EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Sergio Santoro, Eliana Viviano

Optimal trend inflation, misallocation and the pass-through of labour costs to prices

No 2761 / December 2022

Disclaimer: This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Abstract

We show that a sticky price model featuring firms' heterogeneity in terms of productivity and strategic complementarities in price setting delivers a strictly positive optimal inflation in steady state, differently from standard New Keynesian models. Due to strategic complementarities, more productive firms have higher markups in steady state. This leads to a misallocation distortion, as more productive firms produce too little compared to the social optimum. An increase of steady state inflation curbs the markups, especially those of the more productive firms, hence attenuating the inefficient dispersion of markups. At low levels of inflation, the gains from the reduction in misallocation outweigh the cost of inflation. Heterogeneity in productivity and strategic complementarities in price setting, the key ingredients of our model, imply that also firms' response to shocks is heterogenous: less productive firms transmit cost shocks to prices much more than more productive ones. To provide empirical support to our key mechanism we resort to a quasi-natural experiment occurred in Italy in late 2014, when a cut to social security contributions for all new open-ended contracts was announced. Consistently with our theory, we show that the pass-through of this shock to labour costs was much stronger for less productive firms.

JEL classification: D00, D22, E31.

Keynotes: optimal inflation rate, price pass-through, labour costs, firm heterogeneity.

Non-technical summary

A long-standing question in monetary policy is the level of inflation that a central bank should target. When price rigidity is the main source of monetary non-neutrality, a well-established result is that zero inflation is optimal. The reason for the optimality of price invariance is that it eliminates the inefficiencies brought about by the presence of price-adjustment costs.

This prescription is at odds with practice in most central banks of advanced countries, which adopt a quantitative inflation objective gravitating around a level of 2%. Traditional arguments put forward to reconcile theory and practice relate to the role of a positive inflation trend as a buffer against risks of deflation, in the presence of a zero lower bound on nominal interest rates, of downward nominal wage rigidities, or in the case of positive bias in inflation measurement. In this paper we provide an alternative channel that can bridge the gap between theory and practice; this channel is based on production misallocation, namely on the feature that more productive firms don't produce as much as would be desirable from an efficiency point of view, and viceversa for less productive firms: the extent and relevance of misallocation has been documented by a burgeoning literature, but the implications for monetary policy are still relatively under-explored.

To introduce this novel channel, we build on the model by Edmond, Midrigan, and Xu (2018), characterised by heterogeneous firm productivity and strategic complementarities among firms. The latter means that price elasticity of demand increases with the firm's relative price, spurring more consumers to flee from goods with high relative prices, and less to flock to inexpensive ones, compared to a case without strategic complementarity (where the elasticity of demand is constant). As argued in Edmond, Midrigan, and Xu (2018), in such a setup more productive firms that can charge lower prices face a lower demand elasticity for the goods they produce, due to the strategic complementarity assumption; this in turn translates into a stronger market power (higher markup) than less productive firms, allowing more productive firms to set higher markups. This markup dispersion entails a production distribution too much skewed towards less productive firms, resulting in production misallocation.

We add to this setup introducing price rigidity in the form of quadratic price-adjustment costs. Markups get smaller in presence of positive trend inflation: firms incur adjustment costs to keep up with inflation, and to pay them they want to increase revenues and output beyond what would be convenient with flexible prices. This markup reduction is more pronounced, the higher markup is: a firm can charge a higher markup if it faces a weaker demand elasticity, which implies that for the same increase in output the price falls by more. Hence, if a central bank implements a higher inflation target, it curbs misallocation: it spurs the more productive firms to expand their production more than less productive firms, reducing the markup dispersion, with positive effects for aggregate productivity and welfare. The central bank has to weigh this benefit against the cost of inflation related to price rigidity: a numerical example shows that this misallocation channel alone can contribute to a positive inflation target by a quantitatively relevant amount, around 1% or more.

Our theoretical results are based on the joint hypotheses of firms' heterogeneity and strategic complementarity. These assumptions imply a feature that can be empirically estimated: pass-through of shocks to labour costs into prices is decreasing in firm's productivity. To provide empirical support to our model, we estimate the pass-through from labour costs to prices at the micro level drawing on a rich dataset of Italian firms, and analyse the relationship between pass-through and productivity relying on proxies of productivity typically used by the literature, like firm size. A challenge in carrying out this estimation is to disentangle the impact of changes in labour costs on prices, from the effect of prices on labour costs. To isolate the former, we resort to a policy change occurred in Italy in late 2014, when a large 3-year social security contribution cut was announced by the government for hirings taking place the following year. We use this policy change as an exogenous variation in labour costs and relate it to firms' price changes. We merge administrative employer-employee data, firms' balance sheets and a survey on industrial and non-financial service firms conducted every year by the Bank of Italy, which reports information on firms' own price change. The results show that small firms transmit a change in labour costs to prices much more than large firms. These results allow us to conclude that less productive firms have a higher pass-through of labour cost to prices, consistently with our theory.

1 Introduction

A long-standing question in monetary economics is the trend inflation[1] that a central bank should aim at. When price rigidity is the main source of monetary non-neutrality, a well-established result is that zero inflation is optimal. The reason for the optimality of price invariance is that it eliminates the inefficiencies brought about by the presence of price-adjustment costs.[2]

This prescription is at odds with practice in most central banks of advanced countries, which adopt a quantitative inflation objective gravitating around a level of 2%. Traditional arguments put forward to reconcile theory and practice relate to the role of a positive inflation trend as a buffer against risks of deflation, in the presence of a zero lower bound on nominal interest rates, of downward nominal wage rigidities, or in the case of positive bias in inflation measurement.[3] We propose a novel mechanism that rationalizes positive inflation targets regardless of any buffer consideration, based on production *misallocation*, i.e. the welfare loss that occurs when, with respect to a social optimum, more production is allocated to less productive firms, and less production to the most efficient ones in the economy. The extent and relevance of misallocation has been documented by a burgeoning literature,[4] but standard monetary models do not feature misallocation in steady state.

To introduce this novel mechanism, we extend the model by Edmond, Midrigan, and Xu (2018), characterised by heterogeneous firm productivity and strategic complementarities, by adding price rigidity in the form of quadratic price adjustment costs as in Rotemberg (1982). Strategic complementarities are captured in a parsimonious way with a non-constant elasticity of substitution among goods, as in Kimball (1995). It features a price elasticity of demand that, instead of being constant as in the standard constant elasticity of substitution (CES) case, increases with the firm's relative price, generating a smoothed version of a "kink" in the demand curve faced by a given firm. Indeed it implies that, with respect to the CES case, more consumers flee from individual items with high relative prices, but less flock to inexpensive ones. Hence, the costs of deviating from the profit-maximizing relative price are

[1]In what follows we will use the terms "long-term inflation", "trend inflation" and "inflation target" interchangeably.

[2]Schmitt-Grohé and Uribe (2010) provide an exhaustive review of the robustness of the price stability prescription in this class of models.

[3]See, among others, Issing (2003).

[4]See Restuccia and Rogerson (2017) for a survey.

higher than in a world without strategic complementarities, introducing a form of real rigidity.

As Edmond, Midrigan, and Xu (2018) show in a flexible prices model, the interaction of firm heterogeneity and strategic complementarity begets a *misallocation distortion*, due to an inefficient cross-sectional markup dispersion. More productive firms can charge lower prices, hence securing a larger market share. The strategic complementarities á la Kimball imply that these firms also face a lower demand elasticity for the goods they produce, which translates into a stronger market power. The fact that more productive firms can charge higher markups than less productive firms implies that production is lower than in the social optimum.

We show that there exists a trade-off between the cost of inflation and misallocation. Bilbiie, Fujiwara, and Ghironi (2014) and Ascari and Rossi (2012) argue that in a model à la Rotemberg markups get smaller in presence of positive trend inflation, as firms incur adjustment costs to keep up with inflation, and to pay for them they want to increase revenues and output beyond what would be convenient with flexible prices; this incentive is stronger for higher values of trend inflation, due to convexity of adjustment costs.

Our main modelling contribution is to combine price rigidity, heterogeneous firms and strategic complementarities and show that (at low levels of inflation) this negative relation between markups and trend inflation is reinforced for firms with higher markups. If a firm charges a higher markup, it means that it faces a weaker demand elasticity, which implies that for the same increase in output the price falls by more. Hence, a higher steady state inflation spurs the more productive firms to expand their production, and reduces the markup dispersion, with positive effects for aggregate productivity and welfare. As a result, a positive inflation target is optimal also in an environment where no buffer role is accounted for. A numerical example shows that this misallocation channel alone can contribute to a positive trend inflation by a quantitatively relevant amount, around 1% or more.

A testable implication of our crucial assumptions of firm heterogeneity and strategic complementarities in price setting, is that the pass-through from marginal costs to prices is decreasing in firms' productivity. Baqaee, Farhi, and Sangani (2021) show that our assumptions satisfy the Marshall's third law of demand: markup is increasing and pass-through is decreasing in firm size, which in the model is tightly linked to productivity.[5] To provide empirical support

[5]Their argument is derived in a flexible price model. For shocks that are not too large, it can be extended to our sticky price framework; the proof is available from the authors upon request.

to our model, we resort to a quasi-natural experiment and we analyse the price response of a representative sample of Italian firms subject to an unexpected cut in labour costs following a policy change occurred in Italy in late 2014. In October 2014 the Italian government announced a sharp cut in social security contributions for open-ended hires made the following year. The policy applied both to newly hired workers and to fixed-term workers already employed by the firm if their contract was converted from fixed-term to open-ended. The take up of this policy was extremely high, as shown by Sestito and Viviano (2018), especially for firms already employing a large number of fixed-term workers. In fact, these firms were able to convert fixed-term contracts into open-ended ones without having to incur in search and recruiting costs. We use this policy change as an exogenous variation in labour costs and relate it to firms' price changes. Our empirical exercise is based on a unique dataset, obtained by merging administrative employer-employee data, firms' balance sheets and a survey on industrial and non-financial service firms conducted every year by the Bank of Italy, which reports information on firms' own price change. Together, these data constitute an ideal laboratory for studying the price transmission mechanisms of the cut in labour costs. To analyse the relationship between pass-through and productivity, we rely on proxies of productivity typically used by the literature. First, we consider firms' size, an observable firm characteristic that proxies productivity to the extent to which smaller firms cannot benefit from economies of scale in production (e.g. Haltiwanger, Lane, and Spletzer (1999), Bartelsman and Doms (2000) for a review on productivity dynamics and firm size).[6] Recognizing also that highly productive/large firms pays also higher wages, a very-well documented fact in the literature (e.g. Abowd, Kramarz, and Margolis (1999), Dunne, Foster, Haltiwanger, and Troske (2004), Brown and Medoff (1989)), we also split the sample also by the average wage paid by firms to their workers. The results show that small firms and firms paying lower wages transmit labour cost variation to prices much more than large firms. These results allow us to conclude that less productive firms have a higher pass-through, consistently with our theory.

Our analysis extends the already established positive relation between firm size and *level* of markup (e.g, see Atkin, Chaudhry, Chaudry, Khandelwal, and Verhoogen (2015)), by showing that also the *change* in markups is size-dependent. This result allows us to contribute to a recent and growing strand of literature that shows how a non-zero long-term inflation might

[6]As mentioned above, in the model there is a on-to-one onto correspondence between form size and its productivity.

be optimal, if more realistic supply-side features are taken on board in standard monetary models.

Some papers analyze the effect of disposing of the CES assumption in frameworks with a representative firm. Bilbiie, Fujiwara, and Ghironi (2014) argue that, when departing from the CES case, deviations from long-run price stability are optimal in the presence of endogenous entry and product variety in a sticky-price model, as the preference specification entails a wedge between the benefit for consumers of a new variety and the market incentive for creating that variety. In a model with non-CES demand á la Kimball, Kurozumi and Zandweghe (2020) show that higher steady state inflation curbs the markup of the representative firm, hence increasing welfare and calling for a positive inflation target. We take into account firms' heterogeneity, showing that the ensuing misallocation driven by *markup dispersion* is an independent reason to have a positive inflation trend, on top of the channel that goes through the *aggregate markup* in the economy.

More closely related to our work are Adam and Weber (2019) and Adam and Weber (2022), which develop a framework with heterogeneous firms but keeping the standard CES assumption. They show that heterogeneous firm-level productivity trends imply an optimal steady-state inflation rate that generically differs from zero. In our setup firms have not different *productivity trends* along a balanced growth path, but differ in their *productivity levels* in steady state. In particular, we show that, with strategic complementarities, this kind of heterogeneity is sufficient to obtain a positive optimal steady-state inflation.

We also contribute to the literature on misallocation, characterizing its implication for long-run inflation. Baqaee, Farhi, and Sangani (2021) study instead the business cycle properties of an economy that shares with our model the basic features, with heterogeneous firms, sticky prices, and endogenous markups. They show that, under fairly realistic conditions, a monetary easing endogenously increases aggregate TFP and improves allocative efficiency. This endogenous positive supply shock results in a flattening of the Phillips curve. Meier and Reinelt (2020) show that firms heterogeneity in terms of price rigidity can explain the response to a monetary shock of markup dispersion, which in turn significantly affects aggregate TFP. However, since they assume CES demand, in their model optimal trend inflation is zero.

Finally, we provide novel evidence on the pass-through from labour costs to prices, based on micro data and a quasi-natural experiment, which allows us to explore differences among firms. Most of the existing work builds upon macro data, and hence focuses on variation over time of the aggregate pass-through. Peneva and Rudd (2017) find little evidence in the US of pass-through of independent movements in labour costs into price inflation. Instead for the euro area Bobeica, Ciccarelli, and Vansteenkiste (2019) document a significant link between unit labour cost and inflation, and Conti and Nobili (2019) find that the pass-through from wages to consumer prices net of food and energy is less than unity. They also find that it depends on the nature of the shocks hitting the economy, being higher following aggregate demand shocks. Evidence of moderate pass-through is also found by Conflitti and Zizza (2020) and Carlsson and Skans (2012) using micro data for Italy and Sweden respectively.

The paper is organized as follows. In Section 2 we develop a model characterized by non-CES demand and heterogeneous firms, and show in Section 3 that under flexible prices it features a misallocation distortion. In Section 4 we show that under sticky prices, long-run price stability is not optimal. In Section 5 we describe the data and the quasi-experiment to estimate the pass-through, which turns out to be significantly heterogeneous among firms. Lastly, Section 6 concludes.

2 The model

In what follows we set up a sticky price model featuring firm heterogeneity and strategic complementarity; to model the latter, we allow for variable demand elasticity à la Kimball (1995), as in Riggi and Santoro (2015) and Linde and Trabandt (2019).

The other structural features are standard in the New Keynesian literature: we consider an economy inhabited by a representative consumer with preferences over final consumption and labour supply, and who owns firms. The final good is produced by perfectly competitive firms using a bundle of differentiated intermediate goods, whereas the differentiated inputs are produced by monopolistically competitive firms using labour, and facing sticky prices, due to quadratic price adjustemnt costs, following the formalism proposed in Rotemberg (1982). As a first step, in what follows we take the productivity distribution among firms as given, and do not allow for firms' entry and exit. Moreover, we consider a simple production technology, with labour as the only production factor.

2.1 Representative consumer

We consider a representative household, that has an objective function given by:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ U\left(C_t\right) - V\left(N_t\right) \right\} \tag{1}$$

where C_t denotes aggregate consumption and N_t indicates labour supply.

We specify the household's period utility to be given by:

$$U\left(C_t\right) \equiv \log\left(C_t\right) \tag{2}$$

$$V\left(N_t\right) \equiv \frac{N_t^{1+\varphi}}{1+\varphi} \tag{3}$$

where φ is the inverse Frisch elasticity of labor supply. The period-by-period budget constraint takes the form:

$$P_t C_t + Q_t B_t = B_{t-1} + W_t N_t + \int_0^1 \Xi_t\left(i\right) di$$

where P_t and C_t are the price and the quantity consumed of the final good, W_t is the nominal wage, $\Xi_t\left(i\right)$ are the profits of the firm producing the intermediate good i, Q_t is the price of a one-period riskless bond, paying one unit of currency and B_t denotes the quantity of that bond purchased in period t.

The optimal consumption/savings and labour supply decisions are described by the following conditions:

$$Q_t = \beta \mathbb{E}_t \left(\frac{C_t}{C_{t+1}} \right) \frac{P_t}{P_{t+1}} \tag{4}$$

$$\frac{W_t}{P_t} = N_t^{\varphi} \left[C_t\right] \tag{5}$$

2.2 Final good producers

The final good Y_t is obtained using a continuum of differentiated intermediate goods $Y_t(i)$, for $i \in [0,1]$. We assume that the technology to transform intermediate goods into the final good is a flexible variety aggregator à la Kimball (1995), which allows for the possibility that firms face price elasticity of demand which is increasing in firm's relative prices. Whereas in the Dixit-Stiglitz's preferences the elasticity of substitution between a given variety and others is constant, in Kimball's world the elasticity of substitution between differentiated goods is decreasing in the relative quantity consumed of the variety, implying that the desired markup of a producer of an intermediate good is decreasing in the relative price it charges. The Kimball aggregator is

$$\int_0^1 \psi\left(\frac{Y_t(i)}{Y_t}\right) di = 1 \tag{6}$$

where $\psi(\cdot)$ is an increasing, strictly concave function. The standard CES technology is nested within this specification: the Kimball aggregator reduces to the Dixit-Stiglitz when $\psi\left(Y_t(i)/Y_t\right) = \left(Y_t(i)/Y_t\right)^{\frac{\epsilon-1}{\epsilon}}$.

Firms that produce the final good are perfectly competitive: they take the prices $P_t(i)$ of the inputs and the price P_t of the output as given, and choose how much to produce Y_t and the quantities $Y_t(i)$ of intermediate goods to buy, in order to maximize profits

$$P_tY_t - \int_0^1 P_t(i)\, Y_t(i)\, di$$

subject to the Kimball aggregator (6). The solution to the final good producers' problem yields the set of demand equations:

$$P_t(i) = \psi'\left(\frac{Y_t(i)}{Y_t}\right) P_tD_t \tag{7}$$

where:

$$D_t \equiv \left(\int_0^1 \psi'\left(\frac{Y_t(i)}{Y_t}\right)\frac{Y_t(i)}{Y_t} di\right)^{-1}$$

and the zero-profit condition:

$$\frac{\int_0^1 P_t(i)\, Y_t(i)\, di}{Y_t} = P_t. \tag{8}$$

In what follows, we use for the aggregator ψ the functional form introduced by Klenow and Willis (2016) and adopted, among others, by Edmond, Midrigan, and Xu (2018), Gopinath and Itskhoki (2010) and Gopinath and Itskhoki (2011). It results in:

$$\psi'\left(\frac{Y_t(i)}{Y_t}\right) = \frac{\sigma-1}{\sigma}\exp\left(\frac{1-\left(Y_t(i)/Y_t\right)^{\frac{\epsilon}{\sigma}}}{\epsilon}\right)$$

This demand specification is conveniently governed by two parameters, $\sigma > 1$ and $\epsilon \geq 0$; the demand elasticity to prices is decreasing in $\frac{Y_t(i)}{Y_t}$, and equal to:

$$\tilde{\sigma}\left(\frac{Y_t(i)}{Y_t}\right) \equiv -\frac{\partial Y_t(i)}{\partial P_t(i)}\frac{P_t(i)}{Y_t(i)} = -\frac{\psi'\left(\frac{Y_t(i)}{Y_t}\right)}{\psi''\left(\frac{Y_t(i)}{Y_t}\right)\frac{Y_t(i)}{Y_t}} = \sigma\left(\frac{Y_t(i)}{Y_t}\right)^{-\frac{\epsilon}{\sigma}}.$$

If we set $\epsilon = 0$ in the above equation, we are back in the CES world: the demand elasticity is constant and equal to σ.

2.3 Intermediate goods producers

There is a continuum of differentiated intermediate goods distributed uniformly on the unit interval. Each good is produced by a single firm in a monopolistically competitive market. Production requires only one factor, labour; the production function of firm i is

$$Y_t(i) = Z_t(i)N_t(i) \tag{9}$$

where $Z_t(i)$ is a technology shock. Firms face nominal rigidity in the form of a quadratic cost of adjusting prices over time as in the Rotemberg (1982) formulation: the cost of adjusting prices is given by $\frac{\phi}{2}\left(\frac{P_t(i)}{P_{t-1}(i)}-1\right)^2 P_t(i)\,Y_t(i)$, where the parameter ϕ measures the severity of nominal rigidities. Hence, the firm optimization problem at time t becomes:

$$\max_{P_t(i)} E_t\sum_{k=0}^{\infty}\left\{Q_{t,t+k}\left(P_{t+k}(i)\,Y_{t+k}(i) - W_{t+k}N_{t+k}(i) - \frac{\phi}{2}\left(\frac{P_{t+k}(i)}{P_{t-1+k}(i)}-1\right)^2 P_{t+k}(i)\,Y_{t+k}(i)\right)\right\} \tag{10}$$

subject to the demand curve (7) and to the production function (9). Simple algebra yields the following optimality condition[7]:

$$P_t(i) = \frac{\tilde{\sigma}\left(\frac{Y_t(i)}{Y_t}\right)}{\left(\tilde{\sigma}\left(\frac{Y_t(i)}{Y_t}\right)-1\right)\left[1-\frac{\phi}{2}\left(\frac{P_t(i)}{P_{t-1}(i)}-1\right)^2\right]+\phi\frac{P_t(i)}{P_{t-1}(i)}\left(\frac{P_t(i)}{P_{t-1}(i)}-1\right)-\Psi_t(i)}\frac{W_t}{Z_t(i)} \tag{11}$$

where:

$$\Psi_t(i) \equiv \phi E_t\left\{Q_{t,t+1}\left(\frac{P_{t+1}(i)}{P_t(i)}\right)^2\left(\frac{P_{t+1}(i)}{P_t(i)}-1\right)\frac{Y_{t+1}(i)}{Y_t(i)}\right\} \tag{12}$$

Hence, the pricing rule is given by a markup over marginal costs.[8] Two features are worth emphasizing. First of all, this markup depends positively on the *relative quantity* $\frac{Y_t(i)}{Y_t}$ produced by firm i, differently from the Dixit-Stiglitz case. Since firms with a higher relative quantity face a smaller elasticity of demand, they can charge a higher markup with a less severe loss of demand. Second, the markup takes into account *the cost of current and future price changes*, due to the quadratic adjustment cost assumption. If we define the (gross) inflation rate at firm level as $\Pi_t(i) \equiv \frac{P_t(i)}{P_{t-1}(i)}$ and the relative quantity produced by firm i as $y_t(i) \equiv \frac{Y_t(i)}{Y_t}$, the markup can be written as follows:

$$\mu\left(y_t, \Pi_t(i), \Psi_t(i)\right) \equiv \frac{\tilde{\sigma}\left(y_t(i)\right)}{\left(\tilde{\sigma}\left(y_t(i)\right)-1\right)\left[1-\frac{\phi}{2}\left(\Pi_t(i)-1\right)^2\right]+\phi\Pi_t(i)(i)\left(\Pi_t(i)-1\right)-\Psi_t(i)} \tag{13}$$

3 Flexible-price equilibrium

Before studying our model, it is useful to review the efficiency properties of its flexible-price counterpart, obtained setting $\phi = 0$.[9] With flexible prices, the model becomes essentially static, but for the sake of clarity we maintain the time-indexation of variables. Firm i solves the following problem:

$$\max_{Y_t(i)} P_t(i)\, Y_t(i) - W_t N_t(i) \tag{14}$$

[7] In what follows we slightly abuse of notation, indicating with $P_t(i)$ and $Y_t(i)$ also the optimal values of these variables.

[8] It is easy to show that marginal costs of firm i are given by $\frac{W_t}{Z_t(i)}$.

[9] The model with flexible prices boils down to a version of the framework developed in Edmond, Midrigan, and Xu (2018).

subject to the demand curve (7) and to the production function (9). Taking the FOC and rearranging, we get the following optimality condition:

$$P_t(i) = \frac{\tilde{\sigma}(y_t(i))}{\tilde{\sigma}(y_t(i)) - 1} \frac{W_t}{Z_t(i)} \tag{15}$$

As in the sticky price case, the firm sets the price as a markup over the marginal cost; differently from before, markup $\mu^{FP}(y_t)$ depends only on $y_t(i)$, due to the Kimball assumption, but not on price changes, i.e.:

$$\mu^{FP}(y_t) \equiv \frac{\tilde{\sigma}(y_t(i))}{(\tilde{\sigma}(y_t(i)) - 1)} \tag{16}$$

To relate the cross-sectional distribution of relative quantities and markups to aggregate variables, it is useful to introduce the aggregate productivity $\mathcal{Z}_t$ and the aggregate markup $\mathcal{M}_t$. The former relates the total amount of output with the total amount of input

$$\mathcal{Z}_t \equiv \frac{Y_t}{N_t}$$

while the latter is the wedge between aggregate costs and aggregate revenues:

$$\mathcal{M}_t \equiv \frac{P_t Y_t}{W_t N_t}$$

$\mathcal{Z}_t$ and $\mathcal{M}_t$ are the key variable that link the distribution of individual firms' decisions to aggregate variables: as in Edmond, Midrigan, and Xu (2018), we can express $\mathcal{Z}_t$ and $\mathcal{M}_t$ in terms of the distribution of relative quantities and markups of firms. Let $y(Z)$ and $\mu^{FP}(y(Z))$ denote the relative quantity and the markup of the firm i that has productivity Z; it turns out that aggregate productivity is an harmonic average of individual productivities, weighted by relative quantities:

$$\mathcal{Z}_t = \frac{Y_t}{N_t} = \frac{Y_t}{\int_\Omega N(Z)\, dH(Z)} = \left(\int_\Omega \frac{y(Z)}{Z} dH(Z)\right)^{-1} \tag{17}$$

where $H(Z)$ is the CDF of the distribution of productivities and $\Omega \equiv [\underline{Z}, \overline{Z}]$ is the support of

H. Aggregate markup is a weighted arithmetic average of individual markups:

$$\begin{aligned}\mathcal{M}_t &= \frac{P_t Y_t}{W_t N_t} = \frac{\int_\Omega P(Z)\, Y(Z)\, dH(Z)}{W_t N_t} \\ &= \frac{\int_\Omega \mu^{FP}(y(Z))\, \frac{W_t}{Z} Y(Z)\, dH(Z)}{W_t N_t} = \frac{\int_\Omega \mu^{FP}(y(Z))\, \frac{y(Z)}{Z} dH(Z)}{\int_\Omega \frac{y(Z)}{Z} dH(Z)} \end{aligned} \tag{18}$$

Combining the definitions of $\mathcal{Z}$ and $\mathcal{M}$, we get that $\frac{W_t}{P_t} = \frac{\mathcal{Z}_t}{\mathcal{M}_t}$; moreover, we can use the demand curve (7) to substitute out $P_t(i)$ from (15). As a result, equilibrium allocations are given by the set of relative quantities of intermediate goods $\{y_t(Z)\}_{Z\in\Omega}$, consumption and production of the final good C_t and Y_t, and labour N_t that satisfy the following equations

$$Y_t = \mathcal{Z}_t N_t \tag{19}$$

$$N_t^{\varphi} C_t = \frac{\mathcal{Z}_t}{\mathcal{M}_t} \tag{20}$$

$$C_t = Y_t \tag{21}$$

$$\psi'(y_t(Z)) = \frac{\mu^{FP}(y_t(Z))}{\mathcal{M}_t} \frac{1}{D_t} \frac{\mathcal{Z}_t}{Z} \tag{22}$$

where $\mathcal{Z}$ and $\mathcal{M}$ are given by (17) and (18).

The flexible-price equilibrium is not efficient, because the interaction of monopolistic competition and Kimball technology gives rise to an inefficient allocation of resources. To delve more into this issue, we solve the problem of a benevolent planner who faces the same technological and resource constraints as in the decentralized economy, and contrast it with the flexible-price equilibrium.

The planner chooses C_t, N_t and $\{y_t(Z)\}_{Z\in\Omega}$ to maximize:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \log(C_t) - \frac{N_t^{1+\varphi}}{1+\varphi} \right\}$$

subject to the Kimball aggregator

$$\int_\Omega \psi(y_t(Z))\, dH(Z) = 1$$

and to the resource constraint

$$C_t = \mathcal{Z}_t N_t$$

Let λ_t^1 and λ_t^2 be the Lagrange multiplier on the Kimball aggregator and the resource constraint, respectively; then the FOC are as follows

$$\frac{1}{C_t} = \lambda_t^2$$
$$N_t^{\varphi} = \mathcal{Z}_t \lambda_t^2$$
$$\lambda_t^1 \psi' \left(y_t \left(Z\right)\right) = \lambda_t^2 N_t \frac{1}{Z} \mathcal{Z}_t^2$$

Simple manipulations, and the use of the resource constraint, yield

$$N_t^{\varphi} C_t = \mathcal{Z}_t$$
$$\lambda_t^1 \psi' \left(y_t \left(Z\right)\right) = \frac{\mathcal{Z}_t}{Z}$$

To substitute out λ_t^1, we can multiply both sides of the last equation by $y_t\left(Z\right)$ and integrate over Z, to obtain $\lambda_t^1 = D_t$, which implies

$$N_t^{\varphi} C_t = \mathcal{Z}_t \tag{23}$$
$$\psi' \left(y_t \left(Z\right)\right) = \frac{1}{D_t} \frac{\mathcal{Z}_t}{Z} \tag{24}$$

Comparing these equations with the flexible-price counterparts, we recover the existence of two distortions already emphasized in Edmond, Midrigan, and Xu (2018). First, the *aggregate markup* distorts the intratemporal choice between consumption and leisure, inducing the representative household to consume and work less than what is socially optimal, for any level of aggregate productivity:

$$\mathcal{Z}_t > \frac{\mathcal{Z}_t}{\mathcal{M}_t}$$

Second, there is a *misallocation distortion*, due to an inefficient cross-sectional markup dispersion: more productive firms charge higher markups than less productive ones, hence producing too little compared to the social optimum, lowering social welfare. As a result, for any level of aggregate productivity and aggregate markup, relative quantities in the decentralized economy are different than in the planner solution:

$$\frac{1}{D_t} \frac{\mathcal{Z}_t}{Z} \neq \frac{\mu^{FP}\left(y_t\left(Z\right)\right)}{\mathcal{M}_t} \frac{1}{D_t} \frac{\mathcal{Z}_t}{Z}$$

The first distortion would arise also with homogeneous firms and CES preferences, and a well-known result in New Keynesian models shows that appropriately subsidizing labour (or sales of intermediate goods) can undo this distortion. Instead, the misallocation distortion originates from the interplay of heterogeneous firms and Kimball technology, that translate in heterogeneous markups; absent either of these features, we would have $\mu^{FP}(y_t(Z)) = \mathcal{M}_t$ for any Z.

Edmond, Midrigan, and Xu (2018) derive a firm-specific production subsidy that decentralizes the efficient allocation, eliminating both distortions. In the next Section we consider a second-best environment (where such a sophisticated subsidy is not available) and investigate if, in a sticky price version of the model, the central bank would prefer a positive trend inflation to curb the misallocation distortion.

4 Optimal trend inflation

In this Section we present the steady state of the sticky price model developed in Section 2, and investigate whether price stability (i.e., zero inflation in steady state) is optimal. Our welfare metric is the steady state utility of the representative consumer, which is function only of aggregate variables. We concentrate on a steady state where relative quantities produced by firms is constant; this implies that each firm's price increases at the same rate as the aggregate prices, otherwise the firms whose goods become relatively cheaper over time would drive the others out of the market. As a result, markups and, in turn, relative quantities depend on trend inflation Π. Let steady state productivity be distributed according to the CDF H, then the relative quantity and the markup of a firm with productivity Z is denoted as $y(Z;\Pi)$ and $\mu(y(Z;\Pi),\Pi)$, respectively. [10]

For any value of trend inflation Π, we can define aggregate productivity and aggregate markup analogously to the flexible-price case: $\mathcal{Z}(\Pi)$ relates the total amount of output with the total amount of input, while $\mathcal{M}(\Pi)$ is the wedge between aggregate costs and aggregate

[10] With a slight abuse of notation we suppressed the dependence of μ on the function Ψ defined in equation (12), because in steady state the latter depends only on Π.

revenues, and can be expressed as:

$$\mathcal{Z}(\Pi) = \left(\int_{\Omega} \frac{y(Z;\Pi)}{Z} dH(Z)\right)^{-1} \tag{25}$$

and:

$$\mathcal{M}(\Pi) = \frac{\int_{\Omega} \mu(y(Z;\Pi),\Pi) \frac{y(Z;\Pi)}{Z} dH(Z)}{\int_{\Omega} \frac{y(Z;\Pi)}{Z} dH(Z)} \tag{26}$$

Aggregate productivity depends on Π since firms are heterogeneous: if they were homogeneous, in a symmetric equilibrium $\mathcal{Z}$ would be an exogenous constant. Instead, aggregate markup would depend on Π also in a setup with homogeneous firms: as emphasized in Bilbiie, Fujiwara, and Ghironi (2014), in a sticky price model markups are affected by inflation, since firms' margins are curbed by inflation cost.

For any value of Π, aggregate output, labour and consumption satisfy the following equations, where we drop time indices to denote that variables are evaluated in steady state:

$$Y = \mathcal{Z}(\Pi) N \tag{27}$$

$$N^{\varphi} C = \frac{\mathcal{Z}(\Pi)}{\mathcal{M}(\Pi)} \tag{28}$$

$$C = Y\left(1 - \frac{\phi}{2}(\Pi - 1)^2\right) \tag{29}$$

where the first equation is the definition of the aggregate production function, the second one is the intratemporal consumption/leisure optimality condition of the representative consumer, and the third one is the resource constraint, obtained putting together the consumer's budget constraint with intermediate firms' profits (10) (and using the zero profit condition of final goods producers (8)).

In this model there are three distortions that affect welfare. The first two are inherited from the flexible-prices model: the *aggregate markup* and the *misallocation distortion*; on top of them, there is also the *cost of inflation* stemming from the Rotemberg menu cost assumption. The aggregate markup and the cost of inflation would arise also with homogeneous firms and CES preferences. Instead, to have the misallocation distortion we need both heterogeneous firms and Kimball preferences, that translate in heterogeneous markups.

In the absence of misallocation, we recover the standard result in New Keynesian models that zero inflation steady state replicates first best, provided a labour (or sales) subsidy is used to offset the aggregate markup distortion (see Galí (2008) for details). In fact, if we had no misallocation (i.e., homogeneous markups), and a labour subsidy τ chosen by the government as a function of Π, we could rewrite the steady state equilibrium as follows:

$$\begin{aligned} Y &= \mathcal{Z}(\Pi)\, N \\ N^{\varphi} C &= \frac{\mathcal{Z}(\Pi)}{\mathcal{M}(\Pi)\,(1-\tau(\Pi))} \\ C &= Y\left(1-\frac{\phi}{2}(\Pi-1)^2\right) \end{aligned}$$

If we set:

$$\tau(\Pi) = 1-\frac{1}{\mathcal{M}(\Pi)}, \qquad \Pi = 1$$

then the above equations collapse to the equilibrium that would be commanded by the social planner, shown in Section 3.

Instead, in the presence of misallocation, a trade-off emerges, even in presence of a $\tau(\Pi)$ that offsets the aggregate markup distortion: increasing Π from one entails costs as in any sticky price model, while we show that it begets a reduction in the misallocation distortion. The intuition of the latter effect can be traced back to the result that an increase of trend inflation affects markups differently, depending on the demand elasticity to prices. For a given increase in output, the smaller the demand elasticity, the more pronounced the price fall. As a consequence, the increase in revenues and output necessary to pay for the price-adjustment costs necessary to keep up with trend inflation eats more firm's profits and markup. To sum up, a given increase in Π reduces the markup more, the smaller is the demand elasticity.

With Kimball preferences, demand elasticity is decreasing in relative quantities which, in turns, is proportional to productivity; as a result, increasing Π above one reduces markups of more productive firms relatively more than those of less productive firms; this reduction of markups is reflected into lower relative prices, triggering a shift of demand from the less productive to the more productive firms that curbs misallocation.

Hence, we have the following result:

Proposition 1. *If technology is given by the Kimball aggregator* (6)*, firms are heterogeneous in productivity and face quadratic price adjustment costs, a positive trend inflation is optimal.*

Proof. See the Appendix A1. □

4.1 Quantitative results

We quantify the optimal long-run rate of inflation by means of a numerical example. In the simulations below, the intertemporal discount factor β is 0.99, φ is 0.75, equivalent to a labour elasticity of 4, as in Bilbiie, Fujiwara, and Ghironi (2014), and the price adjustment cost parameter ϕ is 200, as estimated for the Euro area by Forni, Gerali, Notarpietro, and Pisani (2015).[11]

We argued above that the misallocation distortion, and the ensuing positive optimal inflation trend, is generated by heterogeneous productivity and strategic complementarities. As regards the former feature, we posit that firms' productivity in steady state is distributed as a Pareto, and compute the optimal inflation as the shape parameter θ of the distribution varies in the range 1-2, which includes the values estimated by Ottaviano, Taglioni, and di Mauro (2007) for manufacturing sectors in European countries. As regards the strength of strategic complementarities, it is captured by the ratio $\frac{\epsilon}{\sigma}$, which determines how much the demand elasticity to prices depends on relative quantities; we show results with $\frac{\epsilon}{\sigma} = \frac{1}{4}$ and $\frac{\epsilon}{\sigma} = \frac{1}{3}$, which are inside the range of values for $\frac{\epsilon}{\sigma}$ considered in Edmond, Midrigan, and Xu (2018).

In the following simulations we focus on the effects of the misallocation distortion on the optimal trend inflation. Hence, we want to consider only how changes in trend inflation affects welfare through changes in markups dispersion, and not through movements in the aggregate markup. To do so, we assume that there exists a labour subsidy $\tau(\Pi)$ that, for each value of trend inflation, fully offsets the wedge in consumption-labour decisions given by the aggregate markup (see the previous Section for more details).

In Figure 1 we plot the optimal long-run annual inflation as θ ranges from 1 to 2. When $\frac{\epsilon}{\sigma} = \frac{1}{4}$, it decreases from 1.5 to 0.7%; as a higher θ corresponds to a Pareto distribution more concentrated around the lower bound of its support, the graph tells us that a less dispersed productivity distribution calls for a smaller optimal inflation in steady state.[12] When $\frac{\epsilon}{\sigma} = \frac{1}{3}$

[11]Empirical evidence on the value of this parameter provided quite dispersed estimates; hence, we run a robustness exercise setting it to 50, more in line with what is found in Gerali, Neri, Sessa, and Signoretti (2010), and the results are basically unaffected.

[12]Changing the θ parameters affects not only the variance of the Pareto distribution, but also its mean; hence, we cannot rule out that the dependence of Π on θ might be non-monotonic for some regions of the latter.

we have a similar decreasing pattern, but with higher values for the optimal Π, due to stronger strategic complementarities.

5 Testing heterogeneity in pass-through: firm-level evidence

We present micro evidence to support the hypothesis of heterogeneity in pass-through based on differences in firm productivity. In particular, we analyse a shock to labour costs in Italy and its effect on firm prices. One of the main challenges in the identification of the pass-through is reverse causality of price changes to labour costs, since firms can always modify the composition of their workforce (and ultimately their labour costs) as a reaction to changes in own prices (present or planned). To identify the causal relationship between labour costs and prices we rely on a quasi-experimental setting based on a cut of labour costs in Italy announced at the end of 2014 and implemented in 2015. We consider a shock to labour costs because, for a given level of TFP, it maps into a change of unit labour costs, which is the key cost variable in the firms' pricing equation. We proxy heterogeneity in productivity with firm size and average wage, under the alternative assumptions that larger firms face higher economies of scale in production or that firms paying higher wages are also those with higher productivity (as in Dunne, Foster, Haltiwanger, and Troske (2004), Abowd, Haltiwanger, Jarmin, Lane, Lengermann, McCue, McKinney, and Sandusky (2005), Haltiwanger, Lane, and Spletzer (1999) Abowd, Kramarz, and Margolis (1999) among others). [13] We first describe the features of the Italian labour market and the policy innovation that we use for identification. Then we present the empirical results.

5.1 The policy

The Great Recession and the sovereign debt crisis led to a significant number of job losses in Italy (around 1 million between 2008 and 2014). Because of the dualism of the Italian labour market (i.e. segmentation between open-ended and fixed-term job contracts), job losses were mainly concentrated among fixed-term workers, who could be fired at no cost for firms. Aimed at reducing dualism and stimulating open-ended employment at the beginning of 2015 the

[13] Unfortunately, our identification strategy does not allow us to evaluate the relevance of the hypothesis about strategic complementarities (see e.g. Amiti, Itskhoki, and Konings (2019) for an empirical assessment).

Italian government introduced a very generous hiring subsidy for that type of contract. [14]

The hiring subsidy was announced in late 2014, when the Budget law was presented to the Parliament. It was a tax rebate paid to firms for new open-ended job contracts signed between January 1 and December 31, 2015. It covered both hirings and conversions of job contracts from fixed-term into open-ended. Eligibility criteria were quite relaxed, as all workers not employed with an open-ended job contract could be potentially eligible. The amount of the subsidy consisted of an exemption from social security contributions for a duration of three years since hiring/conversion (around a 30% reduction of total labour costs for new hires, on average).[15] The subsidy was discontinued in December 2015 and substituted in 2016 by a new subsidy amounting to 40% of total social security contributions for two years.

Thus, Italian firms hiring open-ended workers in 2015 (and partly in 2016) could benefit for a substantial reduction in labour costs for the period 2015-2018. We study whether this labour cost change affected price dynamics in the following years and whether this effect was heterogeneous by firm size and wage.

5.2 The data and some evidence

In this paper we rely on a unique firm-level dataset obtained by merging three data sources. The first is the Bank of Italy's yearly survey on industrial and non-financial service firms (INVIND). Around 4,000 firms with a minimum firm size of 20 employees are interviewed each year and provide data on the average yearly change in their output prices (together with data on sales, investments, etc.). The second dataset comes from the Italian National Social Security Institute (INPS). INPS manages social security payments for all private sector firms in Italy. Firms are required to provide information about their workforce with a monthly frequency. INPS provides the Bank of Italy with complete working histories for all the workers in firms included in the INVIND sample (also if they move to another firm not in the INVIND sample). The INPS dataset includes in particular the starting and end date of each job contract, the

[14] The subsidy was introduced also to foster the adoption of a new type of open-ended contract characterised by lower firing costs (introduced in March 2015 by the so-called Jobs Act reform). We disregard the impact of the Jobs Act, as firing costs entail a cut in labour costs only when firms decide to fire the worker and it is very unlikely that this cost cut translates into a price reduction. See Sestito and Viviano (2018) for and evaluation of the overall impact of the hiring subsidy and the Jobs Act.

[15] There was an upper limit, equal to 8,060 euro per year. At least 80% of new hires, however, were below the thresholds.

contract type, the wages, the number of days worked per year, i.e. all the information needed for determining social security payments. Based on INPS data we can also calculate average (gross) nominal daily wages and social security contribution paid by firms for each worker. The third dataset is CERVED, the Italian archive of all companies' balance-sheet. In particular, from the merge with CERVED we obtain information on firms' value added. We use value added to control for differences in the labour share, under the reasonable assumption that the pass-through from labour costs to prices can depend on the share of labour costs in total (nominal) value added. [16] As already mentioned above we proxy productivity with firm size and average wage paid to workers.

Figure 2 - panel (a) reports the number of subsidized hirings per month in 2015. It shows that subsidized hirings increased gradually from January to March, then stabilized until November 2015. In December, i.e. close to the deadline to access the subsidy, they peaked.

Sestito and Viviano (2018) explain this trend by showing that firms in the first months of 2015 hired already known workers, previously employed as fixed-term workers in the same firm. In the following months, instead, firms preferred to follow a two-step strategy: they first hired new fixed-term workers and tested them. Then firms converted fixed-term contracts into open-ended ones, but only in case of a good match. This strategy allowed firms to minimize the risk of hiring permanently badly-matched workers and to cash the subsidy anyway, as it was paid also in the case of contract conversions. This behaviour is at the base of our identification assumption which we discuss below.

Figure 2 - panel (b) plots the number of workers under the subsidy regime from 2015 to 2018. Their number peaked in January 2016 and then declined because of normal turnover within each firm (mainly driven by voluntary separations).

Figure 3 reports the evolution of open-ended hirings and conversions over a longer period independently on the subsidy (since 2011; panel (a)). It shows that after the decline observed from 2011 to 2014 the number of hires increased again in 2015 and then declined when both the 2015 and 2016 subsidies were suspended. The evolution of labour costs per worker (i.e. wages, salaries and social security contributions paid for the workforce divided by the number of workers) is reported in panel (b). It confirms the slowdown in labour costs after 2015.

[16] We do not use CERVED, however, to recover measures of productivity as book-value of firms' capital stock is poor proxies for physical capital to be used in the estimation of the production function.

Figure 4 reports the delta log of labour costs per worker and prices changes (between time t and time $t + 1$). Panel (a) reports the time series based on our micro sample. The micro correlation (between social security contributions -SSC henceforth- paid at time t and price changes between time t and $t + 1$) is instead reported in panel (b) and it is almost equal to zero. We argue that the correlation in microdata is so low because of the endogeneity of price changes labour costs as firms can adjust workforce composition to endogenously lower labour costs and prices.

5.3 Identification

To identify the relationship between labour costs and prices we use a key characteristics of the policy described in section 5.1. As already mentioned, the subsidy applied to both previously non-employed workers who got hired on an open-ended basis and to fixed-term workers already employed in the firm whose contract got converted into an open-ended one.[17]

The policy discussed in section 5.2 implies that firms that were already employing fixed-term workers at the time of the implementation of the subsidy benefited the most from it, because they could convert fixed-term contracts into open-ended ones, without any additional recruiting cost and time spent in searching for them. We use this idea for our identification strategy.

We calculate the number of fixed-term workers from January to October 2014, i.e. when the law introducing the subsidy was presented to the Parliament. The duration of fixed-term contracts can be highly heterogeneous, ranging from just a few days to many months per year (1.6 months on average in our sample). Here we consider the number of workers with a fixed-term job contract in a firm, independently on the duration of their contracts, under the assumption that the larger the pool of workers already known by the firm, the larger the number of matches that the firm could potentially form in 2015 with no (or very low) additional search cost. We then normalize this number by the effective size of the firm, calculated as the number of workers in 2014 weighted by the number of days employed in the firm. This variable is our instrumental variable (called IV henceforth) used as an exogenous shifter for firms' labour costs from 2015 onward. Note that alternative strategies are also

[17] Eligibility could also be manipulated if workers under open-ended contracts transit into fixed-term contracts (in a new firm) or into unemployment for six months.

possible, for instance using the number of fixed-term workers over a shorter time horizon, for instance the semester March-October 2014. While the main results remain almost unchanged (and are available upon request), shorter time intervals capture also seasonality in some sectors and estimates are less precise.

It is instead important to limit the time interval to October 2014. In this way we exclude the case that firms postponed hirings from November and December 2014 to January 2015 to cash the subsidy.

The cumulative distribution of IV is reported in Figure 5. Around 20% of firms had no fixed-term workers in 2014, whereas around 5% had at least 20% fixed-term workers. The characteristics of firms with and without fixed-term workers can be potentially different. To check for this potential threat to identification, Table 1 reports the main characteristics of firms with IV above and below the mean. We look at firm size, the labour share (directly affecting the pass-through from labour costs to prices), social security contributions per worker (which depend not only on the number of days worked but also on wages), the average daily wage and producer price inflation. All the variables are measured in 2014, i.e. before the implementation of the subsidy.

Indeed, given the characteristics of the sample (firms with at least 20 employees) the differences between the variables are very small and/or insignificant with the exception of size, being firm size lower for firms with higher share of fixed-term workers. To control for this source of heterogeneity and for differences in sector and geographical areas, we match observations with IV below the average with those with IV above the average. Propensity score matching is performed in 2014 and it is based on the following set of variables: sector (2-digit Nace-rev2) labour share (value added divided by firm size), average daily wage, region where the firm is located (20 regions). After matching (columns 3 and 4) also the difference in firm size is not statistically different from zero.

We then focus on the relationship between price changes and the component of labour costs affected by the policy innovation, i.e. SSC, that correspond to a constant fraction of hourly wages (around 30% in the private sector). The distribution of our main variables, namely the logarithm of SSC per worker and price changes are reported in Figure A2.1 in Appendix 2.

5.4 Results

5.4.1 Pass-through: Full sample

Our main identification assumption is that firms with few fixed-term workers at the time of the announcement of the policy had less access to the subsidy than firms already employing fixed-term workers at that time, because the latter had less need to recruit new workers with fixed-term job contracts, test them and then convert their contract.

We expect then that the (normalized) number of fixed-term workers negatively correlates with social security contributions per worker paid by firms. We then estimate, within a diff-in-diff framework, the following model:

$$y_{f,t} = \beta_t IV_f * \gamma_t + \gamma_f + \beta_f * labourshare_{f,t} + u_{f,t} \tag{30}$$

where the variable $y_{f,t}$ is the outcome variable in firm f at time t. It is regressed as a function of IV, which is time-invariant, interacted by time dummies, γ_t. This specification allows to estimate the effect of the variable IV over time. We use data from 2011 to 2018, i.e. before and after the inception of the policy in 2015. Our specification then implies that we should find not statistically significant β_ts before 2015 and significant values afterwards (the so-called parallel trend assumption, necessary for identification in a diff-in-diff framework as the one adopted here).

Other controls are a time-varying measure of the firm labour share, equal to the ratio between labour costs and value added and firm fixed effects to control for long run unobserved characteristics. The inclusion of both $labourshare_{f,t}$ and firm fixed effects allows us to control not only for the long-run labour shares (through fixed effect) but also for its fixed-term deviations (through the time-varying variable).

First, we check the validity of our identification assumption and we analyse the relationship between our IV and SSC per worker (panel (a) of Figure 6). The estimated coefficients (the β_ts) are reported in the y-axis. They are normalized with respect to the the value estimated for 2012 (used as reference). The 95% confidence intervals are represented by the vertical bars. Before 2015 no differences can be detected in the SSC paid by firms with different values of IV. The parallel-trend assumption is then empirically supported. In 2015 instead the share of fixed-term contracts from January to October 2014 positively affects SSC paid since 2015. This first piece of evidence confirms the results of Sestito and Viviano (2018) and the validity of our approach. [18]

Then, we look at price changes between time t and $t+1$. The results of this exercise are reported in Figure 6- panel b. Price dynamics are relatively slower after the decrease in social security contributions, but the effect is statistically significant at 10% only in the full sample. Since the estimates reported in panel (a) can be interpreted as a first stage (the impact of IV on SSC) and those in panel (b) as a reduced form (where SSC changes are substituted by the IV), it follows that in 2015 the pass-through of a change in social security contributions was around 1/6 (derived by dividing the average of the βs after 2014 in panel (b) by the corresponding values in panel (a)). This estimate is fully consistent, for instance, with Conflitti and Zizza (2020) for Italy, Carlsson and Skans (2012) for Sweden.

5.4.2 Pass-through: by firm size and wage

We now investigate possible heterogeneity in firm size and split the sample according to this characteristic (below and above the median size in 2014). Unfortunately our sample refers only to 20+ firms and we cannot analyse the pass-through of very small firms. As stated above, we use firm size as a proxy for productivity, given the theoretical and empirical relationships between the two variables. The estimates reported in Figure 7 support the validity of our macro model. The impact of our IV on social security contributions shows no difference by firm size. Instead, when we look at price changes we find a statistically significant evidence of pass-through in smaller firms and zero pass-through in larger ones. The aggregate pass-through, estimated to be equal to roughly .17, is entirely determined by smaller firms, whose pass-through is around .35.

Then, for robustness, we look at heterogeneity by firm average wage, another proxy of productivity. We look at daily wages in 2014 and we regress them on observable characteristics which can affect wages independently on productivity. Control variables are firm size, as larger firms pay higher wages than smaller ones, sector, geographical areas (regions) to capture the

[18] One may argue that SSC per workers decline also because firms preferred to substitute highly paid workers with cheaper ones after the introduction of the law. This is in general possible but would be in constrast with the evidence of Sestito and Viviano (2018) finding the firms increased their propensity to hire already tested workers.

impact of heterogeneous local labour market conditions and the labour share. We then take the residuals of this regression, under the assumption that firms with positive residuals (i.e. paying higher wages) are those that are also more productive. The results of this exercise are reported in Figure 8 and confirm that firms paying lower wages are also those with higher pass-through. The size of the estimated pass-through, equal to .25, is just slightly lower than one estimated by firm size.

As a further check we perform a falsification test. We run the same specification as in Equation 30 but only for the period 2011-2013, when the policy was neither in place nor announced. We define years 2012 and 2013 as post-periods and we estimate the impact of our IV on labour costs. We find that the variable IV does not impact on labour costs when the law was not in place indirectly confirming the robustness of our identification strategy (results are available upon request).

6 Conclusions

In this paper we show what are the implications of misallocation on the optimal long-run inflation. We set up a sticky price model with heterogeneity in firms productivity and strategic complementarities as in Kimball (1995). Due to strategic complementarities, large firms have higher markups in steady state, and as a consequence have more room to actively play with them in order to maintain their market shares. As more productive firms get larger in this setup, they have higher markups, hence produce too little compared to the social optimum, with a drag on aggregate productivity. Under sticky prices an increase in steady state inflation curbs the markups, especially those of the more productive firms, hence attenuating the inefficient dispersion of markups. At low levels of inflation, the gains from this reduction in misallocation outweigh the cost of inflation, hence long-run price stability is not optimal.

To lend support to our key assumptions, we provide new evidence on the pass-through from labour costs to prices, and how it depends on firm size and wages, two variables that the literature typically recognizes to be associated to productivity. To overcome the difficulties in extracting an exogenous change in labour costs, we resort to a quasi-natural experiment, related to a large social contribution cut announced in late 2014 on all new open-ended contracts, starting from the following year. We show that the pass-through on average is incomplete, and that it is much higher for smaller/low-paying/less productive firms, consistent with what our theory would predict.

We also believe that this result is interesting in itself because policies introducing incentives to hire workers are often used in European countries to support employment growth. We not only show that such policies result in lower prices, but that the magnitude of this effect depends on the structural feature of the economy. Our results also show that pass-through is incomplete at the firm level and a fortiori so at the macro level.

Overall, our results underline the importance of heterogeneity in firms' productivity for the transmission of nominal shocks to prices, and for the design of the optimal monetary policy. However, while consumer heterogeneity has been a subject of active research in recent macroeconomics, firm heterogeneity has been less explored. Since the prescription of positive optimal long-run inflation hinges on the existence of misallocation, our findings provide additional arguments to the need to continue studying this issue. An interesting avenue for future research is to introduce firm dynamics in the setup (i.e. firm entry and exit). Such a model would constitute an ideal laboratory to understand the feedback between monetary policy and the distribution of firms' productivity.

References

Abowd, J., J. Haltiwanger, R. Jarmin, J. Lane, P. Lengermann, K. McCue, K. McKinney, and K. Sandusky (2005): "The Relation among Human Capital, Productivity, and Market Value: Building Up from Micro Evidence," in *Measuring Capital in the New Economy*, ed. by C. Corrado, J. Haltiwanger, and D. Sichel. University of Chicago Press.

Abowd, J., F. Kramarz, and D. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

Adam, K., and H. Weber (2019): "Optimal Trend Inflation," *American Economic Review*, 109, 702–737.

——— (2022): "Estimating the Optimal Inflation Target from Trends in Relative Prices," *American Economic Journal: Macroeconomics*, forthcoming.

Amiti, M., O. Itskhoki, and J. Konings (2019): "International Shocks, Variable Markups, and Domestic Prices," *The Review of Economic Studies*, 86, 2356–2402.

Ascari, G., and L. Rossi (2012): "Trend Inflation and Firms Price-Setting: Rotemberg versus Calvo," *Economic Journal*, 122, 1115–1141.

Atkin, D., A. Chaudhry, S. Chaudry, A. K. Khandelwal, and E. Verhoogen (2015): "Markup and Cost Dispersion across Firms: Direct Evidence from Producer Surveys in Pakistan," *American Economic Review: Papers and Proceedings*, 105, 537–544.

Baqaee, D., E. Farhi, and K. Sangani (2021): "The Supply-Side Effects of Monetary Policy," Working Papers 28345, NBER.

Bartelsman, E. J., and M. E. Doms (2000): "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, 38, 569–594.

Bilbiie, F. O., I. Fujiwara, and F. Ghironi (2014): "Optimal monetary policy with endogenous entry and product variety," *Journal of Monetary Economics*, 64, 1–20.

Bobeica, E., M. Ciccarelli, and I. Vansteenkiste (2019): "The link between labor cost and price inflation in the euro area," Working Papers 2235, European Central Bank.

Brown, C., and J. Medoff (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97, 1027–1059.

Carlsson, M., and O. N. Skans (2012): "Evaluating Microfoundations for Aggregate Price Rigidities: Evidence from Matched Firm-Level Data on Product Prices and Unit Labor Cost," *American Economic Review*, 102, 1571–1595.

Conflitti, C., and R. Zizza (2020): "What's behind firms' inflation forecasts?," *Empirical Economics*, https://doi.org/10.1007/s00181-020-01958-5.

Conti, A. M., and A. Nobili (2019): "Wages and prices in the euro area: exploring the nexus," Occasional Papers 518, Bank of Italy.

Dunne, T., L. Foster, J. Haltiwanger, and K. Troske (2004): "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment," *Journal of Labor Economics*, 22, 397–429.

Edmond, C., V. Midrigan, and D. Y. Xu (2018): "How costly are markups?," *NBER Working Paper*, (24800).

Forni, L., A. Gerali, A. Notarpietro, and M. Pisani (2015): "Euro area, oil and global shocks: An empirical model-based analysis," *Journal of Macroeconomics*, 46, 295–314.

Galí, J. (2008): *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton, NJ: Princeton University Press.

Gerali, A., S. Neri, L. Sessa, and F. M. Signoretti (2010): "Credit and Banking in a DSGE Model of the Euro Area," *Journal of Money, Credit and Banking*, 42, 107–141.

Gopinath, G., and O. Itskhoki (2010): "Frequency of Price Adjustment and Pass-Through," *Quarterly Journal of Economics*, 125, 675–727.

——— (2011): "In Search of Real Rigidities," *NBER Macroeconomics Annual 2010*, 25, 261–309.

Haltiwanger, J., G. Lane, and J. Spletzer (1999): "Productivity Differences Across Employers: The Roles of Employer Size, Age, and Human Capital," *American Economic Review*, 89, 94–98.

Issing, O. (ed.) (2003): *Background Studies for the ECB's Evaluation of its Monetary Policy Strategy*. European Central Bank.

Kimball, M. S. (1995): "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, 27, 1241–1277.

Klenow, P. J., and J. L. Willis (2016): "Real Rigidities and Nominal Price Changes," *Economica*, 83, 443–472.

Kurozumi, T., and W. V. Zandweghe (2020): "Output-Inflation Trade-offs and the Optimal Inflation Rate," Working Papers 20-20, Federal Reserve Bank of Cleveland.

Linde, J., and M. Trabandt (2019): "Resolving the Missing Deflation Puzzle," Discussion paper.

Meier, M., and T. Reinelt (2020): "Monetary Policy, Markup Dispersion, and Aggregate TFP," mimeo.

Ottaviano, G. I., D. Taglioni, and F. di Mauro (2007): "Deeper, wider and more competitive? Monetary integration, Eastern enlargement and competitiveness in the European Union," Working Papers 847, European Central Bank.

Peneva, E. V., and J. B. Rudd (2017): "The Passthrough of Labor Costs to Price Inflation," *Journal of Money, Credit and Banking*, 49, 1777–1802.

Restuccia, D., and R. Rogerson (2017): "The Causes and Costs of Misallocation," *Journal of Economic Perspectives*, 31, 151–174.

Riggi, M., and S. Santoro (2015): "On the Slope and the Persistence of the Italian Phillips Curve," *International Journal of Central Banking*, 11, 157–197.

Rotemberg, J. J. (1982): "Monopolistic Price Adjustment and Aggregate Output," *Review of Economic Studies*, 49, 517–531.

Schmitt-Grohé, S., and M. Uribe (2010): "The optimal rate of inflation," in *Handbook of Monetary Economics*, ed. by B. Friedman, and M. Woodford. Elsevier.

Sestito, P., and E. Viviano (2018): "Firing costs and firm hiring: evidence from an Italian reform," *Economic Policy*, 33, 101–130.

Figures and Tables

Figure 1: Optimal trend inflation as a function of the Pareto shape parameter θ and of the strength of strategic complementarities $\frac{\epsilon}{\sigma}$

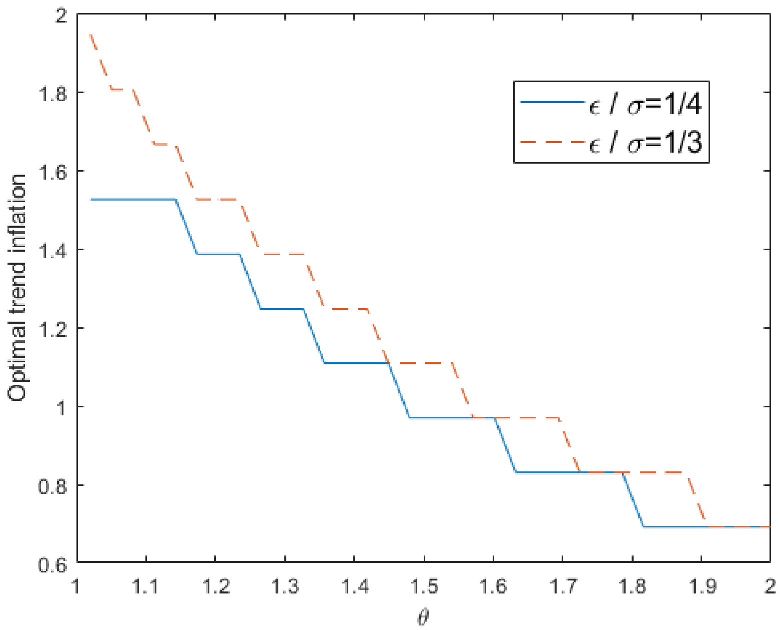
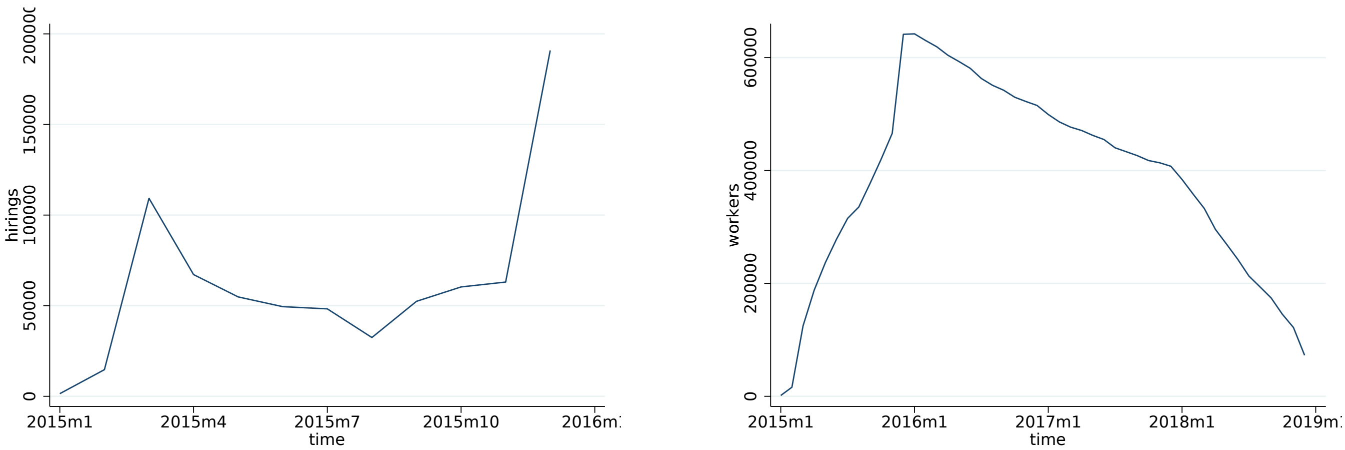


Figure 2: Evolution of subsidized hires (open-ended contracts) and number of workers

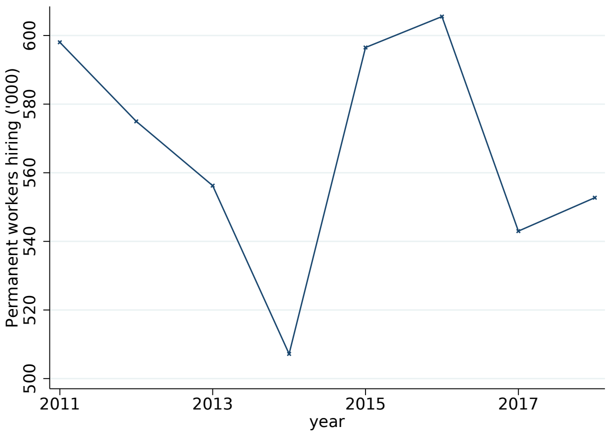
(a) *Subsidized hires* (b) *Subsidized workers*



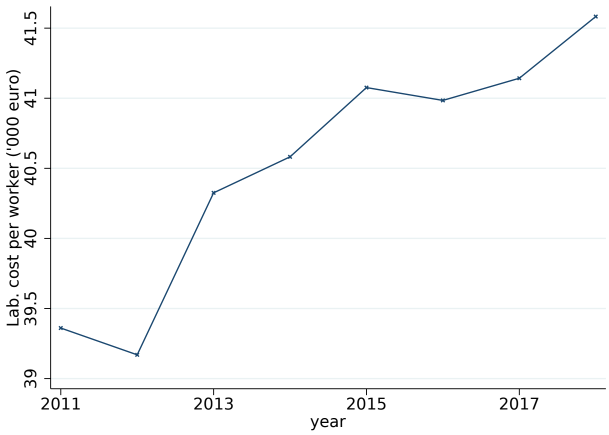
Note: INVIND-INPS, panel 2015-2018. Aggregate hirings in 2015 (panel a) and workers benefiting from the subsidy in 2015-2018, net of job separations.

Figure 3: Evolution of hirings with an open-ended contract, and labour costs per worker, 2011-2018

(a) *open-ended contract - hirings*



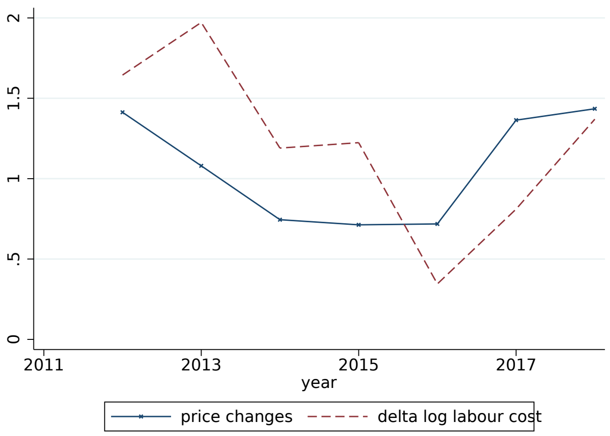
(b) *Labour costs per worker*



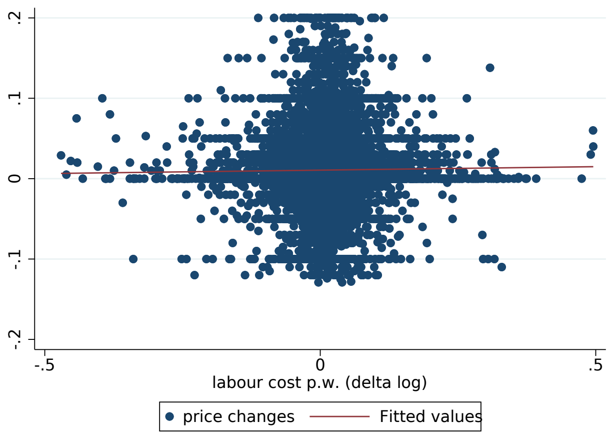
Note: INVIND-INPS-CERVED, 2011-2018. Hirings with an open-ended contract and labour costs per worker (wages and SSC, divided by the total number of employees, i.e. open-ended and fixed-term).

Figure 4: Producer price changes and labour costs per worker over time

(a) *Prices and labour costs*

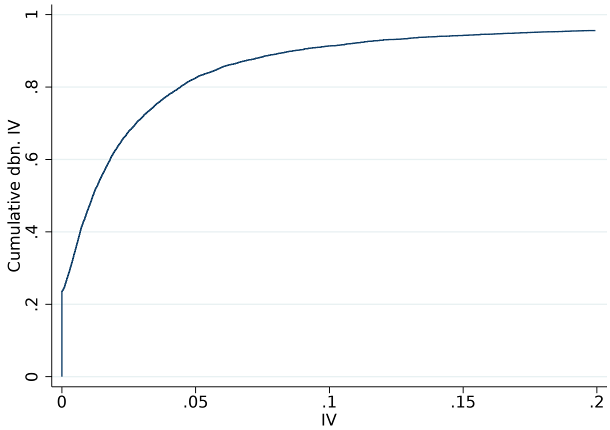


(b) *Correlation*



Note: INVIND-INPS-CERVED, 2011-2018. Firm-level price changes (yearly average over total products sold by the firm; %) and average firm-level labour costs (wage and social security contributions divided by the number of employees; delta logs).

Figure 5: Cumulative distribution of the number of fixed-term workers in the firm from January to October 2014, IV.



Note: INVIND-INPS-CERVED. Number of people employed with a fixed-term job contract from January to October 2014, divided by firm size (i.e. average number of employees in 2014, weighted by the duration of their employment relationship in days).

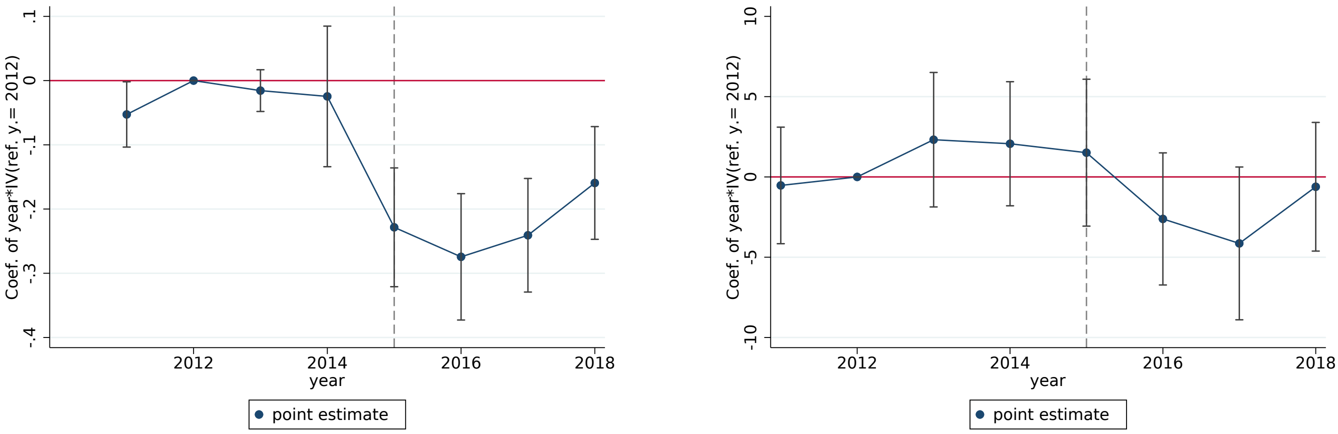
Table 1: Differences in the characteristics of firms by share of fixed-term employment in total firm size (IV). Year 2014

	Unweighted sample		Sample after matching	
	Share fixed-term workers		Share fixed-term workers	
	above mean	below mean	above mean	below mean
	mean/sd	mean/sd	mean/sd	mean/sd
	(1)	(2)	(3)	(4)
Firm size	229.158	86.985	229.158	254.367
	(586.376)	(319.182)	(586.376)	(806.457)
Labour share	0.720	0.716	0.720	0.713
	(0.393)	(0.596)	(0.393)	(0.572)
SSC (per worker)	8.759	8.788	8.759	8.797
	(0.291)	(0.352)	(0.291)	(0.337)
Average daily wage (logs)	4.567	4.556	4.567	4.582
	(0.298)	(0.314)	(0.298)	(0.307)
Change in prices (%)	0.736	0.331	0.736	0.414
	(3.445)	(4.022)	(3.445)	(3.246)

Note: INVIND-INPS-CERVED. Year 2014. Firm size: average number of employees, weighted by the duration of their employment relationship in days; Labour share: balance sheet value added divided by size; SSC: sum of firm-level social security contributions paid by firm divided by firm size; Average daily wage (in logs): sum of wages paid to workers per day divided by size; Change in prices: self reported yearly average change in prices of goods sold by the firm.

Figure 6: Labour costs per worker and producer price changes: estimates for the full sample

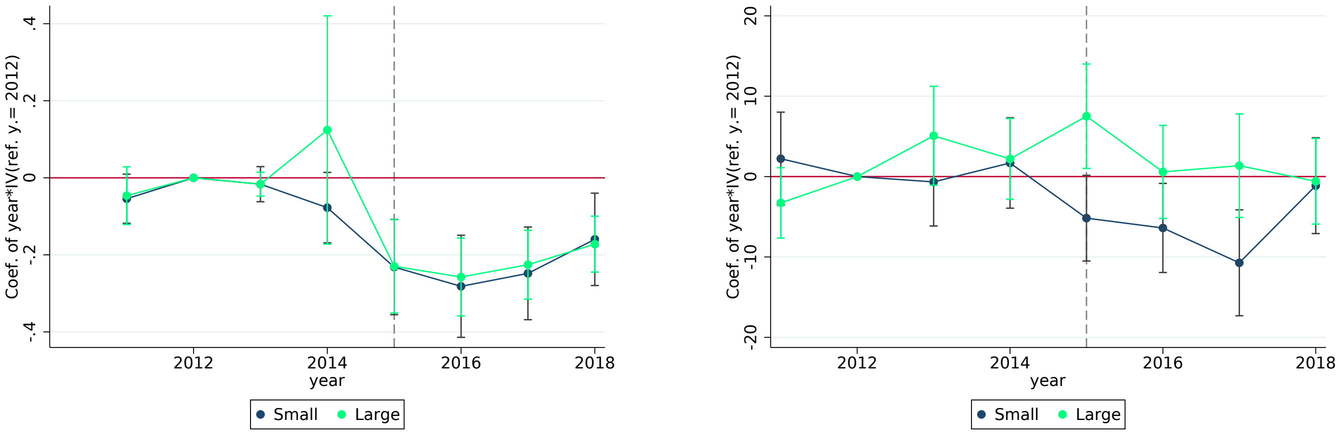
(a) *Estimates - SSC* (b) *Estimates - price changes*



Note: INVIND-INPS-CERVED, 2011-2018. Firm-level social security contributions per worker and firm-level price yearly changes. The vertical bars represent 95% confidence intervals.

Figure 7: Labour costs per worker and producer price changes. Estimates by firm size: large (above the median size) and small (below the median size)

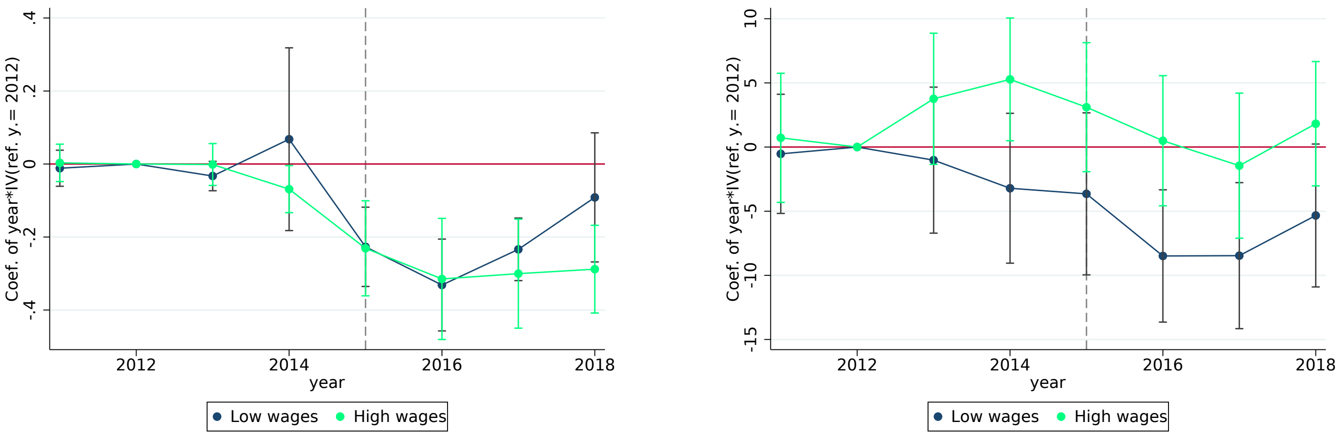
(a) *Estimates - SSC*   (b) *Estimates - price changes*



Note: INVIND-INPS-CERVED, 2011-2018. Firm-level social security contributions per worker and firm-level price yearly changes. The vertical bars represent 95% confidence intervals.

Figure 8: Labour costs per worker and producer price changes. Estimates by firm wage: high paying firms (above the mean daily wage) and small (below the mean)

(a) *Estimates - SSC*   (b) *Estimates - price changes*



Note: INVIND-INPS-CERVED, 2011-2018. Firm-level social security contributions per worker and firm-level price yearly changes. Wages are defined net of observable firms' characteristics (size, sectors, labour share, region). High paying firms are those paying average daily wages above the net average. Low paying firms are defined simmetrically. The vertical bars represent 95% confidence intervals.

Appendix A1

In this Appendix we prove the main result of the paper, Proposition 1. As a first step, we state and prove a series of technical Lemma.

Let's rewrite the optimality pricing condition of a firm with productivity Z in steady state as follows:

$$\psi'(y) = \mu(y, \Pi) \frac{A}{Z} \tag{31}$$

where $A \equiv \frac{W}{PD}$ is a variable summarizing aggregate conditions (different from trend inflation) relevant for firm's pricing; the above equation implicitly gives $\tilde{y}(Z; A, \Pi)$, the relative quantity of a firm of productivity Z for given A and Π. We can prove the following:

Lemma 1. *Let $\tilde{y}(Z; A, \Pi)$ be implicitly given by equation* (31)*; then, for any given A:*

(i) *the derivative of $\tilde{y}$ with respect to inflation, when evaluated at zero inflation, is positive:*

$$\left. \frac{\partial}{\partial \Pi} \tilde{y}(Z; A, \Pi) \right|_{\Pi=1} > 0$$

(ii) *the derivative of $\tilde{y}$ with respect to inflation, when evaluated at zero inflation, is increasing in Z:*

$$\left. \frac{\partial}{\partial \Pi} \tilde{y}\left(Z''; A, \Pi\right) \right|_{\Pi=1} > \left. \frac{\partial}{\partial \Pi} \tilde{y}\left(Z'; A, \Pi\right) \right|_{\Pi=1}$$

for $Z'' > Z'$.

Proof. **(i)** We can use the Implicit Function Theorem to compute the desired derivative; let:

$$J(y, Z; A, \Pi) \equiv \frac{\psi'(y)}{\mu(y, \Pi)} - \frac{A}{Z}$$

then:

$$\left. \frac{\partial}{\partial \Pi} \tilde{y}(Z; A, \Pi) \right|_{\Pi=1} = - \left. \frac{\partial J / \partial \Pi}{\partial J / \partial y} \right|_{\Pi=1}$$

where[19]:

$$- \left. \frac{\partial J / \partial \Pi}{\partial J / \partial y} \right|_{\Pi=1} = \frac{(1-\beta)\phi y \sigma}{\epsilon + \sigma - y^{\epsilon/\sigma}} > 0 \tag{32}$$

The last inequality holds since in equilibrium $\sigma > y^{\epsilon/\sigma}$.

[19] The Mathematica code to compute the derivative below is available upon request.

(ii) Simple monotone comparative statics argument can be used to show that, with zero inflation, $Z'' > Z'$ implies $\tilde{y}\left(Z''; A, \Pi\right) > \tilde{y}\left(Z'; A, \Pi\right)$; using this relation, simple inspection of equation (32) confirms the statement.

□

The above Lemma shows that (at low levels of inflation) an increase in Π triggers an upward shift of the function $\tilde{y}(Z; A, \Pi)$, the more so for high levels of Z; since ψ is increasing, the Kimball aggregator (6) does not hold anymore, and $\int_\Omega \psi(\tilde{y}(Z; A, \Pi))\, dH(Z) > 1$. To restore equilibrium, the aggregate conditions variable A has to adjust. In particular, the following holds.

Lemma 2. *Let $\tilde{y}(Z; A, \Pi)$ be given by equation (31), $\tilde{A}$ be implicitly given by:*

$$\int_\Omega \psi\left(\tilde{y}\left(Z; \tilde{A}, \Pi\right)\right) dH(Z) = 1$$

and $y(Z; \Pi)$ be pinned down by $\tilde{y}(Z; A, \Pi)$ when evaluated in the equilibrium $\tilde{A}$:

$$y(Z; \Pi) \equiv \tilde{y}\left(Z; \tilde{A}, \Pi\right) \tag{33}$$

Then:

(i) *the function $\tilde{y}$ is decreasing in A, when evaluated at zero inflation;*

(ii) *the derivative of $\tilde{A}$ with respect to inflation, when evaluated at zero inflation, is positive:*

$$\left.\frac{\partial}{\partial \Pi}\tilde{A}\right|_{\Pi=1} > 0$$

(iii) *the derivative of y with respect to inflation, when evaluated at zero inflation, is positive for Z larger than a threshold $\hat{Z}$, and negative for $Z < \hat{Z}$.*

Proof. **(i)** We use again the Implicit Function Theorem. We have:

$$\left.\frac{\partial}{\partial A}\tilde{y}(Z; A, \Pi)\right|_{\Pi=1} = -\left.\frac{\partial J/\partial A}{\partial J/\partial y}\right|_{\Pi=1}$$

and:

$$-\left.\frac{\partial J/\partial A}{\partial J/\partial y}\right|_{\Pi=1}=-\frac{\sigma^{3}\exp\frac{y^{\epsilon/\sigma}}{\epsilon}y^{1-\frac{\epsilon}{\sigma}}}{Z(\sigma-1)\exp\frac{1}{\epsilon}\left(\epsilon+\sigma-y^{\epsilon/\sigma}\right)}<0 \tag{34}$$

where the last inequality holds since in equilibrium $\sigma > y^{\epsilon/\sigma}$.

(ii) Let:

$$K(A,\Pi)\equiv\int_{\Omega}\psi\left(\tilde{y}(Z;A,\Pi)\right)dH(Z)-1$$

Again by the Implicit Function Theorem we have:

$$\left.\frac{\partial\tilde{A}}{\partial\Pi}\right|_{\Pi=1}=-\left.\frac{\partial K/\partial\Pi}{\partial K/\partial A}\right|_{\Pi=1}$$

where:

$$\begin{aligned}\frac{\partial}{\partial\Pi}K(A,\Pi) &= \int_{\Omega}\psi'\left(\tilde{y}(Z;A,\Pi)\right)\frac{\partial}{\partial\Pi}\tilde{y}(Z;A,\Pi)\,dH(Z)\\ \frac{\partial}{\partial A}K(A,\Pi) &= \int_{\Omega}\psi'\left(\tilde{y}(Z;A,\Pi)\right)\frac{\partial}{\partial A}\tilde{y}(Z;A,\Pi)\,dH(Z)\end{aligned}$$

As ψ is increasing, and invoking Lemma 1 (i) and Lemma 2 (i), we conclude that $\frac{\partial}{\partial\Pi}K(A,\Pi)$ is positive when evaluated at zero inflation, while $\frac{\partial}{\partial A}K(A,\Pi)$ is negative, hence:

$$\left.\frac{\partial\tilde{A}}{\partial\Pi}\right|_{\Pi=1}>0$$

(ii) We can write:

$$\left.\frac{\partial}{\partial\Pi}y(Z;\Pi)\right|_{\Pi=1}=\left.\left[\frac{\partial}{\partial\Pi}\tilde{y}(Z;A,\Pi)+\frac{\partial}{\partial A}\tilde{y}(Z;A,\Pi)\frac{\partial\tilde{A}}{\partial\Pi}\right]\right|_{\Pi=1}$$

Substituting equation (32) into (34), we get:

$$\left.\frac{\partial}{\partial\Pi}y(Z;\Pi)\right|_{\Pi=1}=\left.\left[\frac{\partial}{\partial\Pi}\tilde{y}(Z;A,\Pi)\left(1-\frac{\partial\tilde{A}}{\partial\Pi}\frac{\exp\frac{y^{\epsilon/\sigma}}{\epsilon}}{Z}\frac{\sigma^{2}y^{-\frac{\epsilon}{\sigma}}}{(\sigma-1)\exp\frac{1}{\epsilon}(1-\beta)\phi}\right)\right]\right|_{\Pi=1}$$

We can now use optimality condition (31) to substitute out Z: with zero inflation (31) is equivalent to:

$$\frac{\exp\frac{y^{\epsilon/\sigma}}{\epsilon}}{Z}=\frac{\frac{\sigma-1}{\sigma}\frac{\sigma-y^{\epsilon/\sigma}}{\sigma}\exp\frac{1}{\epsilon}}{A}$$

which can be plugged in the above equation:

$$\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\bigg|_{\Pi=1} = \left[\frac{\partial}{\partial \Pi}\tilde{y}\left(Z;A,\Pi\right)\left(1-\frac{\partial \tilde{A}}{\partial \Pi}\frac{\frac{\sigma-1}{\sigma^2}\exp\frac{1}{\epsilon}}{A}\frac{\sigma^2\left(\sigma-y^{\epsilon/\sigma}\right)y^{-\frac{\epsilon}{\sigma}}}{(\sigma-1)\exp\frac{1}{\epsilon}(1-\beta)\phi}\right)\right]\bigg|_{\Pi=1}$$

The derivative $\frac{\partial}{\partial \Pi}\tilde{y}\left(Z;A,\Pi\right)$ is positive by Lemma 1 (i), and the term:

$$1-\frac{\partial \tilde{A}}{\partial \Pi}\frac{\frac{\sigma-1}{\sigma^2}\exp\frac{1}{\epsilon}}{A}\frac{\sigma^2\left(\sigma-y^{\epsilon/\sigma}\right)y^{-\frac{\epsilon}{\sigma}}}{(\sigma-1)\exp\frac{1}{\epsilon}(1-\beta)\phi}$$

moves monotonically (and continuously) from $-\infty$ to 1, as $y^{\epsilon/\sigma}$ ranges from zero to σ; hence, there is a threshold $\hat{y}$ such that:

$$\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\bigg|_{\Pi=1} > (<) 0 \qquad \Leftrightarrow \qquad y > (<) \hat{y}$$

Finally, defining implicitly $\hat{Z}$ as the solution of:

$$y\left(\hat{Z};\Pi\right) = \hat{y}$$

proves the statement.

□

We proved in Lemma 2 (iii) that increasing trend inflation from zero triggers a reallocation of production from the less productive firms to the more productive; in light of the discussion in Section 4, this means that the misallocation distortion is curbed. In the following Lemma, we show that it results in an increase in $\mathcal{Z}$ and in $\frac{\mathcal{Z}}{\mathcal{M}}$.

Lemma 3. *Let $y\left(Z;\Pi\right)$ be given by equation* (33). *Then:*

(i) *the derivative of $\mathcal{Z}$ with respect to inflation, when evaluated at zero inflation, is positive:*

$$\frac{\partial}{\partial \Pi}\mathcal{Z}\bigg|_{\Pi=1} > 0$$

(ii) *the derivative of $\frac{\mathcal{Z}}{\mathcal{M}}$ with respect to inflation, when evaluated at zero inflation, is positive:*

$$\frac{\partial}{\partial \Pi}\frac{\mathcal{Z}}{\mathcal{M}}\bigg|_{\Pi=1} > 0$$

Proof. **(i)** We show that $\mathcal{Z}^{-1}$ is decreasing in Π. Its derivative is given by:

$$\frac{\partial}{\partial \Pi}\mathcal{Z}^{-1}\Big|_{\Pi=1} = \int_{\underline{Z}}^{\hat{Z}} \frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) + \int_{\hat{Z}}^{\overline{Z}} \frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right)$$

where, by Lemma 2 (iii) the first term in the RHS is negative, while the second one is positive. To sign the derivative, we can write the following:

$$\begin{aligned}
0 &= \int_{\underline{Z}}^{\overline{Z}} \psi'\left(y\left(Z;\Pi\right)\right)\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) \\
&= \int_{\underline{Z}}^{\overline{Z}} Z\psi'\left(y\left(Z;\Pi\right)\right)\frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) \\
&= \int_{\underline{Z}}^{\hat{Z}} Z\psi'\left(y\left(Z;\Pi\right)\right)\frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) + \int_{\hat{Z}}^{\overline{Z}} Z\psi'\left(y\left(Z;\Pi\right)\right)\frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) \\
&> \hat{Z}\psi'\left(y\left(\hat{Z};\Pi\right)\right)\left[\int_{\underline{Z}}^{\hat{Z}} \frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) + \int_{\hat{Z}}^{\overline{Z}} \frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right)\right]
\end{aligned}$$

where the first equality is obtaining differentiating the Kimball aggregator (6), and the last inequality is a consequence of the fact that the term $Z\psi'\left(y\left(Z;\Pi\right)\right)$ is increasing in Z (it can be see from equation (31), considering that the markup is increasing in y at zero inflation). As Z takes only positive values, and ψ is an increasing function, we get the desired result:

$$\int_{\underline{Z}}^{\hat{Z}} \frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) + \int_{\hat{Z}}^{\overline{Z}} \frac{1}{Z}\frac{\partial}{\partial \Pi} y\left(Z;\Pi\right)\Big|_{\Pi=1} dH\left(Z\right) < 0$$

(ii) We show that $\frac{\mathcal{M}}{\mathcal{Z}}$ is decreasing in Π. From equations (25)-(26), we have:

$$\frac{\mathcal{M}\left(\Pi\right)}{\mathcal{Z}\left(\Pi\right)} = \int_{\Omega} \mu\left(Z;\Pi\right)\frac{y\left(Z;\Pi\right)}{Z} dH\left(Z\right)$$

From equation (31), the total differential of μ with respect to Π can be expressed as:

$$\begin{aligned}
\frac{d}{d\Pi}\mu\left(y\left(Z;\Pi\right),\Pi\right)\Big|_{\Pi=1} &= \frac{d}{d\Pi}\left[Z\frac{\psi'\left(y\left(Z;\Pi\right)\right)}{\tilde{A}\left(\Pi\right)}\right]\Big|_{\Pi=1} \\
&= \left[Z\frac{\tilde{A}\left(\Pi\right)\psi''\left(y\left(Z;\Pi\right)\right)\frac{\partial}{\partial\Pi}y\left(Z;\Pi\right)-\psi'\left(y\left(Z;\Pi\right)\right)\frac{\partial}{\partial\Pi}\tilde{A}\left(\Pi\right)}{\left(\tilde{A}\left(\Pi\right)\right)^{2}}\right]\Big|_{\Pi=1} \\
&= \Bigg[\mu\left(y\left(Z;\Pi\right),\Pi\right)\Bigg(\frac{\psi''\left(y\left(Z;\Pi\right)\right)y\left(Z;\Pi\right)}{\psi'\left(y\left(Z;\Pi\right)\right)}\frac{\partial}{\partial\Pi}y\left(Z;\Pi\right)\frac{1}{y\left(Z;\Pi\right)} \\
&\quad -\frac{\partial}{\partial\Pi}\tilde{A}\left(\Pi\right)\frac{1}{\tilde{A}\left(\Pi\right)}\Bigg)\Bigg]\Big|_{\Pi=1}
\end{aligned}$$

where the second term in the round bracket is an *aggregate* negative impact of inflation on markups, that would be present also in absence of misallocation, with homogeneous firms (see Bilbiie, Fujiwara, and Ghironi (2014)); instead, the first term is *firm-specific*, depends on the demand elasticity, is negative for firms with $Z > \hat{Z}$ (by concavity of ψ and Lemma 2 (iii)) and positive otherwise. This shows how increasing Π curbs misallocation: the markup dispersion is reduced, since for the most productive firms (those with higher markups to begin with) there is an additional force that drives down markups, on top of the aggregate effect, while for the less productive firms the opposite holds. Overall, markups are reduced for all firms[20], and as a consequence:

$$\begin{aligned}
\frac{\partial}{\partial\Pi}\frac{\mathcal{M}\left(\Pi\right)}{\mathcal{Z}\left(\Pi\right)}\Big|_{\Pi=1} &= \int_{\underline{Z}}^{\overline{Z}}\mu\left(y\left(Z;\Pi\right),\Pi\right)\frac{1}{Z}\frac{\partial}{\partial\Pi}y\left(Z;\Pi\right)\Big|_{\Pi=1}dH\left(Z\right)+ \\
&\quad \int_{\underline{Z}}^{\overline{Z}}\frac{d}{d\Pi}\mu\left(y\left(Z;\Pi\right),\Pi\right)\Big|_{\Pi=1}\frac{y\left(Z;\Pi\right)}{Z}dH\left(Z\right) \\
&< \int_{\underline{Z}}^{\overline{Z}}\mu\left(y\left(Z;\Pi\right),\Pi\right)\frac{1}{Z}\frac{\partial}{\partial\Pi}y\left(Z;\Pi\right)\Big|_{\Pi=1}dH\left(Z\right) \\
&= \frac{1}{\tilde{A}\left(\Pi\right)}\int_{\underline{Z}}^{\overline{Z}}Z\psi'\left(y\left(Z;\Pi\right)\right)\frac{1}{Z}\frac{\partial}{\partial\Pi}y\left(Z;\Pi\right)\Big|_{\Pi=1}dH\left(Z\right) \\
&= 0
\end{aligned}$$

where the inequality stems from the previous argument on the total differential of μ with respect to Π, and the final equality from differentiating the Kimball aggregator (6).

□

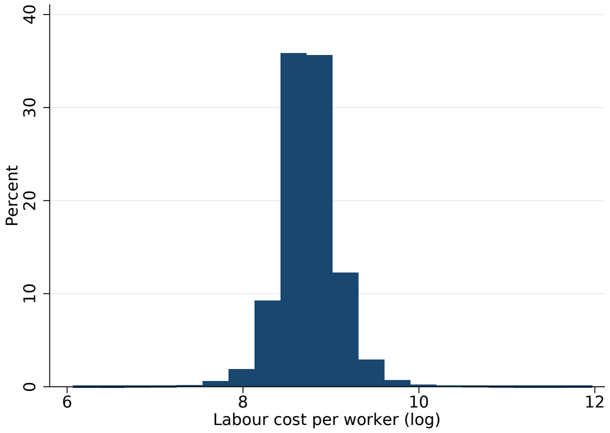
[20]Observe that both the arguments of the function $\mu\left(y\left(Z;\Pi\right),\Pi\right)$ are decreasing in Π, when $Z < \hat{Z}$

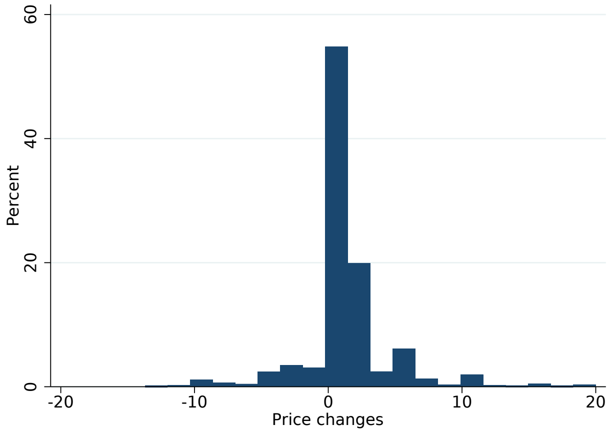
We are now ready to prove Proposition 1.

Proof of Proposition 1. From Lemma 3 we know that a marginal increase of Π from 1 has a first-order effect on $\mathcal{Z}$ and $\frac{\mathcal{Z}}{\mathcal{M}}$, increasing both. It is easy to see from equations (27)-(29) that this raises welfare; on the other side, the cost of increasing inflation is second-order, hence for small values of Π above one the first effect dominates. $\square$

Appendix A2

A2.1-Distribution of social security contributions per worker and price changes. 2011-2018





Note: INVIND-INPS-CERVED. Firm-level labour costs (wage and social security contributions) divided by the number of employees (in logs) and firm-level price yearly changes (average over total products sold by the firm, %).

Acknowledgements

We thank Roberto M. Billi, Federico Cingano, Francesco Lippi, Antonio Mele, Stefano Neri, Alessandro Notarpietro, Andrea Tiseno, Gianluca Violante and Roberta Zizza for useful conversations. All remaining errors are our own. The analysis and conclusions expressed herein are those of the authors and should not be interpreted as those of the ECB or the Bank of Italy.

Sergio Santoro
European Central Bank, Frankfurt am Main, Germany; email: Sergio.Santoro@ecb.europa.eu

Eliana Viviano
Bank of Italy, Rome, Italy; email: eliana.viviano@bancaditalia.it

© European Central Bank, 2022

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from www.ecb.europa.eu, from the Social Science Research Network electronic library or from RePEc: Research Papers in Economics. Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website.

PDF ISBN 978-92-899-5473-0 ISSN 1725-2806 doi:10.2866/268910 QB-AR-22-126-EN-N