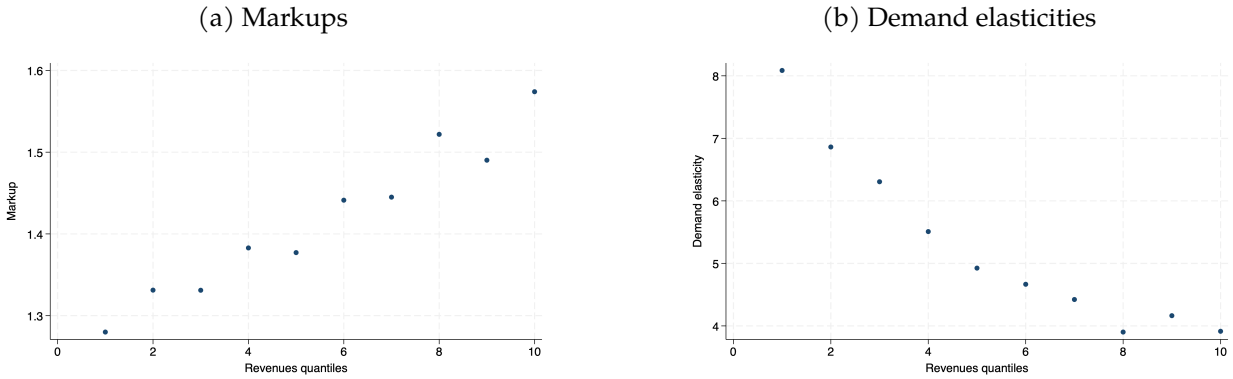
Estimation of the allocative distortion m_i^{-1}. As a reminder, we defined $m_i \equiv (1+\tau_i^v)\frac{\theta_i}{\theta_i - 1}$. It is then straightforward to show that:

$$m_i \propto \frac{p_i y_i}{v_i} = \left(\frac{p_i y_i}{l_i}\right)^{\phi}\left(\frac{p_i y_i}{m_i}\right)^{1-\phi} = \text{MRPL}^{\phi}\text{MRPM}^{1-\phi} \tag{B.35}$$

This formula shows that in the presence of both input-side distortions and markups, TFPR measures the combined distortion, as in Hsieh and Klenow (2009).

When taking this to the data, we consider two points. First, we integrate capital to our definition of wedges. While our baseline model omits capital, consistent with the focus on

FIGURE B.1. Markups and demand elasticities by size



Note: Panel (a) plots average markups by deciles of firms' market shares. Panel (b) plots demand elasticities by deciles of firms' market shares.

the response of marginal costs to monetary shocks, we do account for the fact that there are persistent distortions in marginal revenue products of capital across firms.

Second, constructing the material wedge $\frac{p_{it}y_{it}}{m_{it}}$ requires to divide sales by material quantity. Typical income statement items typically only report the purchase value of materials. This is a major issue: if material wedges are priced, the ratio $\frac{p_{it}y_{it}}{w^m_{it}m_{it}}$ would incorporate the wedge in the denominator of this ratio.

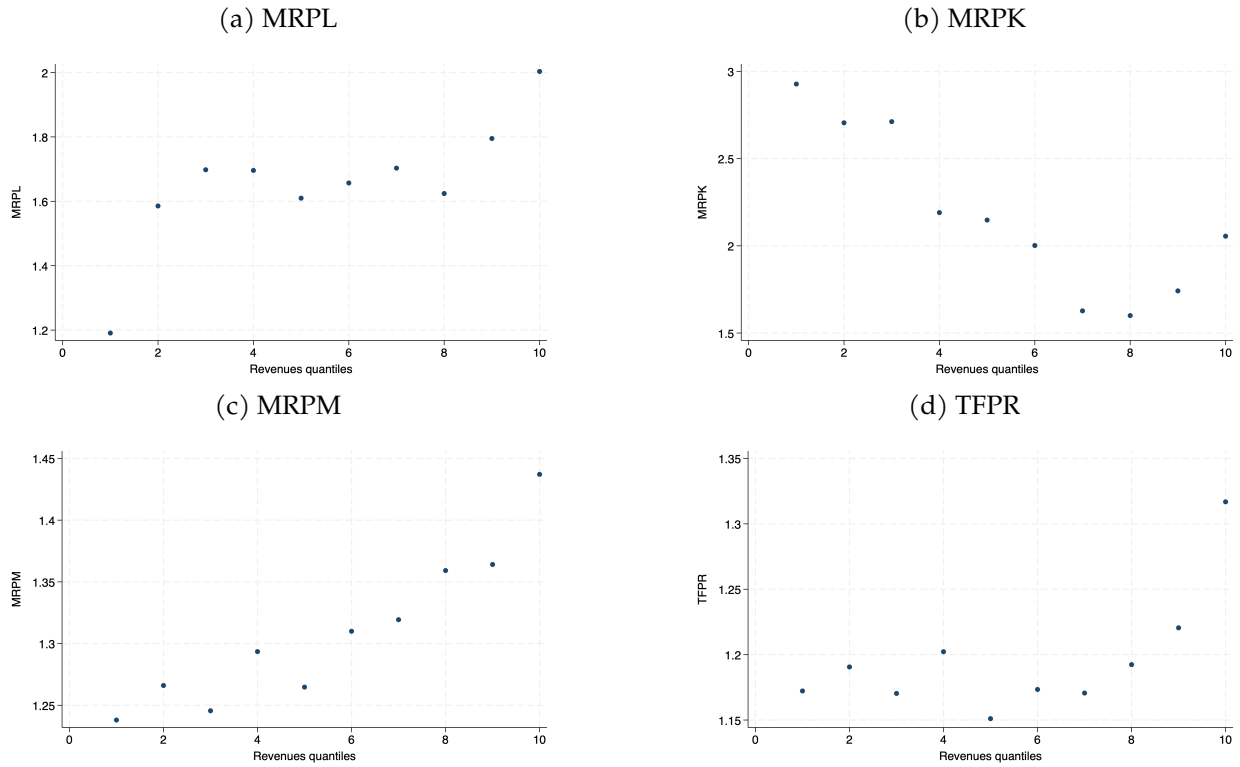
We overcome this issue by exploiting the decomposition of input purchases into quantities and unit values. This exercise requires that we have comparable quantities and unit values for each material inputs *across* firms, which is a more demanding requirement than our baseline tests exploiting within firm$\times$product comparability. We therefore restrict our attention to inputs for which quantities and unit values are reported based on physical quantities: length, areas, volume, mass, and energy units. In our sample, 89% of the total purchase value of material inputs is recorded in physical units.

In practice, we proceed as follows. Let $\mathcal{K}_{PU} \subseteq \mathcal{K}$ be the subset of inputs denominated in a physical unit. Here we consider the set $\mathcal{K}$ to be the original input classification: 5-digit ASICC codes until 2010, and 7-digit NPC codes afterwards. We do so because units are defined at this level, and vary within the more harmonized codes we use in the main empirical exercises (because here we only compare inputs used by different firms in the same year, we do not need harmonized classifications). For each input $k \in \mathcal{K}_{PU}$, we deflate the input purchase value $m_{kit}w^m_{kit}$ by the firm-specific component of the input price w^m_{kit}/w^m_{kt}, where the denominator is computed as the weighted median of the price of input k across

all firms in year t. We can then compute the ratio $\frac{p_{it}y_{it}}{m_{kit}w_{kt}^{m}}$. The denominator includes a constant across all firms, which is innocuous. We proceed in this fashion for all inputs $k \in \mathcal{K}_{PU}$ and aggregate at the firm level using input shares.

Figure B.2 shows the obtain marginal revenue products for labor, capital, intermediates, and the resulting TFPR, by deciles of firm market shares.

FIGURE B.2. MRPXs by size



Note: This figure plots averages of our estimates of marginal revenue products by bins of firm size. Within each industry, we define 10 bins of equal sales density.

Appendix C. Data

C.1. Indian Annual Survey of Industries (ASI)

The ASI is a dataset put together by India's Ministry of Statistics and Programme Implementation (MOSPI). The reference period for each survey is the accounting year, which in

India begins on the 1st of April and ends on the 31st of March the following year. Throughout the paper we reference the surveys by the earlier of the two years covered.

Coverage and sampling methodology. The ASI contains information on a representative sample of manufacturing establishments, conditional on them taking part of the organized sector, and either employing more than 20 employees, or employing more than 10 employees and using electricity. We call the subpopulation of firms satisfying this criteria the ASI population. Within the ASI population, ASI defines a Census sector which is sampled exhaustively and a Sample sector for which the micro-data contains only a representative sample. Details of how the sampling methodology for the ASI changes over time are shown in Table C.1. ASI provides sampling weights, which we use to weight all data moments. We compare total value added by establishments in the ASI population

TABLE C.1. Sampling Methodology for Indian ASI

Period	Census Sector	Sample Sector
1998	Complete enumeration states, plants with > 200 workers, all joint returns	Stratified within state $\times$ 4-digit industry (NIC-98), minimum of 8 plants per stratum
1999-2003	Complete enumeration states, plants with ≥ 100 workers, all joint returns	Stratified within state $\times$ 4-digit industry (NIC-98), minimum of 8 plants per stratum
2004-2006	6 less industrially developed states, 100 or more workers, all joint returns, all plants within state $\times$ 4-digit industry with < 4 units	Stratified within state $\times$ 4-digit industry, 20% sampling, minimum of 4 plants
2007	5 less industrially developed states, 100 or more workers, all joint returns, all plants within state $\times$ 4-digit industry with < 6 units	Stratified within state $\times$ 4-digit industry, minimum 6 plants, 12% sampling fraction: exceptions
2008-2013	6 less industrially developed states, 100 or more workers, all joint returns, all plants within state $\times$ 4-digit industry with < 4 units	Stratified within district $\times$ 4-digit industry, minimum 4 plants, 20% sampling fraction

to total manufacturing value added in India. The latter includes value added by the ASI population, value added by the organized sector establishments below the size threshold, and value added in the informal manufacturing sector. We find that ASI covers 61% of total manufacturing value added on average across years.

Sample selection. We start with 1,068,114 plant $\times$ year observations. We subsequently employ multiple sample selection rules. First, we restrict the sample to factory $\times$ year observations with either positive reported gross sales, or positive reported sales at the factory gate. This drops one third of all observations (360,145). Next, we disregard all observations that exactly copied their sales from the previous year, suspecting these plants to be actually closed. This drops 1,179 additional observations. Third, we drop all plant $\times$ years that reported either no days worked, or no persons employed, dropping a supplementary 206 observations. Moreover, we restrict the sample to observations with correct accounting: we drop all observations where the difference between aggregate items and the absolute value of the sum of corresponding sub-items exceeded 10% of the aggregate items. We regard four aggregate items: (i) the purchase value of basic items (including imports), (ii) the purchase value of non-basic items, (iii) the purchase value of total inputs, and (iv) the gross sales value of output. This accounting rule drops an extra 4,324 observations. Finally, disregarding all factory $\times$ years without any reported positive output or input (including energy) values at the product level, excludes an additional 49,100 observations. The final sample thus includes 189,466 unique plants for a total of 635,551 individual plants $\times$ year observations.

Industry classification. Our data relies on three distinct industry classification systems: NIC-98 (1998–2003), NIC-04 (2004–2007), and NIC-08 (2008 and beyond). We first address issues with the 5-digit industry codes in NIC-98, where codes are sometimes masked with zeroes or absent from the official documentation, by replacing them with the most frequent 5-digit code within each 4-digit grouping following the approach of Martin, Nataraj, and Harrison (2017). Next, we apply concordances from NIC-08 to NIC-04 using the mapping provided by Rijesh (2022). For industries not covered by this concordance, we manually supplement the mapping. For a given firm in NIC-08, we determine the corresponding NIC-04 code by leveraging data from the NIC-04 period (that is, "pre-data"). If no pre-data are available, we assign the candidate with the highest frequency; if pre-data exist, we select the candidate that matches the pre-data and has the highest frequency in the dataset. Finally, we apply concordances from NIC-04 to NIC-98 using the mapping provided by Martin, Nataraj, and Harrison (2017) and manually supplement the mapping for industries not covered by this concordance.

Product classification. Our analysis standardizes product classifications across four distinct classifications used in our sample: NPCMS 2015 (2016-2017), NPCMS 2011 (2010-2015), ASICC 2009 (2008-2009), and ASICC 2008 (pre-2008). We harmonize all product codes to NPCMS 2011, as it provides a well-defined five-digit structure that balances granularity and coverage. Given the absence of an official concordance between NPCMS 2015 and NPCMS 2011, we constructed a mapping using fuzzy matching (based on product codes, descriptions, and units) and semantic embeddings (OpenAI's AA2 model). For ASICC 2009, we utilize the official concordance to NPCMS 2011 but address its limitations (such as missing mappings and invalid classification) by leveraging ASI data and semantic embeddings. The harmonization of ASICC 2008 to ASICC 2009 follows the concordance from Boehm, Dhingra, and Morrow (2022).

TABLE C.2. Example of NPC-MS 2011 5-digit classification

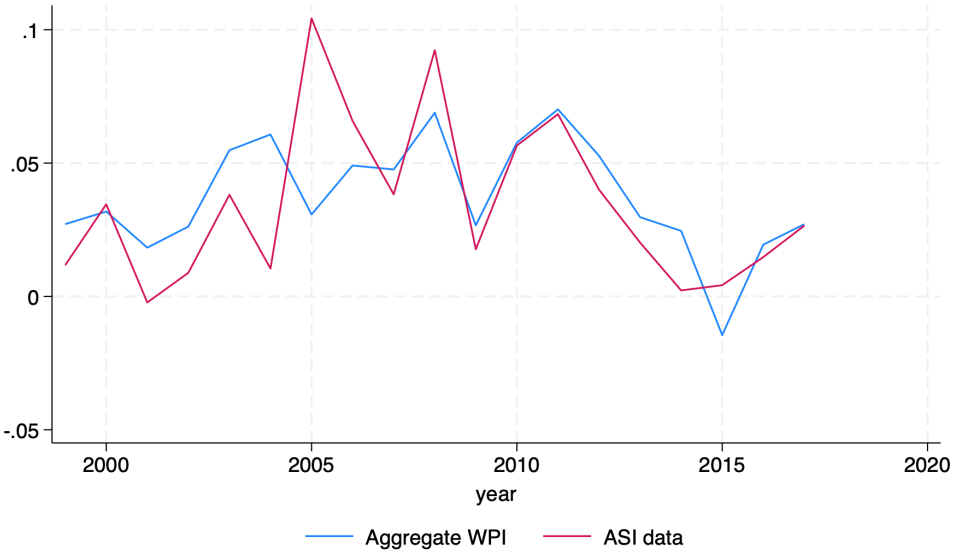
Code	Description
35	Other chemical products; man-made fibres
351	Paints and varnishes and related products; artists' colours; ink
35110	Paints and varnishes and related products
35120	Artists', students' or signboard painters' colours, modifying tints, amusement colours and the like
35130	Printing ink
35140	Writing or drawing ink and other inks
352	Pharmaceutical products
353	Soap, cleaning preparations, perfumes and toilet preparations
354	Chemical products n.e.c.
355	Man-made fibres
36	Rubber and plastics products
361	Rubber tyres and tubes
36111	New pneumatic tyres, of rubber, of a kind used on motor cars
36112	New pneumatic tyres, of rubber, of a kind used on motorcycles or bicycles
36113	Other new pneumatic tyres, of rubber
36114	Inner tubes, solid or cushion tyres, interchangeable tyre treads and tyre flaps, of rubber
36115	Camel back strips for retreading rubber tyres
36120	Retreaded pneumatic tyres, of rubber
362	Other rubber products
36210	Reclaimed rubber
36220	Unvulcanized compounded rubber, in primary forms or in plates, sheets or strip; unvulcanized rubber in forms other than primary forms or plates, sheets or strip
36230	Tubes, pipes and hoses of vulcanized rubber other than hard rubber
36240	Conveyor or transmission belts or belting, of vulcanized rubber
36250	Rubberized textile fabrics, except tyre cord fabric
36260	Articles of apparel and clothing accessories (including gloves) of vulcanized rubber other than hard rubber
36270	Articles of vulcanized rubber n.e.c.; hard rubber; articles of hard rubber
363	Semi-manufactures of plastics
364	Packaging products of plastics
369	Other plastics products

Note: For codes other than 351, 361 & 362, the 5-digit classifications are not shown.

Data quality check. Figure C.1 matches the yearly price index for the manufacturing sector from WPI, constructed as the average of quarterly price indices, to the price index

generated from ASI data. The latter is constructed as the sales-weighted median of log-price changes, with weights being constructed as Törnqvist weights. While the ASI data appears to have a little more volatility, especially between 2005 and 2008, overall the two lines show a very similar pattern.

FIGURE C.1. WPI inflation and inflation from ASI micro-data



Note: This figure plots two inflation series. The blue line is WPI inflation (manufacturing). The pink line is price growth obtained from the micro-data. It is constructed as the sales-weighted median of log-price changes, with weights being constructed as Törnqvist weights.