Discussion Paper Series – CRC TR 224

Discussion Paper No. 406
Project C 02

Inflation Distorts Relative Prices: Theory and Evidence

Klaus Adam [1]
Andrey Alexandrov [2]
Henning Weber [3]

May 2024
(First version: April 2023)

[1] University of Mannheim, CEPR, Email: adam@uni-mannheim.de
[2] Princeton University, Email: alexandrov@princeton.edu
[3] Deutsche Bundesbank, Email: henning.weber@bundesbank.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Inflation Distorts Relative Prices: Theory and Evidence*

Klaus Adam (University of Mannheim & CEPR)
Andrey Alexandrov (Princeton University)
Henning Weber (Deutsche Bundesbank)

April 17, 2023

Abstract

Using a novel identification approach derived from sticky price theories with time or state-dependent adjustment frictions, we empirically identify the effect of inflation on relative price distortions. Our approach can be directly applied to micro price data, does not rely on estimating the gap between actual and flexible prices, and only assumes stationarity of unobserved shocks. Using U.K. CPI micro price data, we document that suboptimally high (or low) inflation is associated with distortions in relative prices that are highly statistically significant. At the aggregate level, fluctuations in inefficient price dispersion are sizable and covary positively with aggregate inflation. In contrast, overall price dispersion fails to covary with inflation because it is mainly driven by trends in the dispersion of flexible prices.

JEL Class. No.: E31, E58

*We thank Jordi Galí, Krzysztof Pytka, Frank Schorfheide and Carsten Trenkler for useful input and discussions. Klaus Adam gratefully acknowledges support from the Deutsche Forschungsgemeinschaft through CRC-TR 224 (project C02). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem.

1 Introduction

The monetary models employed in academia and central banks postulate that too high (or too low) rates of inflation give rise to distortions in relative prices. These distortions drive many of the trade-offs and policy prescriptions, e.g., the recommendation to implement low and stable inflation rates.[1] While being central to monetary theory, there exists no structural empirical evidence in support of inflation distorting relative prices.

The present paper seeks to fill this gap and derives a novel theory-consistent empirical approach that allows estimating the marginal effect of inflation on relative price distortions. The paper uses this approach to estimate the relationship between inflation and relative price distortions in the micro price data underlying the U.K. consumer price index. It documents that inflation is associated - at the product level - with economically significant amounts of price distortions, in line with what sticky price theories predict. At the aggregate level, price distortions covary positively with aggregate inflation over time.

Documenting the relationship between inflation and relative price distortions is challenging for a number of reasons and the present paper makes progress by overcoming a number of these challenges.

First, it is difficult to recover inflation-induced price distortions from price observations. Price distortions consist of the gap between the actual price charged by the firm and the so-called 'flexible price', i.e., the counterfactual price the firm would charge in the absence of price rigidities.[2] We formally show that the flexible price process - and therefore the distribution of price distortions - cannot be identified from micro price data, whenever the dynamics of the flexible price contain some stationary stochastic component.[3]

In light of this finding, it may not be surprising that previous papers documenting the relationship between price dispersion and inflation do *not* decompose price dispersion into the dispersion present under flexible prices and the *additional* dispersion arising from too high or too low inflation„ e.g., Nakamura, Steinsson, Sun and Villar (2018) and Wulfs-

[1] See, for instance, Woodford (2003), Galí (2015), Adam and Weber (2019) or Archarya, Challe and Dogra (2023).

[2] The flexible price may itself be distorted, e.g., due to market power. Distortions due to price stickiness come *on top* of the distortions that are already present under flexible prices.

[3] Price gaps can be estimated in the rare cases in which additional information about marginal costs and the desired mark-up is available. Eichenbaum, Jaimovich and Rebelo (2011) estimate price gaps using such information for supermarket goods, but do not analyze how inflation affects price distortions.

berg (2016). Instead, these papers highlight the difficulties associated with empirically recovering the gaps between the actual and the flexible price.[4]

An important contribution of the present paper is to show how one can empirically identify the *marginal effect* of inflation on price distortions without the need to identify the *level* of price distortions. We show that time and state-dependent pricing models make identical predictions (up to a second-order approximation) on how this can be achieved using micro price data alone: one first computes residual price variation around the life-cycle trend of a product, where a product is a physical object or a service sold in a particular location over time. In a second step, one relates this residual variation - in the cross-section of products - to a squared measure of the deviation of inflation from its product-specific optimal level. This structural approach is valid without imposing any assumptions on the evolution of the cross-sectional distribution of flexible prices over time.

A second challenge is related to the fact that - according to sticky price theory - a higher inflation rate may either increase or decrease price distortions. The direction of the effect depends on whether the current inflation rate lies above or below the optimal (distortion-minimizing) inflation rate.[5] Existing work tends to ignore this issue and often assumes that the optimal inflation rate is zero. Yet, the optimal inflation rate typically differs from zero and has been found to vary systematically in the cross-section of products (Adam and Weber (2022), Adam, Gautier, Santoro and Weber (2022)).

To address this issue, the present paper uses measures of *suboptimal* inflation, i.e., of the difference between the *actual* and the *optimal* level of inflation to obtain estimates of the marginal effect of suboptimal inflation on relative price distortions. We show that price distortions in the data depend on the squared value of *suboptimal* inflation, in line with the theoretical predictions. In particular, we robustly find that the effect of suboptimal inflation on product-level price distortions is statistically significant in 95% of the expenditure categories underlying the U.K. consumer price index. Interestingly, the notion that price distortions at the product level are driven by the square of the inflation rate receives no support in our data.

The third challenge this paper addresses is that it is generally difficult to establish a causal relationship between inflation and inefficient price dispersion by exploiting variation in aggregate inflation over time: out-

[4]See section IV.A in Nakamura et al. (2018).

[5]If the optimal level of inflation lies above (below) the actual level, then sticky price models predict that higher rates of inflation decrease (increase) price distortions.

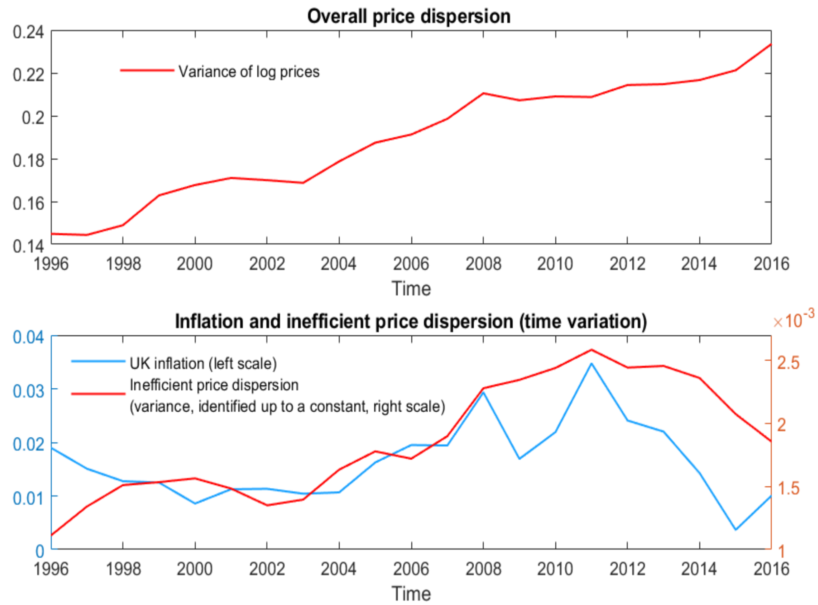

Figure 1: Inflation and price dispersion (United Kingdom, 1996-2016)

side hyperinflationary episodes, aggregate inflation tends to vary slowly over time, so that its movements are often hard to distinguish from a slow-moving time trend. Observed time trends in price dispersion might then reflect either the time trends in inflation or other trends, e.g., a secular increase in the variety of products over time.

Our empirical approach overcomes this issue by exploiting cross-sectional variation in the product-specific optimal inflation rate. This variation is driven by product-specific fundamentals, e.g., the rate of productivity progress at the product level. According to the theory, cross-sectional heterogeneity in these product-specific fundamentals is unrelated to inflation and thus induces quasi-exogenous variation in the gap between actual and optimal inflation that can be exploited to estimate causal effects.

Using these insights, we decompose the observed cross-sectional dispersion of prices into two components: (i) a component reflecting identifiable parts of the flexible price dispersion, and (ii) a residual component. While the residual component fails to identify the level of price distortions, one can use it to identify the *marginal* effect of inflation on price distortions.

Figure 1 illustrates this decomposition. The top panel depicts the evolution of overall cross-sectional dispersion of log prices over time.[6] Overall price dispersion in the U.K. is rising strongly and stands at a level that is more than 50% higher in 2016 than in 1996. Inflation, however, fails to display any trend over this time period, see the bottom panel in figure 1, which suggests that the observed upward trend in price dispersion is unrelated to inflation.[7] Using our product-level approach, we can in fact show that almost all of the increase in overall price dispersion is due to increased dispersion in the identifiable parts of the flexible price distribution, i.e., component (i) mentioned above. Since the flexible price distribution reflects differences in productivities, flexible-price mark-ups, or unobserved qualities across products, our approach suggests that these factors are predominantly driving overall price dispersion in the data.

The bottom panel in figure 1 also depicts the residual component (ii) of overall price dispersion: this component covaries strongly with aggregate inflation over time, with a correlation equal to +0.58 that is statistically significant at the 1% level. According to the theory, time variation in this measure captures time variation in relative price distortions due to changes in inflation.[8] Time variation in price distortions is quantitatively large and implies that changes in inflation alone give rise to an inefficient cross-sectional standard deviation of log prices reaching at least 3.8% over the sample period.[9]

Our finding that an increase in aggregate inflation leads to an increase in aggregate price distortions aligns well with key assumptions made in monetary models and should thus increase confidence in the economic relevance of key policy recommendations derived from these models, e.g., the desirability of targeting low and stable inflation rates. It also aligns well with recent findings in Ascari, Bonmolo and Haque (2022), who show that high inflation rates are associated with a loss in the economy's

[6] The dispersion measure is constructed by computing the variance of log prices at the level of more than 1000 expenditure items and then aggregating across items at each point in time using household expenditure weights. Nakamura et al. (2018) compute aggregate price dispersion using the interquartile range of log prices and aggregate across expenditure items using the expenditure weighted median. This leads to very similar conclusions. We use the variance measure because the underlying theory delivers decomposition and aggregation results for variance-based dispersion measures only.

[7] Even detrended measures of overall price dispersion fail to correlate in a statistically significant way with inflation.

[8] Due to the identification problem, the *level* of this measure does not identify the *level* of relative price distortions.

[9] While the dispersion measure fails to identify the level of price distortions, it can still be used to compute upper an lower bounds for the contribution of inflation to price distortions.

output potential. Relative price distortions are one source of potential output losses associated with high inflation rates, as emphasized in the literature studying price-induced misallocations inferred from product mark-ups (Baqaee, Farhi and Sangani (2022), Meier and Reinelt (2022)).

The paper is related to work by Alvarez, Beraja, Gonzalez-Rozada and Neumeyer (2019) who estimate a nonlinear relationship between the cross-sectional dispersion of prices and inflation using data from Argentina. They find that cross-sectional price dispersion does not respond to inflation for inflation rates below 10%, but rises strongly for higher rates and eventually levels off. Relatedly, Sheremirov (2020) uses supermarket scanner data for the U.S. and documents how the cross-sectional dispersion of prices for products with identical barcodes correlates with inflation over time. He finds that covariation is negative when including all prices, but turns slightly positive when excluding sales prices.[10] Instead of estimating a reduced-form relationship between the *cross-sectional* dispersion of prices and inflation, our structural approach calls for estimating *across-time* dispersion of prices at the level of individual products and relating this dispersion to a product-specific measure of suboptimal inflation.

The remainder of the paper is structured as follows. Section 2 illustrates, using the simplest possible case, the empirical approach developed in this paper for identifying the relationship between suboptimal inflation and inefficient price dispersion. Section 3 introduces the full theory and shows how sticky price models with time or state-dependent pricing frictions imply a regression approach that allows estimating the causal effect of suboptimal inflation on price distortions. Section E discusses econometric issues associated with implementing the approach and section 4 introduces the U.K. micro price data to which it is applied. The main empirical results, including a number of robustness exercises, are presented in section 5. Section 6 discusses variation of aggregate price distortions over time and their covariation with aggregate inflation. A conclusion briefly summarizes. Most technical derivations can be found in the appendix.

2 The Approach in a Nutshell

This section illustrates how one can empirically identify the marginal contribution of suboptimal inflation to inefficient price dispersion from micro price data.

Identification is achieved by considering a set of products for which

[10] Sara-Zaror (2022) extends the empirical approach of Sheremirov (2020) and documents that cross-sectional price dispersion strongly rises with the absolute deviation of inflation from zero, with the relationship becoming flatter for larger inflation rates.

(i) price stickiness and (ii) the shock process driving the idiosyncratic component of the flexible price is homogeneous across products. One can then exploit variation in the optimal inflation rate across products to identify the marginal effect of inflation on inefficient price dispersion. This holds true even if the actual inflation rate is constant over time.

To provide a simple example, we assume that idiosyncratic shocks are simply absent, so that the flexible relative price evolves deterministically, and that prices get adjusted in regular intervals every $N > 1$ periods (Taylor (1979)).[11] Consider product j, which is a physical object or service sold in a specific location.[12] The flexible optimal relative price $p^*_{jt} = P^*_{jt}/P_t$ of product j is the price the firm would like to charge in the absence of any price setting frictions and evolves deterministically according to

$$\ln p^*_{jt} = \ln p^*_j - t \cdot \ln \Pi^*_j, \tag{1}$$

where p^*_j is a product-specific intercept and Π^*_j a product-specific time trend, capturing differences in marginal costs (or other factors) across products. Finally, suppose gross inflation is constant and equal to Π.

In this setting, the optimal inflation rate for product j is given by $\ln \Pi = \ln \Pi^*_j$ because the relative price then gets eroded at the desired rate $\ln \Pi^*_j$: the nominal price for product j can remain constant, so that price setting frictions do no matter for tracking the desired relative price. When $\ln \Pi > \ln \Pi^*_j$ ($\ln \Pi < \ln \Pi^*_j$), the relative price gets eroded too quickly (slowly). As a result, adjustments of the nominal price have to be made to correct for the 'wrong' trend induced by inflation during non-adjustment periods. Due to price stickiness, these adjustments occur only occasionally, so that suboptimal inflation leads to deviations of the relative price from the flexible relative price.

Figure 2 illustrates the situation. It depicts the flexible relative price $\ln p^*_{jt}$ for three products ($j = 1, 2, 3$), for which the flexible relative price falls at rate $\Pi^*_1 < \Pi^*_2 < \Pi^*_3$. Assuming that actual inflation Π is equal to Π^*_1, the flexible relative price of product 1 coincides with the sticky relative price $\ln p_{jt}$, so that there are no relative price distortions. For product $j = 2$, inflation is too low, which means that the relative price falls insufficiently during non-adjustment periods. To compensate for this effect, it becomes optimal to choose a relative price that is lower than the flexible price in adjustment periods, to reduce the gap between

[11] These assumptions are special because they allow identifying the flexible price from micro price data, which fails to be true under the more general assumptions considered later on, but useful for illustrating the approach.

[12] Objects or services that are sold in different locations are treated as different products. The same holds true when an existing product gets substituted by a new product.

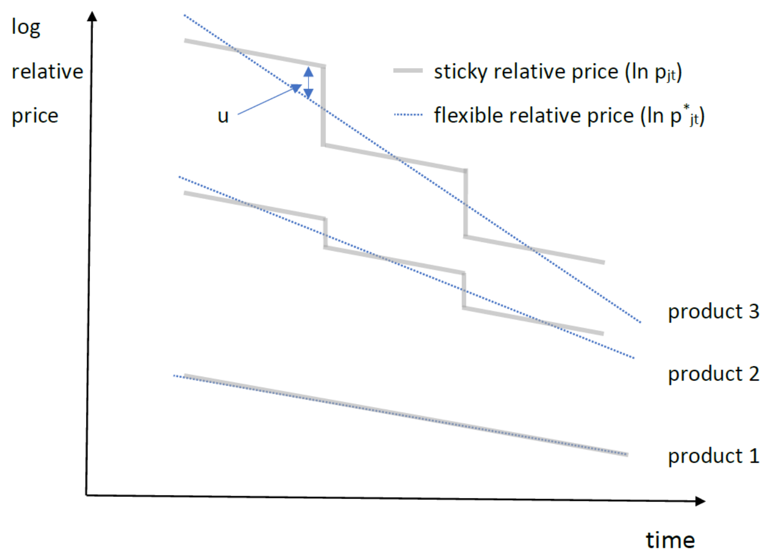

Figure 2: Relative price trends and relative price distortions (u)

the sticky and the flexible relative price over the lifetime of the sticky price. Suboptimally low inflation thus leads to a deviation of the sticky relative price from the flexible relative price. This deviation is even stronger for product $j = 3$, which has a higher optimal inflation rate and - in adjustment periods - a relative price that is even further below the flexible relative price. A larger gap between inflation and the optimal inflation rate thus gives rise to larger deviations of the sticky relative price from the flexible relative price.

Since symmetric arguments apply when inflation is higher than optimal inflation, it is easy to verify that the variance of the gap u between the sticky relative price around its time trend, i.e., the variance of price distortions, is a function of the square of suboptimal inflation:

$$Var(u_j) = c \cdot (\ln \Pi - \ln \Pi_j^*)^2 \tag{2}$$

where

$$c = \frac{N \cdot (N-1) \cdot (N+1)}{12} > 0$$

depends positively on the degree of price stickiness $N > 1$.

An important insight developed in this paper is the fact that the relationship between suboptimal inflation and inefficient price dispersion in equation (2) can actually be estimated using micro price data because (i) the product-specific optimal inflation rate Π_j^* is identified by the time trend in the sticky relative price, see figure 2, and (ii) price distortions, i.e., the gaps between the actual and the flexible price, are identified by the residuals of a regression of actual prices on a time trend, as illustrated in figure 2. Thus micro price data suffices to test whether price distortions vary with suboptimal inflation rates, i.e., whether $c > 0$, as predicted by sticky price theory.

While property (ii) fails to be true when the flexible price also depends on unobserved idiosyncratic shocks, we show in the next section that the presence of such shocks only requires adding a constant to equation (2). This holds true even when considering more plausible pricing setting frictions, such as Calvo or menu-cost frictions, as we show in the following section.

3 Identifying Inefficient Price Dispersion: Theory

This section uses sticky price theory to derive a regression equation that allows identifying the *marginal* effect of suboptimal inflation on inefficient price dispersion using micro price data. The regression approach turns out to be independent (to a second-order approximation) of whether price adjustment frictions are of a time-dependent or state-dependent nature and can be directly applied to micro price data. It does not require imposing any assumptions on the behavior of the cross-sectional distribution of flexible prices over time. Section 3.1 considers time-dependent price-setting frictions and section 3.2 presents the case with state-dependent frictions.

3.1 Time-Dependent Price Setting Frictions

The price setting problem. We consider the finest possible product specification in which a product j is a physical object or service sold in a specific location over time. Otherwise identical objects or services that are sold in different locations are treated as different products in our approach. The same holds true whenever an existing product gets substituted by a new product.

As before, let $p_{jt} \equiv P_{jt}/P_t$ denote the relative price charged for product j, where P_{jt} denotes the nominal product price and P_t the price index of all competing goods within a narrowly defined expenditure item.[13] Similarly, let p^*_{jt} denote the flexible relative price, i.e., the price the firm would like to charge in the absence of price setting frictions. The flexible price can differ from the (socially) efficient relative price.[14]

[13] In our empirical application, we will consider more than 1000 different expenditure items. To simplify notation, we do not separately index the expenditure item in this section.

[14] This may be due the presence of product-specific monopoly mark-ups. In the special case, where desired monopoly mark-ups are identical across products or simply

A second-order approximation to the nonlinear optimal price setting problem with Calvo price adjustment frictions is then given by[15]

$$\max_{\ln p_{jt}} -E_t \sum_{i=0}^{\infty} (\alpha\beta)^i \left(\ln p_{jt} - i \ln \Pi - \ln p^*_{jt+i}\right)^2 , \tag{3}$$

where the parameter $\beta \in (0,1)$ denotes the firm's discount factor, $\alpha \in (0,1)$ the Calvo probability that the price cannot be adjusted in the period, and Π the gross inflation rate. The firm's relative price in period $t+i$ is given by $\ln p_{jt} - i \ln \Pi$, which shows that the reset price $\ln p_{jt}$ chosen by the firm gets eroded over time by inflation, as long as prices fail to adjust. Deviations of the firm's relative from its flexible optimal price $\ln p^*_{jt+i}$ give rise to profit losses that are quadratic in the size of the deviation.

The dynamics of the flexible price. A key object of interest in problem (3) is the flexible (or frictionless) relative price p^*_{jt}. This price is observed by the firm but not by the econometrician. We consider the following general stochastic process:

$$\ln p^*_{jt} = \ln p^*_j - t \cdot \ln \Pi^*_j + \ln x_{jt}. \tag{4}$$

The term $\ln p^*_j$ is an unobserved product fixed-effect that is drawn at the time of product entry from some arbitrary and potentially time-varying distribution. It is a stand-in for unobserved location-specific effects such as difference in the level of marginal costs, wages, rents, service or quality components of the product. It also captures the presence of product and location-specific flexible price mark-ups.

The variable Π^*_j in equation (4) captures a product-specific time trend in the relative price and also denotes the product-specific optimal inflation rate, as discussed in section 2. It is drawn at the time of product entry from an arbitrary distribution that may also depend on time. The trend in relative prices may reflect a product-specific rate of productivity progress, induced for instance by learning-by-doing effects, or product-specific marginal cost trends induced by trends in wages or rents that are specific to the particular location where the product is sold. It is well-known that the strength of these effects varies across products[16] and we will exploit the variation in Π^*_j below to identify the distortionary effects of inflation. We consider a linear time trend in relative prices because the relative price dynamics of newly introduced products are

absent, the frictionless relative price is equal to the efficient relative price.

[15] See Appendix C.1 for a derivation.

[16] Adam and Weber (2022) document this for the U.K and Adam, Gautier, Santoro and Weber (2022) for France, Germany and Italy.

well-approximated by a linear trend.[17] Yet, in our empirical analysis we also consider nonlinear time trends.

Finally, there is an idiosyncratic stochastic component $\ln x_{jt}$ in equation (4), which captures fluctuations induced by changes in productivity or service components at the product level. The absence of a common component in these shocks is justified on the grounds that the left-hand side of equation (4) features the log *relative* price, thus absorbs common components in the nominal price (at the level of a narrowly-defined expenditure category). The stochastic process governing these idiosyncratic components is assumed to be the same for all products within a narrowly-defined expenditure category and satisfies the following restriction:

Assumption 1: Idiosyncratic shocks $\ln x_{jt}$ are stationary and Markov.

Assumption 1 effectively rules out that idiosyncratic shocks $\ln x_{jt}$ follow a random walk. This seems innocuous because our data strongly reject a random walk in $\ln x_{jt}$, as appendix B.[18] We can thus normalize idiosyncratic shocks so that $E[\ln x_{jt}] = 0$.

Note that the cross-sectional distribution of flexible prices is allowed to vary over time in important ways, even when abstracting from idiosyncratic shocks: (i) for a given set of products, heterogeneity in the relative price trends Π_j^* induces changes in the cross-sectional distribution of the flexible relative prices; (ii) as products exit and enter the market, newly entering products may have different product-specific intercepts p_j^* and time trends Π_j^* than exiting products. Since the parameters (p_j^*, Π_j^*) of newly incoming products are drawn from arbitrary time-varying distributions, our setup imposes no restrictions on the evolution of the cross-sectional distribution of flexible relative prices over time.

The optimal reset price. Considering the limit $\beta \to 1$, the optimal reset price $\ln p_{jt}^{opt}$ solving problem (3) is given by[19]

$$\ln p_{jt}^{opt} = (\ln p_{jt}^* - \ln x_{jt}) + \left(\frac{\alpha}{1-\alpha}\right)(\ln \Pi - \ln \Pi_j^*) + f(x_{jt}), \qquad (5)$$

where

$$f(x_{jt}) \equiv (1-\alpha) E_t \sum_{i=0}^{\infty} \alpha^i \ln x_{jt+i}. \qquad (6)$$

[17]See figure A.XI in the November 2018 working paper version of Argente and Yeh (2022), which depicts the relative price dynamics of newly introduced products using scanner data.

[18]This finding does not depend on assuming Calvo frictions.

[19]See appendix C.1 for a derivation.

The first term on the r.h.s. of equation (5), $\ln p^*_{jt} - \ln x_{jt}$, captures the deterministic component of the flexible price (4). The second term captures the effects induced by deviations of actual inflation $\ln \Pi$ from the product-specific optimal inflation rate $\ln \Pi^*_j$. The last term in equation (5) captures effects due to the presence of time-varying idiosyncratic components. Equation (6) shows that it is the expected value of the idiosyncratic shock over the lifetime of the price that matters for this component.

Only the second term on the r.h.s. of equation (5) depends on inflation. If actual inflation exceeds optimal inflation ($\ln \Pi > \ln \Pi^*_j$), then the reset price gets pushed up to compensate for the suboptimally high rate of future erosion of the relative price during periods in which the price does not adjust. The opposite is true if actual inflation falls short of optimal inflation ($\ln \Pi < \ln \Pi^*_j$).

Importantly, the optimal reset price $\ln p^{opt}_{jt}$ is equal to the expected value of the flexible price over the expected lifetime of the price. Therefore, an initial period in which relative prices lie above (below) the flexible price is followed - in expectation - by a period in which the relative price falls short (exceeds) of the flexible price. This explains how - according to the theory - deviations of inflation from its optimal level induce *additional* dispersion of prices around the flexible level. This effect is stronger if prices are more sticky: for a given deviation of inflation from its optimal level, reset prices react by more, the higher is the degree of price stickiness (α).

The dynamics of the actual relative price. While equation (5) determines the optimal reset price in periods where prices adjust, the dynamics of the actual relative price for product j are given by

$$\ln p_{jt} = \xi_{jt}(\ln p_{jt-1} - \ln \Pi) + (1 - \xi_{jt}) \ln p^{opt}_{jt}, \tag{7}$$

where $\xi_{jt} \in \{0, 1\}$ is an *iid* random variable capturing periods with price adjustment ($\xi_{jt} = 0$ with probability $1 - \alpha$) and no-adjustment ($\xi_{jt} = 1$ with probability α). In periods in which the price does not adjust, the relative price falls with inflation.

It also follows from equation (7) that the actual relative price inherits the product-specific time trend present in the optimal price p^{opt}_{jt}, which in inherits the trend from the flexible price p^*_{jt}, see equation (5). We show next that the variability of the actual price $\ln p_{jt}$ around this trend is a function of (i) the deviation of inflation from its optimal level, and (ii) the idiosyncratic shocks $\ln x_{jt}$. This insight turns out to be key for identifying the marginal effects of suboptimal inflation on inefficient price dispersion.

The first-stage regression. The first step in estimating the effects of inefficient price dispersion consists of running OLS regressions of the form

$$\ln p_{jt} = \ln a_j - (\ln b_j) \cdot t + u_{jt}, \tag{8}$$

which regress the relative product price on a product-specific intercept and time trend. To simplify the exposition, we abstract from small sample issues and focus on population regressions.[20] Regression (8) is of interest for two reasons. First, the coefficient estimates deliver[21]

$$\begin{aligned} \widehat{\ln a_j} &\rightarrow \ln p_j^* \\ \widehat{\ln b_j} &\rightarrow \ln \Pi_j^*, \end{aligned} \tag{9}$$

which shows that the regression allows recovering the deterministic components of the flexible relative price, i.e., the intercept term p_j^* and the product-specific optimal inflation rate Π_j^*. Since the actual relative price follows - in terms of its level and time trend - these deterministic dynamics, the effects of inefficient price distortions must be contained in the residuals of regression (8). In fact, these residuals are the second reason why regression (8) is of interest. They are asymptotically given by[22]

$$u_{jt} = \xi_{jt}(u_{jt-1} - (\ln \Pi - \ln \Pi_j^*)) + (1 - \xi_{jt})(f(x_{jt}) + \frac{\alpha}{1-\alpha}(\ln \Pi - \ln \Pi_j^*)) \tag{10}$$

where $\xi_{jt} = 0$ captures periods in which the price gets adjusted and $\xi_{jt} = 1$ captures periods without adjustment, and where $f(x_{jt})$ is defined in equation (6). We next discuss the properties of the the regression residuals (10).

The level of inefficient price dispersion is not identified. Due to price stickiness ($\alpha > 0$), the regression residuals u_{jt} in (10) fail to be very informative about the idiosyncratic shocks, as previously emphasized by Nakamura, Steinsson, Sun and Villar (2018). The underlying intuition is straightforward: in periods where prices do not get adjusted, they reveal no new information about idiosyncratic shocks; and in periods, where prices get adjusted, their adjustment gives considerable weight to expected future values of the idiosyncratic shock, particularly when prices are sticky, see equation (6).

Due to the influence of expected future shock values, the information that becomes available upon a price adjustment, i.e. the term $f(x_{jt})$

[20] Small sample effects are discussed in detail in appendix E.
[21] See appendix C.2 for a formal derivation.
[22] See appendix C.3 for a derivation.

defined in equation (6), fails to identify the underlying process of idiosyncratic shocks $\ln x_{jt}$. Appendix A proves the following result:

Proposition 1 *In the presence of price stickiness, observed prices* $\ln p_{jt}$ *fail to identify the process for idiosyncratic shocks* $\ln x_{jt}$*. Consider, for example, a stationary discrete* N*-state Markov process for* $f(x_{jt})$*. It can be generated either by a stationary Markov processes for* $\ln x_{jt}$ *with* N *states or an infinite number of different Markov processes with* $M \geq N$ *states, where* M *is arbitrary and where* $M - N$ *states in the* M*-state process are not states in the* N*-state process.*

Intuitively, different fundamental processes for $\ln x_{jt}$ give rise to identical processes for $f(x_{jt})$, because they imply the same conditional expectations in equation (6). Since the process for $\ln x_{jt}$ cannot be identified from observed prices, it is impossible to estimate 'price distortions', i.e., the gap between the actual and flexible price. This may explain why the literature has to date not come up with an estimate of how inefficient price dispersion responds to (suboptimal) inflation.

It is worth emphasizing that the result in proposition 1 applies more generally to the case where $\ln x_{jt}$ is non-stationary but still contains some stationary component, e.g., when $\ln x_{jt}$ is the sum of a random walk process $\ln y_{jt}$ plus an independent stationary Markov process $\ln z_{jt}$. We then have $f(\ln x_{jt}) = \ln y_{jt} + f(\ln z_{jt})$, so that the process $\ln z_{jt}$ and thus $\ln x_{jt}$ can again not be identified, even if the process for $\ln y_{jt}$ could be perfectly recovered from the data.

One way to deal with the identification problem is to bring in additional information. This is the strategy pursued in Eichenbaum, Jaimovich and Rebelo (2011) who exploit information on marginal costs in supermarkets to identify price distortions (but do not analyze how they depend on inflation). Yet, information on marginal costs is only rarely available.

An alternative approach to handle the identification problem is to impose additional identification assumptions. This is the approach pursued in Baley and Blanco (2021) and Alvarez, Lippi and Oskolkov (2022), who show that the distribution of price distortions can be recovered from observed price changes, whenever $\ln x_{jt}$ is a pure random walk, i.e., does not contain stationary shock components. With a random walk, we have $f(x_{jt}) = \ln x_{jt}$, so that the size of innovations between price reset times identifies the innovation variance of the random walk. Yet, the hypothesis of a pure random walk in $\ln x_{jt}$ is strongly rejected in our data, as we show in appendix B.

We now show that it is simply not necessary to identify the *level* of price distortions to estimate the marginal effects of suboptimal inflation on price distortions. We discuss this point in the next subsection.

Second-stage regression: the marginal effect of suboptimal inflation. While the level of price distortions cannot be identified from observed prices, the theory predicts that the *marginal effect* of suboptimal inflation on price dispersion can be identified. In fact, equation (5) highlights that any non-zero gap $\ln \Pi - \ln \Pi_j^*$ generates front-loading of prices upon price adjustment times, as captured by the term $\frac{\alpha}{1-\alpha}(\ln \Pi - \ln \Pi_j^*)$. Likewise, during non-adjustment periods, a gap between actual and optimal inflation leads to a drift in the gap between actual and flexible relative prices. Both of these features contribute to increasing the variance of the regression u_{jt} in the first-stage regression (10).

Therefore, the variance of first-stage residuals satisfies the following relationship:[23]

Proposition 2 *The variance of the first-stage residual in equation (8) is given by*

$$Var(u_{jt}) = v + c \cdot (\ln \Pi - \ln \Pi_j^*)^2, \tag{11}$$

where the intercept

$$v \equiv Var\left((1-\alpha)E_t \sum_{i=0}^{\infty} \alpha^i \ln x_{jt+i}\right) \tag{12}$$

is a function of the idiosyncratic shock process $\ln x_{jt}$ *and the price stickiness parameter* α*, and*

$$c \equiv \frac{\alpha}{(1-\alpha)^2}. \tag{13}$$

The intercept term v in equation (11) contains both efficient price components, e.g., the presence of idiosyncratic fundamental shocks, and inefficient price components that arise due to price stickiness, see equation (12). In particular, price stickiness causes the loading on the current idiosyncratic shocks to be too low relative to the flexible price case. Without additional information, it is impossible to further decompose to what extent v reflects efficient or inefficient forces, which is precisely the feature preventing identification of the level of price distortions from observations of actual prices. The second term on the r.h.s. of equation (11) captures the effects of suboptimal inflation on inefficient price dispersion. According to the theory, the coefficient c is an increasing function of the degree of the Calvo price stickiness parameter α.

[23] See appendix C.3 for a derivation.

Equation (11) is a second-stage regression equation and a key equation we shall exploit in the present paper. It uses the residual variance from the first-stage equation (8) as left-hand side variable, and the gap between the (item-level) inflation rate Π and the product-specific optimal inflation Π_j^* as right-hand side variable, where Π_j^* is also identified from the first-stage regression, see equation (9). Equation (11) implies that the marginal effect of suboptimal inflation on inefficient price dispersion can be estimated using a cross-section of products for which price stickiness and the process driving idiosyncratic shocks are the same.

Appendix E describes in detail the two stage estimation approach that allows estimating the coefficient c. It shows that the second-stage estimate for c is biased towards zero, due to the presence of first-stage estimation error. The second-stage estimate of c thus provide a *lower bound* of the true marginal effect of suboptimal inflation on price distortions. Since we are interested in rejecting the null hypothesis of inflation *not* creating inefficient price dispersion ($H_0 : c = 0$), this works against our main finding.

The next section briefly shows that the results derived thus far are not specific to the case with Calvo frictions, but also apply in a setting with menu-cost frictions.

3.2 State-Dependent Price Setting Frictions

We now present a model with state-dependent pricing. To be able to get closed-form solutions, we consider a continuous-time setup and a slightly more restrictive process for the idiosyncratic shocks. Within this setup, we derive continuous-time analogue to proposition 2. The firm's objective (3) becomes:

$$\max_{\{\tau_{ji}, \Delta \ln p_{ji}\}_{i=1}^{\infty}} -E\left[\int_t^{\infty} e^{-\rho(s-t)} \left(\ln p_{jt+s} - \ln p_{jt+s}^*\right)^2 ds + \kappa \sum_{i=1}^{\infty} e^{-\rho(\tau_{ji}-t)}\right] \tag{14}$$

The parameter $\rho > 0$ is the discount rate, τ_{ji} are the random adjustment times and κ is the cost paid at the times of adjustment. As with time-dependent frictions, the firm's relative price in period $\tau_{ji} + s$ is given by $\ln p_{j\tau_{ji}} - s \ln \Pi$ between adjustment periods, reflecting relative price erosion due to inflation.

The flexible relative price $\ln p_{jt}^*$ follows a continuous-time analogue of (4) with an additional restriction on the idiosyncratic process $\ln x_{jt}$, namely that it assumes values from a finite grid $\{\ln x_1, \ldots, \ln x_N\}$ and switches from grid point i to grid point j with Poisson intensity λ_{ij}^X.[24]

[24]The restriction is very mild because we do not impose any assumption on the

Appendix D shows that under $\rho \to 0$ and for sufficiently small adjustment cost κ,[25] the OLS regression (8) recovers the exact same coefficients as in the time-dependent model. Furthermore, the variance of residuals depends on product-specific suboptimal inflation:

$$Var(u_{jt}) = Var(\ln x) + c^{MC} \cdot (\ln \Pi - \ln \Pi_j^*)^2 + O((\ln \Pi/\Pi_j^*)^4), \quad (15)$$

where the intercept is again a function of the idiosyncratic shock process, the quadratic term depends on suboptimal inflation, and $O((\ln \Pi/\Pi_j^*)^4)$ denotes a fourth order approximation error. The coefficient c^{MC} is now a function of the shock process parameters $\lambda_i^X = \sum_{j \neq i}^N \lambda_{ij}^X$:

$$c^{MC} \equiv E\left[\frac{1}{(\lambda_i^X)^2}\right].$$

If λ_i^X is constant across states, then

$$c^{MC} = \frac{1}{\Lambda^2} \quad (16)$$

where Λ is equal to the adjustment frequency (again up to a fourth order approximation error $O((\ln \Pi/\Pi_j^*)^4)$) and thus can be directly estimated from the data. The coefficient c^{MC} differs slightly from the one in the discrete time setup with Calvo friction, see equation (13), for which $\Lambda = 1 - \alpha$. This is so because multiple price adjustments can happen per unit of time under continuous time modeling. Notice also that the coefficient c^{MC} does not depend on the menu cost κ, under the maintained assumption that menu costs are small enough. Differences in κ have only fourth order effects in equation (15). This is the reason why equation (15) now holds only up to a fourth-order approximation error, while it was exact in the Calvo setup (given the quadratic approximation to the firm objective), see equation (11).

Perhaps surprisingly, the results obtained from the state-dependent model are (to the consider order of approximation) virtually the same as for the time-dependent model.

4 Data Description

We use the micro price data underlying the official U.K. consumer price index (CPI) for the sample period February 1996 to December 2016,

switching intensities. Even though we are ruling out all processes with continuous paths, we can still approximate them well with a sufficiently fine grid.

[25]Note that we *do not* consider a limiting case $\kappa \to 0$, instead our result holds *for all* $\kappa \leq \bar{\kappa}$ for some $\bar{\kappa} > 0$.

as obtained from the Office of National Statistics (ONS). The data are monthly and classified into narrowly defined expenditure items (e.g., flat panel TV 33inch, men's shoes trainers, vegetarian main course, etc.). Given the sample selection described further below, we consider 1033 different expenditure items and 15.4 million price observations over the considered 20 year period.

A product within an item is a sequence of price observations for a particular object or service sold in a particular store. The product life ends whenever a product substitution takes place or when the product rotates out of the sample.

We then estimate the first-stage equation (8) for every product in the sample and estimate the second-stage equation (11) at the level of the expenditure item $z = 1, ...1033$, considering all products j belonging to the item, i.e., we estimate

$$\widehat{Var}(u_{jzt}) = v_z + c_z \cdot (\ln \widehat{\Pi_z/\Pi^*_{jz}})^2 + \varepsilon_{jz} \tag{17}$$

where $\widehat{Var}(u_{jzt})$ is the variance of first-stage residuals of product j in item z and $\ln \widehat{\Pi_z/\Pi^*_{jz}}$ the corresponding first-stage estimate of the gap between the item-level inflation rate and product-specific optimal inflation.[26] Estimation of equation (17) delivers 1033 estimates c_z, one for each expenditure item. We focus in our analysis on the item-level rather than on the aggregate level because doing so increases the chances that our two key identifying assumptions (identical degrees of price rigidity & identical stochastic processes driving idiosyncratic shocks) are satisfied.

The data methodology follows the one used in Adam and Weber (2022), who provide further details. Starting from the raw micro price data, we delete products with duplicate price observations in a given month[27] and also delete all price observations flagged by ONS as "invalid." Furthermore, we split observed price trajectories for ONS product identifiers, whenever ONS indicates a change in the underlying product, i.e., a comparable or non-comparable product substitution, and whenever price quotes are missing for two months or more. This conservative splitting approach insures that we do not lump together products that might in fact be different. It leads to a refined product definition that we use to compute relative prices by deflating nominal product prices with a quality-adjusted item price index.

[26] See appendix E for details of the estimation approach, including arguents showing why two-stage estimation approach only biases the coefficient c_z towards zero, i.e., against finding a role for suboptimal inflation on inefficient price dispersion.

[27] Duplicate price quotes can arise because the U.K. Office of National Statistics (ONS) does not disclose all available locational information underlying the data, so that in rare cases we cannot uniquely identify the product price.

Total number of price quotes used	15.4 million			
	mean	median	min	max
Number of products per item	816	637	101	3,490
Number of price quotes per item	14,861	11,095	608	73,301

Table 1: Basic product and price statistics

We only include expenditure items for which the item price index, computed from our micro price data, replicates the official item price index provided by ONS sufficiently well. This leads to a selection of 1093 expenditure items from the 1233 contained in the raw data. Furthermore, we only consider products with a minimum length of six price observations and expenditure items containing at least one hundred of such products.[28] This leads us to the 1033 expenditure items that we use in our empirical analysis.[29] Table 1 reports basic statistics on the number of products and price observations.

4.1 Descriptive Statistics of the Regression Inputs

This section presents key descriptive statistics about the variables entering the first and second-stage regression equations. Since we run these regressions for more than one thousand expenditure items, we report the distribution of key moments of the variables of interest in the cross-section of items.

The left column in figure 3 depicts the distribution of the mean and standard deviation of the length of product life. For most items, the mean product length ranges between 10 and 25 months, which is long enough to estimate an intercept and slope parameter in our first-stage. The bottom left panel in figure 3 highlights that there is a considerable amount of variation in the length of product lives within each item. We exploit this feature below to present estimates that are based on products whose price can be observed for at least twelve (instead of six) months.

The top right panel in figure 3 reports the distribution of the mean R^2 values of the first-stage regression (46) across items. For most items, the intercept and time trend tend to capture on average between 30% and 50% of the observed variation in relative prices. The remainder of the

[28] We also eliminate expenditure items for which the estimated residual variances are zero for all products. The latter occurs when prices never adjust within an item, which is the case for less than a handful of items capturing administered prices.

[29] Not all these items are present throughout the sample period, as expenditure items get added and removed. For the average year, we have 503 expenditure items.

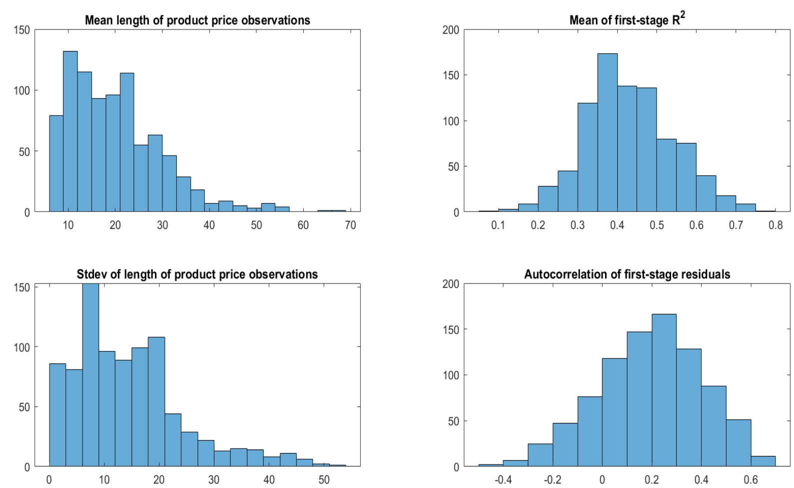

Figure 3: Descriptive statistics: first-stage regression

variation goes into the regression residual, the variance of which enters our second-stage regression. The bottom right panel in figure 3 depicts the distribution of the mean autocorrelation of these residuals. The autocorrelation is significantly below one, showing that the assumption of a random walk is implausible given our data.[30]

The top left panel of figure 4 reports the mean standard deviation of the regression residual across items.[31] For most items, the average standard deviation ranges between 2% and 4%. The standard deviation of the standard deviation of residuals is shown in the bottom left panel of figure 4. It highlights that there is a considerable amount of variation in the left-hand side variable of our second-stage regression, which is desirable.

The top right panel in figure 4 depicts the distribution of item-level means of the suboptimal inflation rate.[32] For the vast majority of items, the average suboptimal inflation rate lies between ±0.5% per month. The lower right panel in figure 4 shows the within-item standard deviation of suboptimal inflation. The cross-product variation is significant, with a standard deviation ranging between 1/3 and 2/3 of a percent on

[30]See appendix B for formal tests of the random walk hypothesis, which are based on price observations form price adjustment periods.

[31]We report moments of the non-squared variables entering the second-stage regression to increase readability of the figures.

[32]See appendix F for information on the cross-sectional distribution of the product-specific optimal inflation rate Π^*_{jz}.

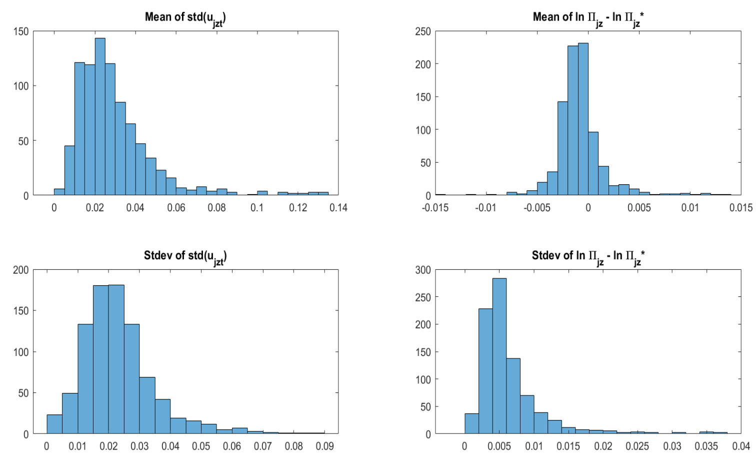

Figure 4: Descriptive statistics: second-stage regression

a monthly basis in most items. This shows that our second-stage right-hand side variable also displays a considerable amount of variation.

5 Price Distortions at the Product Level: Empirical Results

This section reports our estimates of the coefficient c_z in equation (48), which captures how suboptimal inflation distorts relative prices. We perform the estimation at the level of 1033 finely disaggregated expenditure categories, the so-called U.K. expenditure items, to ensure that the key identifying assumptions of identical degrees of price rigidity and identical stochastic processes for idiosyncratic disturbances hold.

Our baseline results are presented in section 5.1 and the subsequent section documents their robustness along a number of dimensions. An alternative estimation approach, which exploits only within-product variation, is presented in section 5.3 and leads to very similar conclusions.

5.1 Baseline Results

Figure 5 presents overwhelming evidence of the notion that suboptimal inflation gives rise to price distortions at the product level. It depicts the distributions of the estimated intercepts (top left panel) and the coefficients of interest c_z (top right panel) obtained from estimating equation (17) across 1033 expenditure categories z. The coefficient c_z captures the marginal effect of suboptimal inflation on inefficient price dispersion. In line with the underlying sticky-price theories, 97% of the estimated co-

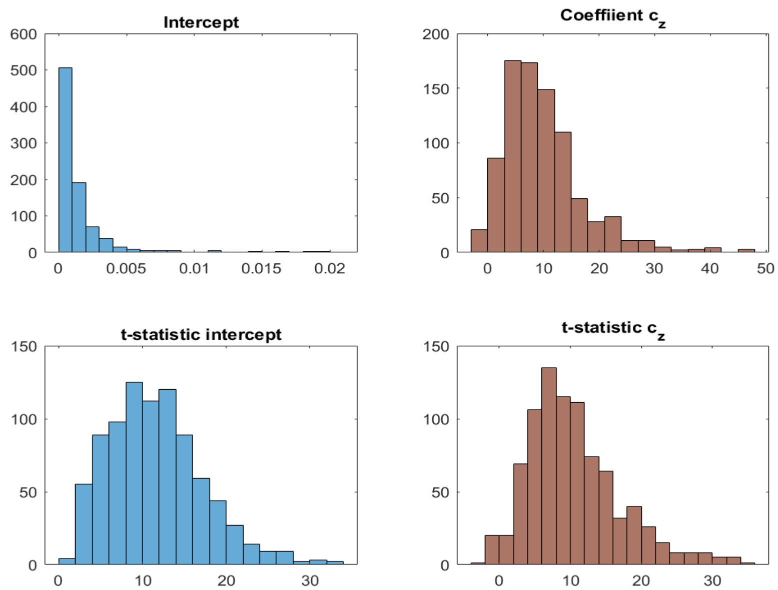

Figure 5: Baseline results from estimating equation (17)

efficients are positive.

The bottom row in figure 5 reports the distribution of t-statistics of the coefficients shown in the corresponding top panels. It shows that 95% of the estimated coefficients c_z have a t-statistic larger than two, while only 0.5% have a t-statistic below minus two.

Moreover, the mean R^2 value of the second-stage regression (17) is 17%, which highlights that suboptimal inflation explains a sizable part of the cross-product variance of first-stage residuals.[33] This is the case despite first-stage estimation error contributing to unexplained variance in our second stage regression.

The point estimates for c_z are not only positive and statistically significant for the vast majority of expenditure categories, but also quantitatively large: the average point estimate is equal to 10 and implies that a monthly inflation rate that lies 1% above (or below) its optimal level[34] increases the variance of the first-stage residual by 0.1% and thus

[33] Recall that item-specific constants do not contribute to the R^2 values of the second-stage regressions (17).

[34] The 1% number is approximately equal to two times the median value of the standard deviation of the inflation gap in the data, see the lower right panel in figure 4.

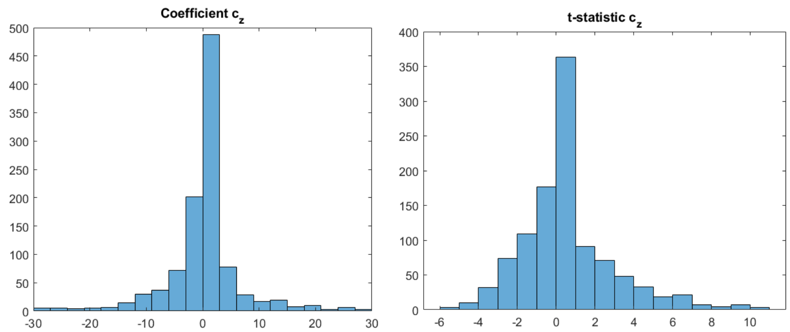

Figure 6: Estimation with actual inflation only on r.h.s. (equation (18))

its standard deviation by 3.2 percentage points, which appears sizable.[35] Since first-stage estimation error causes the second-stage estimates of c_z to be biased towards zero, we refrain here from a further quantitative interpretation of the point estimates. Instead, we will assess in section 6 the quantitative importance of relative price distortions using (unbiased) first-stage estimates only.

Inflation versus Suboptimal Inflation. It turns out to be important for our empirical results that the right-hand side of equation (17) features *suboptimal* inflation rather than simply the level of inflation. To show this, we estimate the alternative second-stage regression

$$\widehat{Var}(u_{jzt}) = v_z + c_z \cdot (\widehat{\ln \Pi_{jz}})^2 + \varepsilon_{jz}, \tag{18}$$

where $\widehat{\ln \Pi_{jz}}$ denotes the average item-level inflation rate during the lifetime of product j. This specification counterfactually imposes zero optimal inflation rate for all products. Figure 6 shows that the coefficients c_z are then more or less symmetrically centered around zero with most of them being statistically insignificant: only 20% of coefficients have a t-statistic larger than two, while 12% have a t-statistic below minus two. Thus, one would wrongly conclude that inflation produces no price dispersion, if one assumed that the product-specific optimal inflation is equal to zero, as the simplest sticky price models actually suggest. This result also highlights that our baseline findings emerge due to cross-sectional variation in the product-specific *optimal* inflation rate.

[35] The reported increase in the standard deviation assumes that fundamental shocks and measurement error are absent ($v_z = \varepsilon_{jz} = 0$) in equation (17).

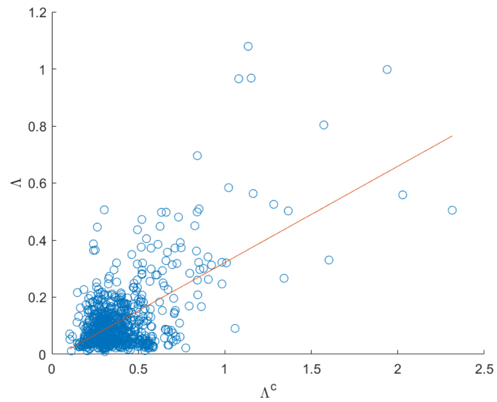

Figure 7: Obseved (Λ_z) and estimation-implied (Λ_z^c) price adjustment rates.

Price Stickiness. The underlying sticky price theories suggest that the coefficient c_z increases in the degree of price stickiness, see equations (13) and (16). We now investigate the empirical relationship between the estimated coefficient c_z and the observed price adjustment rate at the item level.

In a continuous time setup, with constant price adjustment rate Λ_z in item z, the share of non-adjusters per unit of is equal to $e^{-\Lambda_z}$. The monthly share of non-adjusters α_z in item z can be measured directly from the data, which provides an estimate of the implied instantaneous price adjustment rate:

$$\Lambda_z = -\ln \alpha_z.$$

We compare this estimate to the estimate implied by our regression coefficients c_z. To this end, we back out an *implied* instantaneous price adjustment rate from c_z using our theory. This delivers[36]

$$\Lambda_z^c = \sqrt{1/\left(c_z + \frac{1}{12}\right)}.$$

[36] This follows from equation (13), which implies

$$\begin{aligned} c_z &= \frac{\alpha_z}{(1-\alpha_z)^2} = \frac{e^{-\Lambda_z}}{(1-e^{-\Lambda_z})^2} \\ &= \frac{1}{(\Lambda_z)^2} - \frac{1}{12} + O(\Lambda_z^2), \end{aligned}$$

where the approximation in the last equality is taken for the limit $\Lambda_z \to 0$. Using the menu cost setup, we arrive at very similar implications, namely: $c_z = 1/(\Lambda_z)^2 - 1/12$.

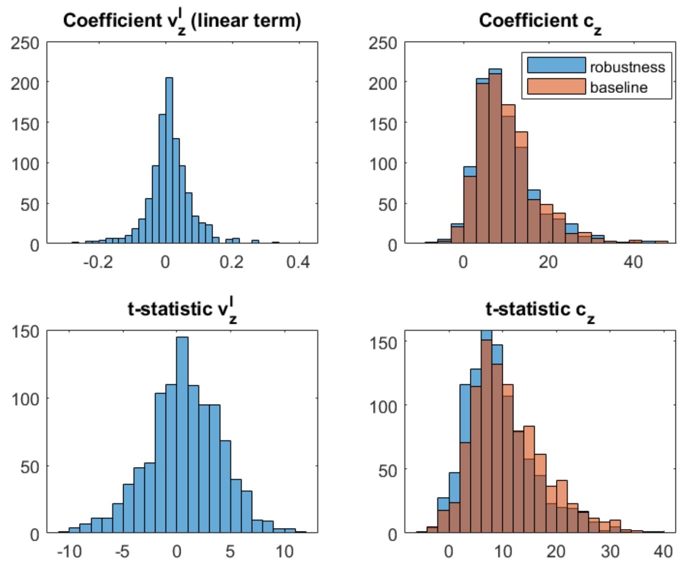

Figure 8: Robustness to adding a linear term (equation (19))

Figure 7 presents a scatter plot with observed price adjustment rate (Λ_z) on the y-axis and the estimation-implied adjustment rate (Λ_z^c) on the x-axis. While both measures display a strongly positive correlation equal to +0.6, the linear regression line in figure 7 has a slope that is equal to 0.34 only, while the predicted slope is unity. This is consistent with a substantial downward bias in our estimated coefficients c_z, due to the presence of first-stage estimation error. Nevertheless, the positive correlation between observed and implied price adjustment rates is encouraging and in line with the underlying theory.

5.2 Robustness of Baseline Approach

We now explore the robustness of our baseline results in a number of directions.

Adding Linear Terms. Sticky price theories predict that only the squared deviation of inflation from its optimal level should explain variance of the first-stage regression residuals. In particular, a linear term consisting of the gap between inflation and its optimal level should have a zero coefficient. We test this overidentifying restriction by running regressions of the form

$$\widehat{Var}(u_{jzt}) = v_z + v_z^l \cdot \ln \widehat{\Pi_z/\Pi_{jz}^*} + c_z \cdot (\ln \widehat{\Pi_z/\Pi_{jz}^*})^2 + \varepsilon_{jzt} \qquad (19)$$

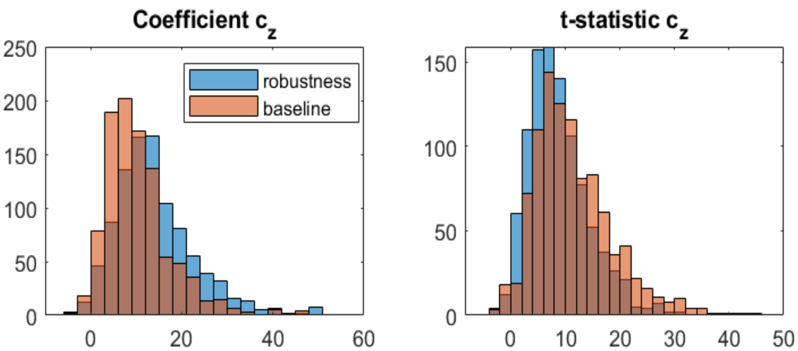

Figure 9: Longer price series in first stage (at least 12 price observations)

and checking whether $v_z^l \approx 0$. We also check whether the estimates for the coefficients c_z remain unaffected by the presence of the linear term.

Figure 8 reports the distribution of the estimates of v_z^l (top left panel), the estimated coefficient c_z (top right panel), and the distribution of t-statistics for these coefficients (corresponding bottom panels). In line with sticky price theory, the coefficients v_z^l are indeed tightly centered around zero. Moreover, the distribution of estimated c_z and the distribution of associated t-statistics hardly move relative to the baseline.

Reducing First-Stage Estimation Error. One possible econometric concern with the baseline estimation approach is that first-stage estimation errors are large and lead to substantial attenuation in the second stage. The quantitative relevance of possible attenuation effects can be evaluated by restricting attention to products in the first stage sample to products with a sufficiently large number of price observations. Thus, we consider only products with at least 12 monthly price observations (rather than six). If the concern is valid, this subsample should increase the point estimates in the second stage.

The left panel in figure 9 compares the distribution of the estimated coefficients c_z to the baseline outcome. As expected, the distribution of estimates moves to the right, suggesting that first stage estimation errors lead to a downward bias in the second stage baseline estimates. However, the distribution of t-statistics, displayed in the right panel of figure 9, remains largely unaffected, possibly due to the decline in the number of observations in the second-stage.

Including Sales Prices. Our baseline estimation removes all sales prices from the sample, mainly because the underlying sticky price theo-

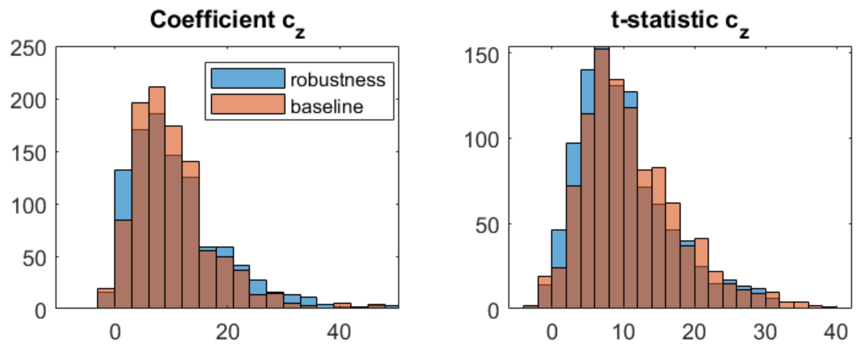

Figure 10: Estimation including sales prices

ries typically do not model sales. Our results are robust to including also sales prices in the estimation, as shown in figure 10: the distribution of estimated coefficients c_z (left panel) and t-statistics (right panel) remain almost unaffected by the inclusion of sales prices.

Nonlinear Time Trends. Our baseline approach allows for a linear time trend in relative prices in the first-stage regression equation (8). Importantly, the linearity assumption is no crucial for our empirical results. To illustrate this, we recompute the first-stage residuals allowing also a quadratic time trend and then use the residuals obtained this ways in our second-stage regression (11). The outcome is depicted in figure 11. It shows that the point estimates for the coefficient c_z are then somewhat closer to zero, possibly due to attenuation effects associated with a more noisy first-stage estimate of the linear coefficient. Still, more than 97% of the point estimates are positive and 93% of all coefficients have a positive t-statistic larger than two, which is very similar to the baseline result.

5.3 Exploiting Within Product Variation

The baseline estimation exploits cross-sectional variation across products within narrowly defined expenditure categories to identify the coefficient c_z. An important identifying assumption is that idiosyncratic shocks are driven by the same stochastic process for all products within an item.

Equations (11) and (15) show that product-specific processes for idiosyncratic shocks would give rise to product-specific intercepts in the second-stage regression. If unobserved shock heterogeneity varies systematically with our second-stage regressor (the squared of suboptimal inflation), then second stage estimates could be driven by heterogeneity in unobserved shock processes, rather than the effect of inflation on price

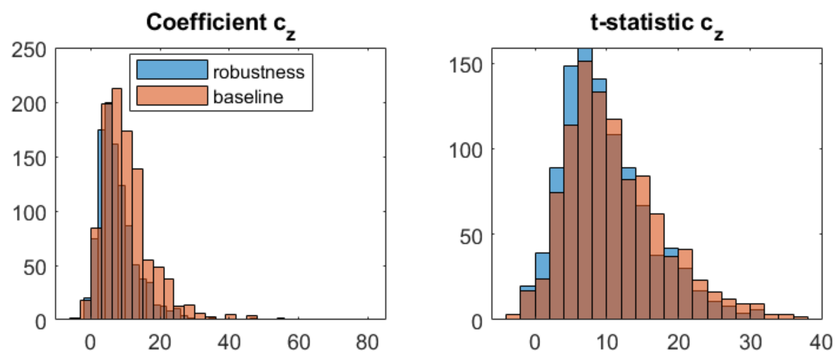

Figure 11: Nonlinear time trend in first-stage regression

dispersion.

We address this concern by also estimating the coefficient of interest c_z using only within-product variation. Specifically, we split the life of each product into its first and second half and exploit variation in item-level inflation across the two lifetime subsamples to estimate c_z. Since equation (17) holds for either subsample, we can then take first differences of equation (17) and estimate

$$\widehat{Var}_1(u_{jzt}) - \widehat{Var}_2(u_{jzt}) = c_z \left(\left(\ln \widehat{\Pi_z^1/\Pi_{jz}^*} \right)^2 - \left(\ln \widehat{\Pi_z^2/\Pi_{jz}^*} \right)^2 \right) + \varepsilon_{jz}, \tag{20}$$

where $\widehat{Var}_1(u_{jzt})$ and $\widehat{Var}_2(u_{jzt})$ denote the residual variances in the first and second half of the product lifetime, respectively, and $\ln \widehat{\Pi_z^1/\Pi_{jz}^*}$ and $\ln \widehat{\Pi_z^2/\Pi_{jz}^*}$ the respective suboptimal inflation rate.[37] Specification (20) is considerably more demanding than our baseline specification, as it relies purely on variation of suboptimal inflation over the product lifetime, rather than on cross-product variation in suboptimal inflation, as is the case in our baseline.

Figure 12 depicts the regression outcomes and compares them to our baseline findings.[38] The estimated coefficients in the left panel tend to be even larger than in our baseline specification. At the same time, they tend to be estimated less precisely, causing a shift in the distribution of t-statistics in the right panel of figure 12 towards zero. Nevertheless, the

[37] This rate can be obtained by estimating equation (47) in appendix E separately for the first and second half of product lifetime.

[38] To alleviate concerns about a missing constant, we also add a constant to regression equation (20). The specification without the constant leads to very similar outcomes.

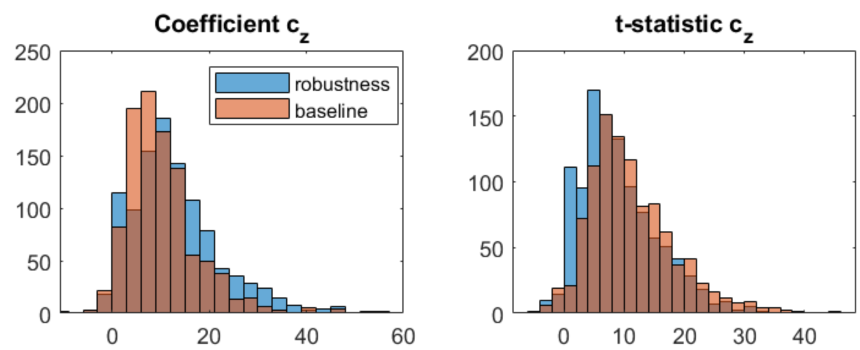

Figure 12: Estimation using within product variation (equation (20))

share of point estimates with a t-statistic larger than two still stands at 88%, while only 2% of the estimated coefficients have a t-statistic below minus two. This strongly suggests that our baseline results are not driven by unobserved product heterogeneity in the underlying idiosyncratic shock processes.

6 Cross-Sectional Price Dispersion over Time

The analysis so far focused on price distortions at the product and item level. This section considers the cross-sectional distribution of prices and shows how to decompose it into (i) a component capturing elements of the flexible price distribution, and (ii) and a remainder identifying variation in inefficient price dispersion over time. Inefficient price dispersion covaries positively with aggregate inflation over time, as shown in our introductory figure 1, and is sizable in absolute terms. At the same time, inefficient price dispersions is modest in size relative to overall price dispersion. As one might expect, most time-series variation in overall price dispersion is explained by variation in component (i).

6.1 Decomposing Cross-Sectional Price Dispersion

From the sticky price theories analyzed in section 3, it follows that the price of product j in expenditure category z evolves *over time* according to[39]

$$\ln p_{jzt} = \ln p^*_{jz} - \ln \Pi^*_{jz} \cdot t + u_{jzt}, \tag{21}$$

[39]See equation (38) in appendix C.2 for the case with Calvo frictions and equation (39) in appendix D for the case with menu costs.

where the residuals u_{jzt} have mean zero, are independent across j and z, and have variance *over time* equal to

$$Var(u_{jzt}) = v_z + c_z \cdot (\ln \Pi_z - \ln \Pi_{jz}^*)^2. \tag{22}$$

We are now interested in decomposing the *cross-sectional* variance of prices, denoted by $Var^j(\ln p_{jzt})$, of the products present at some time t in some item z. We then evaluate how this measure of *cross-sectional* price dispersion depends on the item-level inflation rate Π_z.

Suppose there is a unit mass of products in item z and that a share of products randomly exits the sample each period and gets replaced by newly sampled products. Newly sampled products may have different characteristics than the products that leave the sample. Thus the distribution of product characteristics $\{p_{jz}^*, \Pi_{jz}^*\}$ within the item may change over time.[40]

Appendix F shows that time variation in the distribution of optimal inflation rates is minor, which allows us to consider a time-invariant cross-sectional distribution of optimal inflation rates $\{\Pi_{jz}^*\}$. Specifically, we assume that upon entry of a product into the sample, the optimal inflation rate Π_{jz}^* is an i.i.d. draw from $\{\Pi_z^{*1}, \Pi_z^{*2}, ..., \Pi_z^{*I}\}$, where Π_z^{*i} is chosen with probability $m_z^i \geq 0$, $\sum_i m_z^i = 1$. In contrast, the distribution of estimated intercepts $\{p_{jz}^*\}$ is strongly moving over time in the data. We thus allow for arbitrary time variation in the distribution of intercepts for newly incoming products.[41]

Given this setup, we obtain the following decomposition result:[42]

Proposition 3 *Let $Var^j(\cdot)$ denote the variance in the cross-section of products j. Then, the cross-sectional dispersion of prices in expenditure category z at time t is given by*

$$Var^j(\ln p_{jzt}) = Var^j(\ln p_{jz}^* - \ln \Pi_{jz}^* \cdot t) + Var^j(u_{jzt}), \tag{23}$$

where

$$Var^j(u_{jzt}) = v_z + c_z \cdot \sum_i (\ln \Pi_z - \ln \Pi_z^{*i})^2 m_z^i. \tag{24}$$

[40] We assume that upon the time of entry, the residual u_{jzt} is drawn from the stationary residual distribution for products with characteristics (p_{jz}^*, Π_{jz}^*). This is justified by the fact that newly sampled products in our data typically do not represent truly new products, instead products that are newly sampled by the Office of National Statistics.

[41] The covariance between the distribution of intercepts $\{p_{jz}^*\}$ and optimal inflation rates $\{\Pi_{jz}^*\}$ is also left unrestricted.

[42] See appendix 3 for the proof.

Equation (23) decomposes the cross-sectional price dispersion into two components. The first component, $Var^j(\ln p^*_{jz} - \ln \Pi^*_{jz} \cdot t)$, captures identifiable elements of the flexible price distribution. We obtain this component using our first stage estimates of (a_{jz}, b_{jz}) from equation (8). The second component, $Var^j(u_{jzt})$, depends on the constant v_z and $c_z \cdot \sum_i (\ln \Pi_z - \ln \Pi_z^{*i})^2 m_z^i$, which captures - according to the theory - price distortions induced by inflation.[43] Using our our first-stage residuals, we can estimate $Var^j(u_{jzt})$ as $Var^j(\widehat{u}_{jzt})$. The decomposition in proposition 3 holds at each point in time in setting where inflation Π_z is constant.

The decomposition also applies in a setting where inflation is changing slowly over time.[44] Equation (24) thus provides a theory-implied relationship linking (yearly) inflation rates Π_{zt} to the cross-sectional distribution of first-stage residuals $Var^j(u_{jzt})$ (at the end of the year). Depending on the distribution $\{\Pi_z^{*i}\}$ of optimal inflation rates, an increase in Π_{zt} can lead to either an increase or a decrease in inefficient price dispersion: if average optimal inflation rate $(\sum_i \Pi_z^{*i} m_z^i)$ lies below actual inflation, then inefficient dispersion is predicted to increase with inflation. The opposite is true if the average optimal inflation rate lies above actual inflation. It is thus an empirical question whether higher inflation rates lead to more or less inefficient price dispersion in the data, which is investigated in the next section.

6.2 Inflation and Cross-Sectional Price Distortions over Time

The present section investigates the comovement between inflation and inefficient cross-sectional price dispersion over time. It considers first comovement at the level of expenditure items and then comovement at the aggregate level.

6.2.1 Item Level Results

From proposition 3 and the subsequent discussion follows that inefficient price distortions correlate positively (negatively) with item-level inflation over time, whenever the average optimal inflation rate $(\sum_i \Pi_z^{*i} m_z^i)$ lies below (above) actual inflation. For inflation rates close to the average optimal level, this correlation is predicted to be close to zero.

[43] In the absence of price stickiness, we have $c_z = 0$ so that price dispersion does not depend on inflation.

[44] Suppose inflation changes from year to year and that the inflation rate is equal to Π_{zt} in year t. If price setters expect future inflation to be equal to Π_{zt}, then our steady-state pricing results continue to apply. (Such expectations are justified whenever Π_{zt} follows a random walk.) And since the vast majority of prices adjusts over the course of a year, inflation rates from earlier years will not matter for the cross-sectional dispersion of prices.

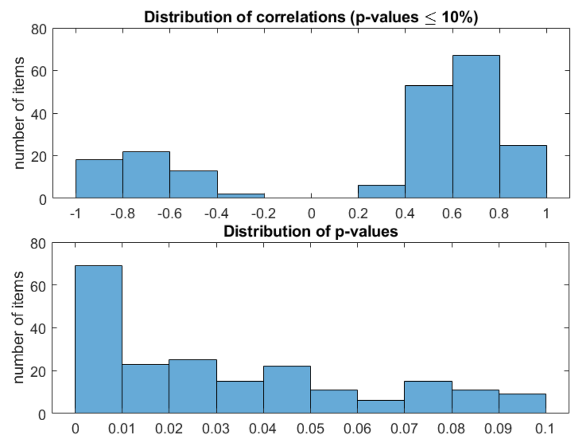

Figure 13: Correlation between inflation and price distortions at the item level

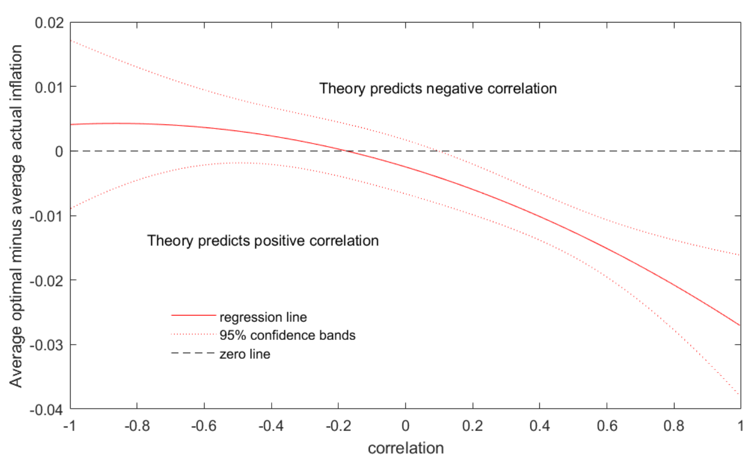

Figure 14: The gap between average optimal and actual inflation (y-axis) determines the correlation between inflation and inefficient price dispersion (x-axis)

To test this prediction, we compute the correlation between the cross-sectional variance $Var^j(u_{jzt})$ and item-level inflation Π_{zt} over time, using all items for which we have at least three years of data.[45] The top panel in figure 13 depicts the resulting distribution of correlations across items, using all correlations with a p-value less or equal to 10%. The bottom panel depicts the distribution of p-values.[46] Figure 13 shows that there are significantly positive and significantly negative correlations, but more positive than negative ones.[47] Proposition 3 implies that positive (negative) correlations should emerge whenever optimal average inflation ($\sum_i \Pi_z^{*i} m_z^i$) lies above the average actual inflation rate in the item. Figure 14 shows that this is indeed the case: the figure depicts the outcome of a regression of the gap between optimal and actual inflation on the correlation and its square. The regression line behaves in line with theory, with the statistically significant parts being fully aligned with the theoretical prediction.[48]

6.2.2 Aggregate Results

We now consider an economy-wide measure of inefficient price dispersion. To this end, we aggregate the residual variances (24) across items to an economy-wide dispersion measure using item-level expenditure weights. We then compare time variation in aggregated residual dispersion with time variation in aggregate inflation, as aggregate inflation is also an expenditure-weighted average of item level inflation.

The bottom panel of our introductory figure 1 depicts the resulting aggregate price dispersion measure together with the aggregate inflation rate.[49] Both measures covary positively, with a correlation equal to +0.58 that is significant at the 1% level. This shows that higher aggregate inflation rates are associated with larger amounts of relative price distortions because time variation in the dispersion of first-stage residuals captures time variation in relative price distortions.

Importantly, this result is not driven by outliers in the distribution of first-stage residuals. For instance, results are similar when removing at the item level the 2.5% highest and 2.5% lowest residuals before com-

[45] This is the case for 696 expenditure items.

[46] 206 of the 696 correlations have p-values smaller than 10%.

[47] This result is robust to choosing tighter p-values, e.g., a value of 5%, or to considering all correlations, independently of their p-value.

[48] This continues to be true when restricting consideration to a linear regression or when including a third order term into the regression. However, the regression coefficient on the third order term is not statistically signficant.

[49] Figure 1 displays annual dispersion and annual inflation to remove within-year seasonalities in price dispersion and inflation. Both measures are computed as a 12 month average of monthly dispersion and monthly year-over-year inflation rate.

puting the variance of first-stage residuals. Likewise, computing instead a robust dispersion measure, as in Nakamura, Steinsson, Sun and Villar (2018), leads to very similar outcomes.[50]

By proposition 3 , the positive correlation between aggregate inflation and inefficient price dispersion is driven by products with optimal inflation rate Π^*_{jz} that lie below the actual inflation rate.

This theory prediction can again be tested. To do so, we now group individual products according to their optimal inflation rate. Specifically, we consider the 1/3 of products with the highest and the 1/3 of products with the lowest optimal inflation rate in each expenditure item and then recompute inefficient price dispersion for these two subgroups.[51]

The top group of products has an (unweighted) average optimal inflation rate that varies between +2.5% and +5.5% over time, which roughly covers the range in which actual inflation moves. The bottom group, however, has a deeply negative optimal inflation rate that ranges between −6% and −10% over time. According to the theory, this group should display a strong positive correlation between inefficient price dispersion and inflation over time. In contrast, the top group should display no or only a weak correlation with inflation.

This is indeed what we find: the correlation between inflation and inefficient dispersion is weak (+0.19) and statistically insignificant (p-value of 0.40) for the top group, but strongly positive (+0.54) and significant (p-value of 0.01) for the bottom group. In line with sticky price theory, the positive correlation between inflation and inefficient price dispersion at the aggregate level is thus driven by products with optimal inflation rates that lie below actual inflation.

6.3 Bounds on Price Distortions and Flexible Price Dispersion

This section derives upper and lower bounds on the amount of relative price distortions from inflation and discusses the drivers of the trends present in overall price dispersion. Inefficient price distortion is sizable, giving rise - on its own - to a standard deviation of log prices of several percentage points. At the same time, the amount of price dispersion

[50] Following their approach, we aggregated the interquartile range (IQR) of first-stage residuals at the level of each expenditure item and then use the expenditure-weighted median to aggregate across items.

[51] We split products within each expenditure item, rather splitting products across all items combined, to avoid that results are driven by compositional effects. As is well-known, the average optimal inflation rates varies systematically in the cross section of items.

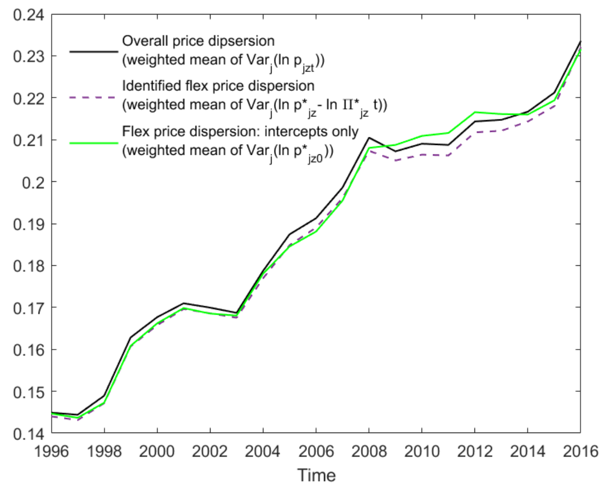

Figure 15: Overall price dispersion versus flexible price dispersion (various identified components)

due to inflation-induced distortions is dwarfed by the dispersion already present under flexible prices.

It follows from proposition 3 that the (aggregated) variance of first stage residuals represents an upper bound on the amount of inefficient price dispersion that is due to inflation.[52] The upper bound of the variance reached in the lower panel of figure 1 is approximately $2.5 \cdot 10^{-3}$. Therefore, absent any flexible price dispersion, inefficient price dispersion gives rise to a standard deviation of prices of at most $\sqrt{2.5 \cdot 10^{-3}} = 5\%$. While this is quantitatively large, inefficient dispersion accounts for about 1% of the overall price dispersion, see the upper panel of figure 1.

A lower bound on the contribution of inflation to inefficient price dispersion is given by the min-max range of the variance of first-stage residuals, as the time-varying component is - according to the theory - due to inflation. This range is approximately equal to $1.5 \cdot 10^{-3}$ and implies (in the absence of flexible price dispersion) that inflation would induce variation in the standard deviation of prices of up to $\sqrt{1.5 \cdot 10^{-3}} = 3.87\%$ over time. Again, this appears sizable in absolute terms.

The overall dispersion of prices, however, is overwhelmingly driven by price dispersion that is also present under flexible prices. Figure 15 depicts overall price dispersion (the expenditure-weighted item level variances $Var^j(\ln p_{jzt})$) together with the dispersion of the identifiable components of the flexible price dispersion (the expenditure-weighted

[52] This is so because the intercept v_z in equation (24) is not due to inflation.

item level variances $Var^j(\ln p^*_{jz} - \ln \Pi^*_{jz} \cdot t)$).

Figure 15 shows that the identifiable component of flexible price dispersion makes up for the vast majority of the observed price dispersion and also closely tracks it over time. Since time variation in the distribution of optimal inflation rates ($\ln \Pi^*_{jz}$) is very limited, virtually all time-series variation is coming from time-series variation in the cross-sectional dispersion of the intercepts ($\ln p^*_{jz}$).[53]

This illustrates that time series variation in overall price dispersion is to a large extent driven by time series variation in flexible price dispersion, which is also strongly rising over time.

Increasing flexible price dispersion may reflect a number of economic forces, such as widening in the distribution of (flexible price) mark-ups, the productivity distribution across products, or the distribution of unobserved product qualities. The large increase in flexible price dispersion explains why aggregate inflation fails to covary with overall price dispersion (top panel in our introductory figure 1).

7 Conclusions

This paper documents three important new facts: (1) at the product level, deviations of inflation from its product-specific optimal level are robustly associated with an increase in inefficient price dispersion; (2) at the aggregate level, time-series variation in aggregate inflation covaries positively with aggregate measures of inefficient price dispersion, with this relationship being driven mainly by products whose optimal inflation rate is low; and (3) the time series variation in overall price dispersion is closely tracked by the identifiable part of flexible price dispersion, which strongly suggests that time series fluctuations in overall price dispersion are driven by fundamentals other than inflation.

Taken together, these findings provide considerable support for the economic mechanisms enshrined in sticky price models and the monetary policy conclusions they give rise to.

In future work, we seek to investigate to what extent inefficient price dispersion is associated with inefficient dispersion in product demand (misallocation). This requires observing product quantities, in addition to observing product prices. While this is feasible only for a much narrower set of products, studying the misallocation of quantities is an important step that could provide further support for the view that sub-optimal inflation gives rise to significant welfare costs.

[53] To make comparisons meaningful over time, figure 15 reports the dispersion coming from intercepts using the normalized intercepts $\ln p^*_{jz} - \Pi^*_{jz} \cdot t^0_{jz}$, where t^0_{jz} is the time period in which the product first enters the sample.

References

Acharya, S., E. Challe, and K. Dogra (2023): "Optimal Monetary Policy According to HANK," *American Economic Review (forthcoming)*.

Adam, K., E. Gautier, S. Santoro, and H. Weber (2022): "The Case for a Positive Euro Area Inflation Target: Evidence from France, Germany and Italy," *Journal of Monetary Economics (forthcoming)*.

Adam, K., and H. Weber (2019): "Optimal Trend Inflation," *American Economic Review*, 109(2), 702–737.

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

Alvarez, F., M. Beraja, M. Gonzalez-Rozada, and P. A. Neumeyer (2019): "From Hyperinflation to Stable Prices: ArgentinaŠs Evidence on Menu Cost Models," *The Quarterly Journal of Economics*, 134(1), 451–505.

Alvarez, F., F. Lippi, and A. Oskolov (2022): "The Macroeconomics of Sticky Prices and Generalized Hazard Functions," *Quarterly Journal of Eonomics*, 137, 989–1038.

Argente, D., and C. Yeh (2022): "Product Life Cycle, Learning, and Nominal Shocks," *The Review of Economic Studies*, 89(6), 2992Ű3054.

Ascari, G., P. Bonmolo, and Q. Haque (2022): "The Long-Run Phillips Curve is ... a Curve," Dutch Central Bank mimeo.

Baley, I., and A. Blanco (2021): "Aggregate Dynamics in Lumpy Economies," *Econometrica*, 89, 1235–1264.

Baqaee, D., E. Farhi, and K. Sangani (2022): "Supply Side Effects of Monetary Policy," NBER Working Paper 28345.

Eichenbaum, M., N. Jaimovich, and S. Rebelo (2011): "Reference prices, costs, and nominal rigidities," *American Economic Review*, 101(1), 234–62.

Galí, J. (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New KeynesianFramework and Its Applications*. Second Edition, Princeton University Press.

Meier, M., and T. Reinelt (2022): "Monetary Policy, Markup Dispersion, and Aggregate TFP," *Review of Economics and Statistics (forthcoming)*.

Nakamura, E., J. Steinsson, P. Sun, and D. Villar (2018): "The Elusive Costs of Inflation: Price Dispersion during the U.S. Great Inflation," *The Quarterly Journal of Economics*, 133(4), 1933–1980.

Sara-Zaror, F. (2022): "Expected Inflation and Welfare: The Role of Consumer Search," Federal Reserve Bank of San Francisco mimeo.

Sheremirov, V. (2020): "Price Dispersion and Inflation: New Facts and Theoretical Implications," *Journal of Monetary Economics*, 114, 59–70.

Taylor, J. B. (1979): "Staggered wage setting in a macro model," *American Economic Review, Papers and Proceedings*, 69(2), 108Ű13.

Woodford, M. (2003): *Interest and Prices*. Princeton University Press, Princeton.

Wulfsberg, F. (2016): "Inflation and Price Adjustment: Micro Evidence from Norwegian Consumer Prices 1975-2004," *American Economic Journal: Macroeconomics*, 8, 175–193.

A Proof of Proposition 1

In this section we prove that it is impossible to recover the price gap distribution if shocks are stationary. Suppose an econometrician observes the infinite path of actual prices $\ln p_{jt}$ and it is known that this path is generated under the time-dependent friction and stationary shocks $\ln x_{jt}$. The econometrician can recover the N values of the vector $\mathbf{f} \equiv [f_1, \ldots, f_N]'$ of $f(x_{jt})$ as defined in (6):

$$f(x_{jt}) \equiv (1-\alpha)E_t \sum_{i=0}^{\infty} (\alpha)^i \ln x_{jt+i}.$$

In addition, the econometrician can recover the $N \times N$ transition matrix Λ^f:

$$\Lambda^f = \begin{bmatrix} \lambda_{11}^f & \cdots & \lambda_{1N}^f \\ \vdots & \ddots & \vdots \\ \lambda_{N1}^f & \cdots & \lambda_{NN}^f \end{bmatrix},$$

where λ_{ij}^f is the probability of observing f_j in the subsequent period, conditional on observing f_i in the previous period.[54] From the definition of $f(x_{jt})$ it follows that:

$$\mathbf{f} = (1-\alpha)\ln \mathbf{x} + \alpha \Lambda^x \mathbf{f}$$

where $\ln \mathbf{x}$ is the state vector of the process $\ln x_{jt}$ and Λ^x is its transition matrix. Setting $\Lambda^x = \Lambda^f$ and solving the above equation for $\ln \mathbf{x} \equiv [\ln x_1, \ldots, \ln x_N]$ provides a candidate for the process $\ln x_{jt}$ that leads to the observed process $f(x_{jt})$. However, as we show below, this candidate solution is not unique and the observed N-state process of $f(x_{jt})$ can be equally supported by an $(N+1)$-state process $\ln \tilde{x}_{jt}$, defined on the grid $\ln \tilde{\mathbf{x}} \equiv [\ln \tilde{x}_1, \ldots, \ln \tilde{x}_N, \ln \tilde{x}_{N+1}]$ with $(N+1) \times (N+1)$ transition matrix $\tilde{\Lambda}^x$. Such a process would lead to an $(N+1)$-state process of $\tilde{f}(x_{jt})$, with $\tilde{f}_i = f_i$ for all $i < N$ and $\tilde{f}_N = \tilde{f}_{N+1} = f_N$, making $\tilde{f}(x_{jt})$ and $f(x_{jt})$ observationally equivalent, provided the transition probabilities of $\tilde{\Lambda}^x$ imply Λ^f. To construct such a process, set $\ln \tilde{x}_i = \ln x_i$ for all $i < N$, $\ln \tilde{x}_N = \ln x_N - \varepsilon$ and $\ln \tilde{x}_{N+1} = \ln x_N + \varepsilon$ for a sufficiently small $\varepsilon > 0$.[55]

[54] This can be achieved by conditioning on price spells of length one.

[55] One requirement for ε is that $\ln \tilde{x}_N$ and $\ln \tilde{x}_{N+1}$ do not coincide with existing values of $\ln x_i$. A stricter condition on the size of ε is introduced below.

We now construct the transition matrix $\tilde{\Lambda}^x$ in the following way:

$$\tilde{\Lambda}^x = \left[\begin{array}{cccc|cc} \lambda_{11}^x & \lambda_{12}^x & \dots & \lambda_{1(N-1)}^x & \lambda_{1N}^x/2 & \lambda_{1N}^x/2 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \\ \lambda_{(N-1)1}^x & \lambda_{(N-1)2}^x & \dots & \lambda_{(N-1)(N-1)}^x & \lambda_{(N-1)N}^x/2 & \lambda_{(N-1)N}^x/2 \\ \hline \tilde{\lambda}_{N1}^x & \lambda_{N2}^x & \dots & \lambda_{N(N-1)}^x & \tilde{\lambda}_N^x & \tilde{\lambda}_N^x \\ \tilde{\lambda}_{(N+1)1}^x & \lambda_{N2}^x & \dots & \lambda_{(N+1)(N-1)}^x & \tilde{\lambda}_{N+1}^x & \tilde{\lambda}_{N+1}^x \end{array}\right]$$

All elements in black are borrowed directly from the Λ^x matrix, whereas elements in red are to be solved for.[56] The first $(N-1)$ rows of $\tilde{\Lambda}^x$ ensure that for all $i < N$:

$$\begin{aligned} \tilde{f}_i &= (1-\alpha)\ln \tilde{x}_i + \alpha \sum_{j=1}^{N+1} \tilde{\lambda}_{ij}^x \tilde{f}_j \\ &= (1-\alpha)\ln x_i + \alpha \sum_{j=1}^{N-1} \lambda_{ij}^x f_j + \left(\frac{\lambda_{iN}^x}{2} + \frac{\lambda_{iN}^x}{2}\right) f_N = f_i \end{aligned}$$

We now have to set the elements in red ($\tilde{\lambda}_{N1}^x$, $\tilde{\lambda}_N^x$, $\tilde{\lambda}_{(N+1)1}^x$, $\tilde{\lambda}_{N+1}^x$) such that $\tilde{f}_N = \tilde{f}_{N+1} = f_N$. For $i = N$ it requires:

$$\begin{aligned} \tilde{f}_N &= (1-\alpha)(\ln x_N - \varepsilon) + \alpha \tilde{\lambda}_{N1}^x f_1 + \alpha \sum_{j=2}^{N-1} \lambda_{Nj}^x f_j + 2\tilde{\lambda}_N^x f_N \\ &= f_N - (1-\alpha)\varepsilon + \alpha(\tilde{\lambda}_{N1}^x - \lambda_{N1}^x) f_1 + \alpha(2\tilde{\lambda}_N^x - \lambda_{NN}^x) f_N \stackrel{!}{=} f_N \end{aligned}$$

Denote $\sum_{j=2}^{N-1} \lambda_{Nj}^x \equiv \lambda$, then it must be the case that $\tilde{\lambda}_{N1}^x + \lambda + 2\tilde{\lambda}_N^x = 1$ to ensure that $\tilde{\Lambda}^x$ is a proper transition matrix. The same applies to the elements of Λ^x: $\lambda_{N1}^x + \lambda + \lambda_{NN}^x = 1$. Substituting $\tilde{\lambda}_N^x$ and λ_{NN}^x in the above equation and rearranging terms yields:

$$\tilde{\lambda}_{N1}^x = \lambda_{N1}^x + \frac{1-\alpha}{\alpha} \frac{\varepsilon}{f_1 - f_N}$$

For $i = N+1$, a similar line of arguments leads to:

$$\tilde{\lambda}_{(N+1)1}^x = \lambda_{N1}^x - \frac{1-\alpha}{\alpha} \frac{\varepsilon}{f_1 - f_N}$$

and the remaining elements $\tilde{\lambda}_N^x$ and $\tilde{\lambda}_{N+1}^x$ can then be recovered using the fact that all rows of $\tilde{\Lambda}^x$ sum up to one. ε must be small enough to ensure

[56] We order states such that $\lambda_{N1}^x > 0$ and $\lambda_{NN}^x > 0$. This is without loss of generality since $\ln x_{jt}$ is a stochastic process, implying that there exists a state i such that for at least two states j_1 and j_2, $\lambda_{ij_1}^x > 0$ and $\lambda_{ij_2}^x > 0$.

that $\tilde{\lambda}^x_{N1}$, $\tilde{\lambda}^x_{N}$, $\tilde{\lambda}^x_{(N+1)1}$ and $\tilde{\lambda}^x_{N+1}$ are all $\in [0,1]$. Such ε always exists since we have ordered the states to ensure $\lambda^x_{N1} > 0$ and $\lambda^x_{NN} > 0$ and there are infinitely many of them. It remains to show that transition probabilities in $\tilde{\Lambda}^x$ imply Λ^f. This holds trivially for all transitions between states f_i and f_j such that $i, j < N$. It is also true for transitions from f_i to f_N when $i < N$ since the probability of transiting from f_i to f_N is then equal to $\frac{\lambda^x_{iN}}{2} + \frac{\lambda^x_{iN}}{2} = \lambda^x_{iN}$. Finally, note that states $\ln x_N$ and $\ln x_{N+1}$ have the same unconditional probability,[57] and therefore the probability of moving from f_N to f_i is equal to $\frac{1}{2}\left(\tilde{\lambda}^x_{Ni} + \tilde{\lambda}^x_{(N+1)i}\right) = \lambda^x_{Ni}$ for all $i < N$. This implies that the probability of staying in f_N is also the same as in the original process (λ^x_{NN}).

Therefore, we have constructed an N+1-state process $\ln \tilde{x}_{jt}$ that leads to the same process $f(x_{jt})$ as the N-state process $\ln x_{jt}$. By induction this step can be repeated arbitrary many times.

B Testing for a Random Walk in Idiosyncratic Shocks

This appendix shows that our data strongly rejects the presence of a pure random walk in $\ln x_{jt}$. One can test for a random walk in $\ln x_{jt}$ by exploiting the fact that the optimal reset price upon price adjustment involves a constant gap relative to the flexible price, whenever $\ln x_{jt}$ is a random walk. This holds true with Calvo frictions, see equation (5), but also for the case with menu cost frictions.

Consider the times t_n ($n = 1, 2, ...N_{jz}$) during which the price of some product j in expenditure item z adjusts. Given the constant gap property, we have

$$\ln p^{opt}_{jzt_{n+1}} - \ln p^{opt}_{jzt_n} = -\ln \Pi^*_{jz} \cdot (t_{n+1} - t_n) + \ln e_{jzn+1} \tag{25}$$

where

$$\ln e_{jzn+1} \equiv \ln x_{jzt_{n+1}} - \ln x_{jzt_n}.$$

With a random walk in $\ln x$, the residuals $\ln e$ are uncorrelated over time and have adjustment-time-dependent variance $(t_{n+1} - t_n)\sigma^2_z$, where σ^2_z denotes the innovation variance in the random walk in expenditure item z. These two properties can be tested.

To test for the adjustment-time-dependent variance, we use all observations $(t_{n+1} - t_n, e_{jzn+1})$ within some item z to run the regression

$$(\ln e_{jzn+1})^2 = a_z + b_z(t_{n+1} - t_n) \tag{26}$$

[57] The unconditional probability satisfies $p = (\tilde{\Lambda}^x)'p$, and the last two columns of $\tilde{\Lambda}^x$ are identical, implying identical values of p_N and p_{N+1}.

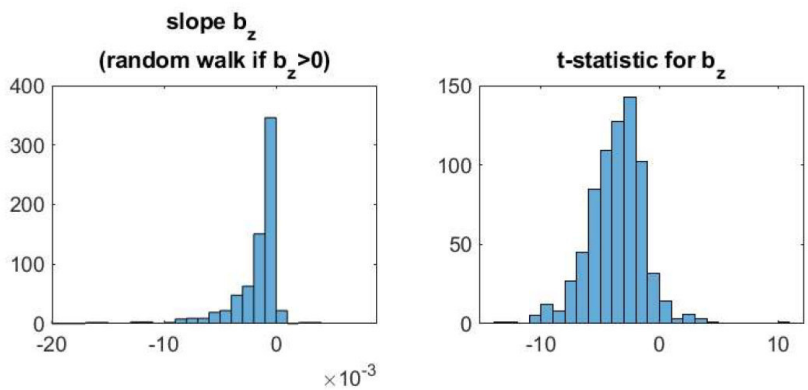

Figure 16: Random walk test, equation (26)

and check whether $b_z = \sigma_z^2 > 0$ as predicted by the random walk. Figure 16 reports the distribution of the estimated b_z and the associated t-statistics using all products with $N_{jz} > 3$. It shows that the random walk hypothesis $b_z > 0$ is strongly rejected by the data.

Second, we can also test if the residuals $\ln e$ in (25) are uncorrelated over time. To do so, we re-scale residuals according to $(\ln e_{jzn+1}) / \sqrt{t_{n+1} - t_n}$ to make them homoskedastic under the null hypothesis of a random walk. We then compute the autocorrelations $\widehat{Corr}_z = \widehat{Cov}_z / \widehat{Var}_z$ of these re-scaled residuals within each item z, using the variance and covariance estimates for all products with $N_{jz} > 3$:

$$\widehat{Var}_z = \sum_j \left(\frac{N_{jz} - 2}{\sum_k (N_{kz} - 2)} \sum_{n=2}^{N_{jz}} \frac{\left(\frac{\ln e_{jzn}}{\sqrt{t_n - t_{n-1}}} \right)^2}{N_{jz} - 2} \right)$$

$$\widehat{Cov}_z = \sum_j \left(\frac{N_{jz} - 3}{\sum_k (N_{kz} - 3)} \sum_{n=2}^{N_{jz}-1} \frac{\frac{\ln e_{jzn}}{\sqrt{t_n - t_{n-1}}} \frac{\ln e_{jzn+1}}{\sqrt{t_{n+1} - t_n}}}{N_{jz} - 3} \right)$$

The left panel in figure 17 depicts the estimated autocorrelations across items. Almost all of the estimates are negative, and most of them sizably so, which is inconsistent with $\ln x_{jzt}$ following a random walk. The right panel in the figure reports the bootstrapped p-values for the autocorrelation being weakly larger than zero, as implied by the random walk, and shows that these values are very low.

We then repeat the analysis when exogenously imposing $\Pi_{jz}^* = 0$ for all products in the first-stage regression. This is motivated by the possibility that the estimated time trends Π_{jz}^* could be purely spurious

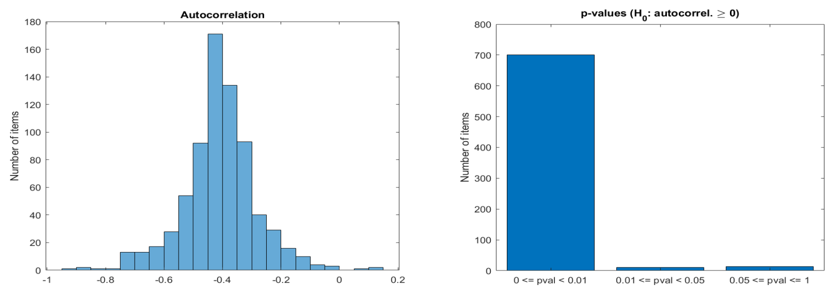

Figure 17: Autocorrelation of residuals and bootstrapped p-values (random walk implies autocorrelation of zero)

in the presence of a random walk in $\ln x_{jzt}$. While the estimated coefficients b_z in (26) are then symmetrically centered around zero (but still not predominantly positive), the evidence on the auto-correlation of the residuals remains almost identical to the one shown in figure 17 for the case with an estimated time trend Π^*_{jz} in the first-stage regression.

Based on these findings, we conclude that unobserved shocks in our data do *not* follow a pure random walk.

C Details of the Calvo Model

C.1 Quadratic approximation and optimal reset price

The price-setting problem of firm j in price-adjustment period t is to set its price P_{jt} such that it maximizes the expected discounted sum of period profits,

$$E_t \sum_{i=0}^{\infty} \alpha^i \Omega_{t,t+i} \left[(1+\tau) \frac{P_{jt}}{P_{t+i}} - mc_{jt+i} \right] Y_{jt+i} \tag{27}$$

$$\text{subject to} \quad Y_{jt+i} = \left(\frac{P_{jt}}{P_{t+i}} \right)^{-\theta} Y_{t+i}, \tag{28}$$

where $\Omega_{t,t+i}$ denotes the stochastic discount factor of the representative household, Y_{jt} denotes output of product j, mc_{jt} real marginal costs and τ is a sales subsidy. Equation (28) states product demand of the cost-minimizing household that derives utility from a CES consumption composite aggregating the products in the economy with substitution elasticity $\theta > 1$.

We approximate the profit objective (27) around a balanced growth path in which aggregate output and consumption grow at constant rate γ^e, and where gross inflation is equal to Π. With growth-consistent preferences that exhibit constant relative risk aversion, the household discount factor is given by $\Omega = \omega\left(\gamma^e\right)^{-\sigma} < 1$, where σ denotes relative risk aversion and ω is the rate of time preference. Substituting this expression for the discount factor into equation (27) yields

$$E_t \sum_{i=0}^{\infty} \left(\alpha\omega\left(\gamma^e\right)^{1-\sigma}\right)^i \left[(1+\tau)\widetilde{p}_{jt+i} - mc_{jt+i}\right]\left(\widetilde{p}_{jt+i}\right)^{-\theta} y \qquad (29)$$

where $\widetilde{p}_{jt+i} = p_{jt}\Pi^{-i}$ and $p_{jt} \equiv P_{jt}/P_t$, and y denotes detrended aggregate output. Real marginal costs evolve according to

$$mc_{jt} = mc_j e^{-(\ln \Pi_j^*)t} x_{jt}, \qquad (30)$$

where the product-specific time-fixed effect mc_j is drawn randomly at the time of product entry from some arbitrary distribution. As described in the main text, Π_j^* is the efficient rate of relative price decline under flexible prices and x_{jt} is a stationary process with $E[\ln x_{jt}] = 0$.

By equation (29), the objective for period $t+i$ is given by

$$D_{jt+i} = \left[(1+\tau)e^{\ln \widetilde{p}_{jt+i}} - e^{\ln mc_{jt+i}}\right]\left(e^{\ln \widetilde{p}_{jt+i}}\right)^{-\theta} y. \qquad (31)$$

We approximate this objective to second order in the variables $\ln \widetilde{p}_{jt+i}$ and $\ln mc_{jt+i}$ around the deterministic paths of the flexible price and marginal costs, respectively. The deterministic path of the flexible price is equal to

$$\vartheta mc_{jt+i}^{\text{det}}$$

where mc_{jt}^{det} denotes the deterministic path of marginal costs which is equal to the value of marginal costs mc_{jt} imposing $x_{jt} = 1$, and $\vartheta = \frac{\theta}{\theta-1}\frac{1}{1+\tau}$ denotes the flexible-price markup.

The second-order Taylor approximation of equation (31) yields

$$\begin{aligned} D_{jt+i} &= \left(y\vartheta^{-\theta}\right) e^{(1-\theta)\ln mc_{jt+i}^{\text{det}}}\big[-\theta(\ln \widetilde{p}_{jt+i} - \ln(\vartheta mc_{jt+i}^{\text{det}}))^2 \\ &\quad + 2\theta(\ln \widetilde{p}_{jt+i} - \ln(\vartheta mc_{jt+i}^{\text{det}}))(\ln mc_{jt+i} - \ln mc_{t+i}^{\text{det}})\big] + O(3) \\ &= \left(-\theta y\vartheta^{-\theta}\right)\left(mc_{jt+i}^{\text{det}}\right)^{1-\theta}\left[\ln \widetilde{p}_{jt+i} - \ln(\vartheta mc_{jt+i}^{\text{det}}) - (\ln mc_{jt+i} - \ln mc_{t+i}^{\text{det}})\right]^2 + \text{t.i.p.} + O(3) \\ &= \left(-\theta y\vartheta^{-\theta}\right)\left(mc_{jt+i}^{\text{det}}\right)^{1-\theta}\left[\ln \widetilde{p}_{jt+i} - \ln(\vartheta mc_{jt+i})\right]^2 + \text{t.i.p.} + O(3), \end{aligned} \qquad (32)$$

where t.i.p. collects terms independent of policy and it follows from equation (30) that $mc_{jt+i}^{\text{det}} = mc_j e^{-(\ln \Pi_j^*)(t+i)}$. Thus, we rewrite the Taylor approximation coefficient in the previous equation according to

$$-\theta y\vartheta^{-\theta}\left(mc_j e^{-(\ln \Pi_j^*)(t+i)}\right)^{1-\theta} = -\theta y\vartheta^{-\theta} mc_j^{1-\theta}(\Pi_j^*)^{(\theta-1)(t+i)}.$$

We can now express the expected discounted sum of period profits in equation (29) accurate to second order according to

$$-\theta y\vartheta^{-\theta} mc_j^{1-\theta}(\Pi_j^*)^{(\theta-1)t} E_t \sum_{i=0}^{\infty}(\alpha\omega\left(\gamma^e\right)^{1-\sigma}(\Pi_j^*)^{\theta-1})^i \left[\ln \widetilde{p}_{jt+i} - \ln(\vartheta mc_{jt+i})\right]^2 + \text{t.i.p.} + O(3)$$

which is proportional to

$$-E_t \sum_{i=0}^{\infty}(\alpha\beta_j)^i \left[\ln p_{jt} - i\ln\Pi - \ln(p_{jt+i}^*)\right]^2 + \text{t.i.p.} + O(3) \qquad (33)$$

after substituting $\widetilde{p}_{jt+i} = p_{jt}\Pi^{-i}$ and denoting the firm discount factor by $\beta_j = \omega\left(\gamma^e\right)^{1-\sigma}(\Pi_j^*)^{\theta-1}$ and defining

$$p_{jt+i}^* = \vartheta mc_{jt+i}$$

which implies using equation (30)

$$p_{jt}^* = p_j^* e^{-(\ln \Pi_j^*)t} x_{jt},$$

which is equal to (4) for $p_j^* = \vartheta mc_j$. While p_{jt}^* denotes the firm's flexible price, the ratio of two firms' flexible prices is equal to the efficient relative price for these firms, whenever price mark-ups are constant across firms and time. In this special case, p_{jt}^* denotes also the efficient relative price.

We can then express the flexible price in period $t+i$ as

$$p_{jt+i}^* = p_{jt}^* e^{-(\ln \Pi_j^*)i} x_{jt+i} x_{jt}^{-1}.$$

and substitute into equation (33), which delivers

$$\max_{\ln p_{jt}} -E_t \sum_{i=0}^{\infty}(\alpha\beta_j)^i \left(\ln p_{jt} - i\ln(\Pi/\Pi_j^*) - \ln p_{jt}^* - \ln x_{jt+i} + \ln x_{jt}\right)^2. \qquad (34)$$

The first-order condition is given by

$$0 = -2E_t \sum_{i=0}^{\infty}(\alpha\beta_j)^i \left(\ln p_{jt}^{opt} - i\ln(\Pi/\Pi_j^*) - \ln p_{jt}^* - \ln x_{jt+i} + \ln x_{jt}\right),$$

which implies that the optimal price is given by

$$\ln p_{jt}^{opt} = \ln p_{jt}^* - \ln x_{jt} + \left(\frac{\alpha\beta_j}{1-\alpha\beta_j}\right)\ln(\Pi/\Pi_j^*) + E_t(1-\alpha\beta_j)\sum_{i=0}^{\infty}(\alpha\beta_j)^i \ln x_{jt+i} \qquad (35)$$

since $\sum_{i=0}^{\infty}(\alpha\beta_j)^i i = \sum_{i=1}^{\infty}(\alpha\beta_j)^i i = \frac{\alpha\beta_j}{(1-\alpha\beta_j)^2}$ with $\alpha\beta_j < 1$. For the limit $\beta_j \to 1$, this reduces to equation (5).

C.2 Asymptotics of the first-stage regression

Starting with equation (7), we substitute $\ln p_{jt}^{opt}$ using equation (5) and also use (4) to obtain

$$\ln p_{jt} = \xi_{jt}(\ln p_{jt-1} - \ln \Pi) + (1-\xi_{jt}) \left(\ln p_j^* - t \ln \Pi_j^* + \frac{\alpha}{1-\alpha} \ln(\Pi/\Pi_j^*) + f(x_{jt}) \right), \tag{36}$$

where $f(x_{jt})$ is defined in equation (6).

To derive the OLS estimates of the parameters in equation (8), we rearrange equation (36) to

$$\begin{aligned} \ln p_{jt} + t \ln \Pi_j^* = {} & \xi_{jt}(\ln p_{jt-1} + (t-1) \ln \Pi_j^* - \ln(\Pi/\Pi_j^*)) \\ & + (1-\xi_{jt}) \left(\ln p_j^* + \frac{\alpha}{1-\alpha} \ln(\Pi/\Pi_j^*) + f(x_{jt}) \right). \end{aligned} \tag{37}$$

Computing the unconditional expectation yields

$$\begin{aligned} E[\ln p_{jt} + t \ln \Pi_j^*] = {} & \alpha E[\ln p_{jt-1} + (t-1) \ln \Pi_j^*] - \alpha \ln(\Pi/\Pi_j^*) \\ & + (1-\alpha) \left(\ln p_j^* + \frac{\alpha}{1-\alpha} \ln(\Pi/\Pi_j^*) \right), \end{aligned}$$

using independence of ξ_{jt} and $E[f(x_{jt})] = 0$. Given stationarity of the detrended relative price $\ln p_{jt} + t \ln \Pi_j^*$, the previous equation yields

$$E[\ln p_{jt} + t \ln \Pi_j^*] = \ln p_j^*,$$

or

$$\ln p_{jt} = \ln p_j^* - t \ln \Pi_j^* + u_{jt}, \tag{38}$$

where u_{jt} denotes an expectation error with zero mean. This shows that for regression (8) we get

$$\begin{aligned} \widehat{\ln a_j} &\rightarrow \ln p_j^* \\ \widehat{\ln b_j} &\rightarrow \ln \Pi_j^*, \end{aligned}$$

as the number of price observations becomes large.

C.3 Proof of proposition 2

This appendix derives equations (10) and (11) in the main text. We substitute equation (38) into equation (37), which yields directly yields equation (10). Squaring equation (10), taking unconditional expectations, and using independence of ξ_{jt} yields

$$E[u_{jt}^2] = E[\xi_{jt}^2] E[(u_{jt-1} - \ln(\Pi/\Pi_j^*))^2] + E[(1-\xi_{jt})^2] E\left[\left(f(x_{jt}) + \frac{\alpha}{1-\alpha} \ln(\Pi/\Pi_j^*) \right)^2 \right],$$

where we also used $E[(1-\xi_{jt})\xi_{jt}] = 0$. We can rewrite the previous equation using $E[\xi_t^2] = \alpha$ and $E[(1-\xi_t)^2] = 1-\alpha$, completing the squares to obtain

$$E[u_{jt}^2] = \alpha E[u_{jt-1}^2 + \ln(\Pi/\Pi_j^*)^2 - 2u_{jt-1}\ln(\Pi/\Pi_j^*)] \\ + (1-\alpha)E\big[f(x_{jt})^2 + \left(\frac{\alpha}{1-\alpha}\ln(\Pi/\Pi_j^*)\right)^2 + 2f(x_{jt})\frac{\alpha}{1-\alpha}\ln(\Pi/\Pi_j^*)\big].$$

Recognizing that the expectation of the cross terms in the previous equation are zero because $E[u_{jt}] = 0$ and $E[f(x_{jt})] = 0$ yields

$$E[u_{jt}^2] = \alpha E[u_{jt-1}^2] + \alpha\ln(\Pi/\Pi_j^*)^2 + (1-\alpha)E[f(x_{jt})^2] + (1-\alpha)\left(\frac{\alpha}{1-\alpha}\ln(\Pi/\Pi_j^*)\right)^2.$$

Using $E[u_{jt}^2] = E[u_{jt-1}^2]$ and simplifying terms yields

$$E[u_{jt}^2] = E[f(x_{jt})^2] + \frac{\alpha}{(1-\alpha)^2}(\ln\Pi - \ln\Pi_j^*)^2.$$

Recognizing that $Var[u_{jt}] = E[u_{jt}^2]$, as $E[u_{jt}] = 0$, and $Var[f(x_{jt})] = E[f(x_{jt})^2]$, as $E[f(x_{jt})] = 0$, delivers equation (11).

D Details of the State-Dependent Model

D.1 Setup and OLS regression

Let $z_{jt} = \ln p_{jt} - \ln p_{jt}^*$ be the deviation of the current relative price of product j from the flexible price optimum. Then in between adjustments z_{jt} follows:

$$dz_{jt} = d\ln p_{jt} - d\ln p_{jt}^* = -\underbrace{(\ln\Pi - \ln\Pi_j^*)}_{\mu_j}dt - d\ln x_{jt}$$

$$d\ln x_{jt} = \sum_{i=1}^{N}(\ln x_i - \ln x_{jt})dJ_t^i(\ln x_{jt})$$

where $dJ_t^i(\ln x_{jt})$ is a Poisson jump process with intensity dependent on the current state $\ln x_{jt}$. Since $\ln p_{jt} = \ln p_{jt}^* + z_{jt}$, it follows that:

$$\ln p_{jt} = \ln p_j^* + \ln x_{jt} - t\ln\Pi_j^* + z_{jt} \tag{39}$$

$$E\left[\ln p_{jt} + t\ln\Pi_j^*\right] = \ln p_j^* + \underbrace{E[\ln x_{jt}]}_{=0} + E[z_{jt}]$$

And thus the estimates of OLS regression (8) converge to

$$\widehat{\ln a_j} \to \ln p_j^*$$
$$\widehat{\ln b_j} \to \ln\Pi_j^*,$$

if $E[z_{jt}] = 0$, which is true in the limiting case as $\rho \to 0$, as shown below.[58] Furthermore, residuals and their variance can be written as:

$$\begin{aligned} u_{jt} &= \ln p_{jt} - \ln p_j^* + t \ln \Pi_j^* = z_{jt} + \ln x_{jt} \\ Var(u_{jt}) &= E[z_{jt}^2] + 2E[z_{jt} \ln x_{jt}] + Var(\ln x_{jt}) \end{aligned} \tag{40}$$

D.2 Solution

To simplify notation, we omit the product index j. The firm's objective is to maximize its value from equation (14), given by:

$$V(z, x_i) = \max_{\{\tau_i, \Delta z_{\tau_i}\}_{i=1}^{\infty}} -E\left[\int_0^{\infty} e^{-\rho t} z_t^2 dt + \kappa \sum_{i=1}^{\infty} e^{-\rho \tau_i} \,\middle|\, z_0 = z, x_0 = x_i\right]$$

The firm's policy consists of a collection of inaction region boundaries $\{\underline{z}(x_i), \overline{z}(x_i)\}$ and reset price gaps $\hat{z}(x_i)$, for all $i \in N$. The HJB equation for the inaction region is given by:

$$\begin{aligned} \rho V(z, x_i) = &-z^2 - \mu \partial_z V(z, x_i) \\ &+ \sum_{j \neq i}^{N} \lambda_{ij}^X \big(V(z - (\ln x_j - \ln x_i), x_j) - V(z, x_i)\big) \end{aligned}$$

The optimal policy satisfies the usual smooth pasting and optimality conditions: $\partial_z V(\hat{z}(x_i), x_i) = \partial_z V(\underline{z}(x_i), x_i) = \partial_z V(\overline{z}(x_i), x_i) = 0$ and $V(\underline{z}(x_i), x_i) = V(\overline{z}(x_i), x_i) = V(\hat{z}(x_i), x_i) - \kappa$. Define $v(z, x_i) = V(z, x_i) - V(\hat{z}(x_1), x_1)$. Then:

$$\begin{aligned} \rho v(z, x_i) = &-z^2 - \mu \partial_z v(z, x_i) \\ &+ \sum_{j \neq i}^{N} \lambda_{ij}^X \big(v(z - (\ln x_j - \ln x_i), x_j) - v(z, x_i)\big) - \rho V(\hat{z}(x_1), x_1) \end{aligned}$$

with $\partial_z v(\hat{z}(x_i), x_i) = \partial_z v(\underline{z}(x_i), x_i) = \partial_z v(\overline{z}(x_i), x_i) = 0$ and $v(\underline{z}(x_i), x_i) = v(\overline{z}(x_i), x_i) = v(\hat{z}(x_i), x_i) - \kappa$. We now take the limit as $\rho \to 0$.

Proposition 4 *As $\rho \to 0$, the scaled value function $\rho V(z, x)$ at any state $\{z, x\}$ converges to a constant: $\lim_{\rho \to 0} \rho V(z, x) = A \in \mathbb{R}\ \forall z, x$.*

[58] While this result is shown formally under the assumption of sufficiently small κ, it holds more generally. As $\rho \to 0$, the firms' value until adjustment becomes the negative expected squared deviation of price gaps from zero, maximizing which requires setting the expected price gap to zero.

All proofs are provided in section D.3. By Proposition 4, $\lim_{\rho \to 0} \rho v(z, x_i) = 0$ and $\lim_{\rho \to 0} \rho V(\hat{z}(x_1), x_1) = A$, so that:

$$\lambda_i^X v(z, x_i) = -z^2 - \mu \partial_z v(z, x_i) + \sum_{j \neq i}^{N} \lambda_{ij}^X v(z - (\ln x_j - \ln x_i), x_j) - A$$

where $\lambda_i^X = \sum_{j \neq i}^N \lambda_{ij}^X = -\lambda_{ii}^X$ is the intensity with which $\ln x_t$ is exiting state i. Evaluate the above expression at $z = \hat{z}(x_1), x_i = x_1$ to obtain:

$$A = -\left(\hat{z}(x_1)\right)^2 + \sum_{j \neq 1}^{N} \lambda_{1j}^X v(\hat{z}(x_1) - (\ln x_j - \ln x_1), x_j)$$

Lemma 5 *There exists $\overline{\kappa} > 0$ such that firms find it optimal to adjust after every change in x for all $\kappa < \overline{\kappa}$.*

Suppose that κ is small enough in the sense of Lemma 5. Then firms find it optimal to adjust whenever idiosyncratic state x changes its value. The HJB equation becomes:

$$\lambda_i^X v(z, x_i) = -z^2 - \mu \partial_z v(z, x_i) + \sum_{j \neq i}^{N} \lambda_{ij}^X v(\hat{z}(x_j), x_j) - \lambda_i^X \kappa - A$$

with

$$A = -\left(\hat{z}(x_1)\right)^2 + \sum_{j \neq i}^{N} \lambda_{1j}^X v(\hat{z}(x_j), x_j)$$

and value function satisfies:

$$v(z, x_i) = C_i^v e^{-\alpha_i z} - \frac{z^2}{\lambda_i^X} + \frac{2z}{\alpha_i \lambda_i^X} - \frac{2}{\alpha_i^2 \lambda_i^X} + \frac{C_i}{\lambda_i^X}$$

$$C_i = \sum_{j \neq i}^{N} \lambda_{ij}^X v(\hat{z}(x_j), x_j) - \lambda_i^X \kappa - A$$

$$\partial_z v(\hat{z}(x_i), x_i) = \partial_z v(\underline{z}(x_i), x_i) = \partial_z v(\overline{z}(x_i), x_i) = 0$$

$$v(\hat{z}(x_i), x_i) - \kappa = v(\underline{z}(x_i), x_i) = v(\overline{z}(x_i), x_i)$$

with $\alpha_i = \frac{\lambda_i^X}{\mu}$. As long as state x remains unchanged, price gaps evolve deterministically with drift $-\mu$. It thus suffices to solve for the reset price gap and only one boundary of the inaction region. From now on, we consider $\mu > 0$ and solve for $\hat{z}(x_i)$ and $\underline{z}(x_i)$ since the upper boundary of the inaction region is irrelevant. Because of symmetry properties of the model, it is straightforward to then recover the solution and all statistics for $\mu < 0$. To ease notation, let $\hat{z}(x_i) = \hat{z}_i$ and $\underline{z}(x_i) = \underline{z}_i$.

Lemma 6 *Suppose $\mu > 0$. Then for each state x_i, optimal policy is determined by the following two conditions:*

$$\underline{z}_i^2 - \hat{z}_i^2 = \lambda_i^X \kappa \tag{41}$$

$$e^{\alpha_i \hat{z}_i}(1 - \alpha_i \hat{z}_i) = e^{\alpha_i \underline{z}_i}(1 - \alpha_i \underline{z}_i) \tag{42}$$

where $\alpha_i = \frac{\lambda_i^X}{\mu}$.

Conditional on state x_i, the price gap distribution satisfies:

$$\lambda_i^X f_i(z) = \mu \partial_z f_i(z)$$

$$\int_{\underline{z}_i}^{\hat{z}_i} f_i(z) dz = 1$$

and is thus given by:

$$f_i(z) = \frac{\alpha_i e^{\alpha_i z}}{e^{\alpha_i \hat{z}_i} - e^{\alpha_i \underline{z}_i}}$$

It follows that:

$$E[z|x_i] = \int_{\underline{z}_i}^{\hat{z}_i} z f_i(z) dz = 0$$

$$E[z] = 0$$

$$E[z^2|x_i] = \int_{\underline{z}_i}^{\hat{z}_i} z^2 f_i(z) dz = \frac{\hat{z}_i + \underline{z}_i}{\alpha_i} - \hat{z}_i \underline{z}_i \tag{43}$$

$$E[z^2] = E_x \left[\frac{\hat{z}_i + \underline{z}_i}{\alpha_i} - \hat{z}_i \underline{z}_i \right] \tag{44}$$

where $E_x[\cdot]$ is the expectation with respect to stationary distribution of x.

Proposition 7 *For μ close to zero, $E[z^2] = E\left[\frac{1}{\left(\lambda_i^X\right)^2}\right] \mu^2 + O(4)$.*

Finally, note that $E[zx] = E\big[x_i E[z|x_i]\big] = 0$ and the main object of interest – the variance of residuals from the OLS regression (8) – is given by:

$$Var(u_{jt}) = Var(\ln x_{jt}) + E\left[\frac{1}{(\lambda_i^X)^2}\right]\mu_j^2 + O(4)$$
$$= Var(\ln x_{jt}) + E\left[\frac{1}{(\lambda_i^X)^2}\right](\ln\Pi - \ln\Pi_j^*)^2 + O(4)$$

D.3 Proofs

Proof of Proposition 4. The proof here extends Lemma 3 in Online Appendix of Alvarez et al. (2019) to a setting with two state variables. Let $V(z,x,\rho)$ be the value function in state $\{z,x\}$ under discount rate ρ. We can write $\rho V(z,x,\rho)$ as follows:

$$\rho V(z,x,\rho) = -E\left[\rho\int_0^{\tau_N} e^{-\rho t} z_t^2 dt\right] - \kappa E\left[\rho\sum_{i=1}^{N} e^{-\rho\tau_i}\right]$$
$$\underbrace{-\rho E\left[\int_0^{\infty} e^{-\rho(\tau_N+t)} z_{\tau_N+t}^2 dt + \kappa\sum_{i=1}^{\infty} e^{-\rho\tau_{N+i}}\right]}_{\rho E\left[e^{-\rho\tau_N}V(z_{\tau_N},x_{\tau_N},\rho)\right]}$$

where τ_N is the N-th adjustment and all expectation operators are conditional on $z_0 = z, x_0 = x$. Subtract $\rho E\left[e^{-\rho\tau_N}V(z,x,\rho)\right]$ from both sides and divide by $\left(1 - E\left[e^{-\rho\tau_N}\right]\right)$ to obtain:

$$\rho V(z,x,\rho) = -\frac{\rho}{1-E\left[e^{-\rho\tau_N}\right]}E\left[\int_0^{\tau_N} e^{-\rho t} z_t^2 dt\right] - \frac{\rho\kappa}{1-E\left[e^{-\rho\tau_N}\right]}E\left[\sum_{i=1}^{N} e^{-\rho\tau_i}\right]$$
$$\frac{\rho}{1-E\left[e^{-\rho\tau_N}\right]}E\left[e^{-\rho\tau_N}\left(V(z_{\tau_N},x_{\tau_N},\rho) - V(z,x,\rho)\right)\right]$$

Take the limit as $\rho \to 0$. Note that $\frac{\rho}{1-E\left[e^{-\rho\tau_N}\right]} \to \frac{1}{E[\tau_N]}$ and thus:

$$\lim_{\rho\to 0}\rho V(z,x,\rho) = -\frac{1}{E\left[\tau_N\right]}E\left[\int_0^{\tau_N} z_t^2 dt\right] - \frac{\kappa N}{E\left[\tau_N\right]}$$
$$\frac{1}{E\left[\tau_N\right]}\lim_{\rho\to 0}E\left[V(z_{\tau_N},x_{\tau_N},\rho) - V(z,x,\rho)\right]$$

By Lemma 8, $|V(z_{\tau_N},x_{\tau_N},\rho) - V(z,x,\rho)| \le C \in \mathbb{R}$ for all $\rho > 0$ and thus this also holds in the limit as $\rho \to 0$. As we take the limit with $N \to \infty$, the first term converges to the unconditional expected squared gap $E[z^2]$, the second term converges to adjustment frequency λ_a times adjustment cost κ, and the third term vanishes as $E\left[\tau_N\right] \to \infty$. Thus $\lim_{\rho\to 0}\rho V(z,x,\rho) = -E[z^2] - \kappa\lambda_a \equiv A$ for all z,x. ■

Lemma 8 *There exists $C \in \mathbb{R}$ such that for any $\rho > 0$ and any z, x, z', x', $|V(z,x) - V(z',x')| \leq C$.*

Proof. First, we show that $\rho V(z, x_i)$ is bounded from below. To see that, recall that $V(z, x_i)$ is achieved under the optimal adjustment policy, meaning that the value of any feasible policy is weakly lower. Consider the following policy: the firm adjusts its price gap whenever it is hit by a Poisson x shock. In addition, it also adjusts at random times with Poisson intensity λ_i, which is specific to each state x_i. These intensities satisfy the following condition: $\lambda_i^X + \lambda_i = \max_i \lambda_i^X \equiv \lambda$, such that in every state x_i firms adjust with equal intensity λ. Since adjustments occur exogenously, firms only choose the reset price gap $\hat{z}_i$ to maximize expected profits until the next adjustment:

$$\max_{\hat{z}_i} E\left[-\int_0^\tau e^{-\rho t} z_t^2 \,\middle|\, z_0 = \hat{z}_i\right] = \max_{\hat{z}_i} E\left[-\int_0^\infty e^{-(\rho+\lambda)t} z_t^2 \,\middle|\, z_0 = \hat{z}_i\right]$$

Because in between adjustments price gaps drift deterministically ($z_t = \hat{z}_i - \mu t$) and adjustment intensities are equalized across states, optimal reset price gap does not depend on x and satisfies FOC:

$$\int_0^\infty e^{-(\rho+\lambda)t}(\hat{z} - \mu t) = 0 \Longrightarrow \hat{z} = \frac{\mu}{\rho + \lambda}$$

Denote by $\tilde{V}(z, x)$ the value function under this policy. Since $\partial_z \tilde{V}(\hat{z}, x_i) = 0$, evaluating the HJB equation at $\hat{z}$ yields:

$$\begin{aligned}
\rho \tilde{V}(\hat{z}, x_i) &= -\hat{z}^2 + \lambda_i\left(\tilde{V}(\hat{z}, x_i) - \kappa - \tilde{V}(\hat{z}, x_i)\right) + \sum_{j \neq i}^N \lambda_{ij}^X \left(\tilde{V}(\hat{z}, x_j) - \kappa - \tilde{V}(\hat{z}, x_i)\right) \\
&= -\hat{z}^2 + \sum_{j \neq i}^N \lambda_{ij}^X \left(\tilde{V}(\hat{z}, x_j) - \tilde{V}(\hat{z}, x_i)\right) - \kappa \underbrace{\left(\lambda_i + \sum_{j \neq i}^N \lambda_{ij}^X\right)}_{=\lambda}
\end{aligned}$$

It is straightforward to show that $\tilde{V}(\hat{z}, x_i) = \tilde{V}(\hat{z}, x_j)$ for all i and j. Assume the opposite and let $\overline{v} = \max_i \tilde{V}(\hat{z}, x_i)$ and $\underline{v} = \min_i \tilde{V}(\hat{z}, x_i)$. Then:

$$\begin{aligned}
\rho \overline{v} &= -\hat{z}^2 + \sum_{j \neq i(\overline{v})}^N \lambda_{i(\overline{v})j}^X \underbrace{\left(\tilde{V}(\hat{z}, x_j) - \overline{v}\right)}_{\leq 0} - \lambda\kappa \\
&\leq -\hat{z}^2 + \sum_{j \neq i(\underline{v})}^N \lambda_{i(\underline{v})j}^X \underbrace{\left(\tilde{V}(\hat{z}, x_j) - \underline{v}\right)}_{\geq 0} - \lambda\kappa = \rho \underline{v}
\end{aligned}$$

Meaning $\underline{v} = \overline{v}$. As a result, $\rho\tilde{V}(\hat{z}, x_i) = -\hat{z}^2 - \lambda\kappa = -\frac{\mu^2}{(\rho+\lambda)^2} - \lambda\kappa \geq -\frac{\mu^2}{\lambda^2} - \lambda\kappa$ for any $\rho > 0$. Thus for the true value function evaluated at the true optimal reset price gap $\hat{z}(x_i)$ it holds that $\rho V(\hat{z}(x_i), x_i) \geq \rho\tilde{V}(\hat{z}, x_i) \geq -\frac{\mu^2}{\lambda^2} - \lambda\kappa$ for all $\rho > 0$.

Consider now the true value function $V(z, x_i)$ and pick i such that $V(\hat{z}(x_i), x_i) = \max_j V(\hat{z}(x_j), x_j)$. The HJB equation for this value function satisfies:

$$
\begin{aligned}
-\frac{\mu^2}{\lambda^2} - \lambda\kappa \leq \rho V(\hat{z}(x_i), x_i) &= \underbrace{-\left(\hat{z}(x_i)\right)^2}_{\leq 0} - \mu \underbrace{\partial_z V(\hat{z}(x_i), x_i)}_{=0} \\
&\quad + \sum_{j\neq i}^{N} \lambda_{ij}^X \left(\underbrace{V(\hat{z}(x_i) - (\ln x_j - \ln x_i), x_j)}_{\leq V(\hat{z}(x_j), x_j)} - V(\hat{z}(x_i), x_i) \right) \\
&\leq \sum_{j\neq i}^{N} \lambda_{ij}^X \underbrace{\left(V(\hat{z}(x_j), x_j) - V(\hat{z}(x_i), x_i)\right)}_{\leq 0} \leq 0
\end{aligned}
$$

It follows that whenever $\lambda_{ij}^X > 0$:

$$
\left(-\frac{\mu^2}{\lambda^2} - \lambda\kappa\right) / \lambda_{ij}^X \leq V(\hat{z}(x_j), x_j) - V(\hat{z}(x_i), x_i) \leq 0
$$

For the states j where $\lambda_{ij}^X = 0$ we can bound the difference $V(\hat{z}(x_j), x_j) - V(\hat{z}(x_i), x_i)$ iteratively because the network of x_i is connected (every two states are connected by some path). In addition, for any z, x_i:

$$
V(\hat{z}(x_i), x_i) - \kappa \leq V(z, x_i) \leq V(\hat{z}(x_i), x_i)
$$

Therefore there exists $C \in \mathbb{R}$ such that for all $\rho > 0$, $|V(z, x) - V(z', x')| \leq C$ for all z, x, z', x'. ■

Proof of Lemma 5. Consider a model M in which firms are forced to adjust after every change in x, but can also adjust at other times and choose the boundaries of inaction regions and reset price gaps. Suppose we now allow the firms to adjust whenever they find it to be optimal. They will adjust their policies $\{\underline{z}(x_i), \hat{z}(x_i), \overline{z}(x_i)\}_{i=1}^N$ only if changes in x keep price gaps within the bounds of inaction regions. Otherwise the optimal policy in model M and in the model of interest coincide, meaning that firms find it optimal to adjust after every change in x. To see that, compare the HJB equations in the original model (first line) and model

M (second line):

$$\begin{aligned}\lambda_i^X v(z,x_i) &= -z^2 - \mu\partial_z v(z,x_i)\\ &\quad + \sum_{j\neq i}^{N} \lambda_{ij}^X v(z-(\ln x_j - \ln x_i), x_j) - A\\ \lambda_i^X v(z,x_i) &= -z^2 - \mu\partial_z v(z,x_i)\\ &\quad + \sum_{j\neq i}^{N} \lambda_{ij}^X (v(\hat{z}(x_j), x_j) - \kappa) - A\end{aligned}$$

If upon the change in x, $z-(\ln x_j - \ln x_i) \notin [\underline{z}(x_j), \overline{z}(x_j)]$, then $v(z-(\ln x_j - \ln x_i), x_j) = v(\hat{z}(x_j), x_j) - \kappa$ and the value functions in the two models coincide. Therefore, $\overline{\kappa}$ is such that $\min_{ij} |\ln x_i - \ln x_j| = \max_i \overline{z}(x_i) - \min_i \underline{z}(x_i)$ in model M. Such $\overline{\kappa} > 0$ always exists since for all i $\lim_{\kappa\to 0} \overline{z}(x_i) = \lim_{\kappa\to 0} \underline{z}(x_i) = 0$. ■

Proof of Lemma 6. From $\partial_z v(\hat{z}_i, x_i) = 0$ and $\partial_z v(\underline{z}_i, x_i) = 0$ it follows:

$$\begin{aligned}-\alpha_i C_i^v e^{-\alpha_i \hat{z}_i} - \frac{2\hat{z}_i}{\lambda_i^X} + \frac{2}{\alpha_i\lambda_i^X} &= 0 = -\alpha_i C_i^v e^{-\alpha_i \underline{z}_i} - \frac{2\underline{z}_i}{\lambda_i^X} + \frac{2}{\alpha_i\lambda_i^X}\\ -\alpha_i C_i^v - \frac{2\hat{z}_i e^{\alpha_i\hat{z}_i}}{\lambda_i^X} + \frac{2e^{\alpha_i\hat{z}_i}}{\alpha_i\lambda_i^X} &= 0 = -\alpha_i C_i^v - \frac{2\underline{z}_i e^{\alpha_i\underline{z}_i}}{\lambda_i^X} + \frac{2e^{\alpha_i\underline{z}_i}}{\alpha_i\lambda_i^X}\\ e^{\alpha_i\hat{z}_i}(1-\alpha_i\hat{z}_i) &= e^{\alpha_i\underline{z}_i}(1-\alpha_i\underline{z}_i)\end{aligned}$$

Similarly:

$$\begin{aligned}-\alpha_i C_i^v e^{-\alpha_i \hat{z}_i} - \frac{2\hat{z}_i}{\lambda_i^X} + \frac{2}{\alpha_i\lambda_i^X} &= -\alpha_i C_i^v e^{-\alpha_i \underline{z}_i} - \frac{2\underline{z}_i}{\lambda_i^X} + \frac{2}{\alpha_i\lambda_i^X}\\ C_i^v e^{-\alpha_i\hat{z}_i} + \frac{2\hat{z}_i}{\alpha_i\lambda_i^X} &= C_i^v e^{-\alpha_i\underline{z}_i} + \frac{2\underline{z}_i}{\alpha_i\lambda_i^X}\\ C_i^v e^{\alpha_i(\underline{z}_i - \hat{z}_i)} &= C_i^v + e^{\alpha_i\underline{z}_i}\frac{2(\underline{z}_i - \hat{z}_i)}{\alpha_i\lambda_i^X} \qquad (45)\end{aligned}$$

From $v(\hat{z}_i, x_i) - \kappa = v(\underline{z}_i, x_i)$ it follows:

$$\begin{aligned}C_i^v e^{-\alpha_i\hat{z}_i} - \frac{\hat{z}_i^2}{\lambda_i^X} + \frac{2\hat{z}_i}{\alpha_i\lambda_i^X} - \kappa &= C_i^v e^{-\alpha_i\underline{z}_i} - \frac{\underline{z}_i^2}{\lambda_i^X} + \frac{2\underline{z}_i}{\alpha_i\lambda_i^X}\\ C_i^v e^{\alpha_i(\underline{z}_i - \hat{z}_i)} + e^{\alpha_i\underline{z}_i}\left[\frac{2(\hat{z}_i - \underline{z}_i)}{\alpha_i\lambda_i^X} + \frac{\underline{z}_i^2 - \hat{z}_i^2}{\lambda_i^X} - \kappa\right] &= C_i^v\\ \underline{z}_i^2 - \hat{z}_i^2 &= \lambda_i^X\kappa\end{aligned}$$

where the last line follows from (45). ■

Lemma 9 *For every state x_i, $\hat{z}_i \frac{\partial \hat{z}_i}{\partial \mu} = \underline{z}_i \frac{\partial \underline{z}_i}{\partial \mu} = \frac{E[z^2|x_i]}{\mu}$.*

Proof. The first equality follows directly from the first order derivative of equilibrium condition (41) with respect to μ. For the second equality, differentiate equilibrium condition (42) and collect terms:

$$
\begin{aligned}
e^{\alpha_i \hat{z}_i} \left[\frac{\partial \hat{z}_i}{\partial \mu} - \frac{\hat{z}_i}{\mu} \right] \hat{z}_i &= e^{\alpha_i \underline{z}_i} \left[\frac{\partial \underline{z}_i}{\partial \mu} - \frac{\underline{z}_i}{\mu} \right] \underline{z}_i \\
(1 - \alpha_i \underline{z}_i) \left[\frac{\partial \hat{z}_i}{\partial \mu} - \frac{\hat{z}_i}{\mu} \right] \hat{z}_i &= (1 - \alpha_i \hat{z}_i) \left[\frac{\partial \underline{z}_i}{\partial \mu} - \frac{\underline{z}_i}{\mu} \right] \underline{z}_i \\
\underline{z}_i \frac{\partial \underline{z}_i}{\partial \mu} (\alpha_i \hat{z}_i - \alpha_i \underline{z}_i) &= \frac{\hat{z}_i^2 (1 - \alpha_i \underline{z}_i) - \underline{z}_i^2 (1 - \alpha_i \hat{z}_i)}{\mu} \\
\underline{z}_i \frac{\partial \underline{z}_i}{\partial \mu} &= \frac{1}{\mu} \frac{\hat{z}_i^2 - \underline{z}_i^2 - \alpha_i \hat{z}_i \underline{z}_i (\hat{z}_i - \underline{z}_i)}{\alpha_i (\hat{z}_i - \underline{z}_i)} \\
&= \frac{1}{\mu} \left[\frac{\hat{z}_i + \underline{z}_i}{\alpha_i} - \hat{z}_i \underline{z}_i \right] = \frac{E[z^2|x_i]}{\mu}
\end{aligned}
$$

where the second line uses (42) and the third line uses $\hat{z}_i \frac{\partial \hat{z}_i}{\partial \mu} = \underline{z}_i \frac{\partial \underline{z}_i}{\partial \mu}$. ■

Lemma 10 *As $\mu \to 0$, $\hat{z}_i \to 0$, $\underline{z}_i \to -\sqrt{\lambda_i^X \kappa}$ and $E[z^2] \to 0$.*

Proof. Combine equilibrium conditions (41) and (42) to obtain:

$$
\underbrace{\left(\mu + \lambda_i^X \sqrt{\lambda_i^X \kappa + \hat{z}_i^2} \right)}_{>0} = \left(\mu - \lambda_i^X \hat{z}_i \right) \underbrace{e^{\frac{\lambda_i^X}{\mu} \left(\hat{z}_i + \sqrt{\lambda_i^X \kappa + \hat{z}_i^2} \right)}}_{>0}
$$

Since the LHS is always positive, and so is the exponent on the RHS, $\lim_{\mu \to 0} \hat{z}_i = 0$. It then follows from (41) that $\lim_{\mu \to 0} \underline{z}_i = -\sqrt{\lambda_i^X \kappa}$ and from (44) that $\lim_{\mu \to 0} E[z^2] = 0$. ■

Proof of Proposition 7. From Lemmas 9 and 10, and equation (43) it follows that:

$$
\begin{aligned}
\underline{z}_i' \equiv \frac{\partial \underline{z}_i}{\partial \mu} &= \frac{1}{\lambda_i^X} + \frac{\mu \hat{z}_i - \lambda_i^X \hat{z}_i \underline{z}_i}{\mu \lambda_i^X \underline{z}_i} \\
\lim_{\mu \to 0} \underline{z}_i' &= \frac{1}{\lambda_i^X} - \lim_{\mu \to 0} \frac{\hat{z}_i}{\mu} = \frac{1}{\lambda_i^X} - \lim_{\mu \to 0} \hat{z}_i'
\end{aligned}
$$

At the same time, by Lemma 9: $\hat{z}_i = \frac{\underline{z}_i \underline{z}_i'}{\hat{z}_i'}$, and by Lemma 10: $\lim_{\mu \to 0} \frac{\underline{z}_i'}{\hat{z}_i'} = 0$. It then follows that:

$$
O(1) = \frac{\underline{z}_i'}{\hat{z}_i'} = \frac{\frac{1}{\lambda_i^X} - \hat{z}_i' + O(1)}{\hat{z}_i'} = \frac{1 + O(1)}{\lambda_i^X \hat{z}_i'} - 1
$$

And therefore $\lim_{\mu\to 0} \hat{z}_i' = \frac{1}{\lambda_i^X}$. From (41) it follows that $\lim_{\mu\to 0} \underline{z}_i' = 0$ and from (44) that $\lim_{\mu\to 0} \frac{\partial E[z^2]}{\partial \mu} = 0$. If $\hat{z}_i$ is twice differentiable at $\mu = 0$, then due to anti-symmetry ($\hat{z}_i(\mu) = -\hat{z}_i(-\mu)$), $\hat{z}_i''(0) = 0$. It follows that $\hat{z}_i' = \frac{1}{\lambda_i^X} + O(2)$ and $\hat{z}_i = \frac{\mu}{\lambda_i^X} + O(3)$. Using Lemma 9 we obtain that:

$$E[z^2] = E\left[\frac{1}{(\lambda_i^X)^2}\right]\mu^2 + O(4)$$

■

Lemma 11 *Suppose $\lambda_i^X = \Lambda$ for all i. Then, as $\mu \to 0$, adjustment frequency $\Lambda_a = \Lambda + O(4)$.*

Proof. Since $\lambda_i^X = \Lambda$, we can omit the i index. The expected stopping time $\tau(z)$ solves the following ODE: $\Lambda\tau(z) = 1 - \mu\partial_z\tau(z)$, together with boundary condition $\tau(\underline{z}) = 0$, and is given by $\tau(z) = \frac{1}{\Lambda}\left(1 - e^{\alpha(\underline{z}-\hat{z})}\right)$. It follows from Lemma 6 and equation (44) that:

$$\Lambda_a \equiv \frac{1}{\tau(\hat{z})} = \frac{1}{\kappa}\left(\underline{z}^2 - E[z^2]\right)$$

Lemma 10 implies that as $\mu \to 0$, $\Lambda_a \to \Lambda$. Furthermore:

$$\begin{aligned}\frac{\partial \Lambda_a}{\partial \mu} &= \frac{1}{\kappa}\left(2\underline{z}\frac{\partial \underline{z}}{\partial \mu} - \frac{\partial E[z^2]}{\partial \mu}\right) \\ &= \frac{1}{\kappa}\left(2\frac{E[z^2]}{\mu} - 2\frac{\mu}{\Lambda^2} + O(3)\right) = O(3)\end{aligned}$$

where the last line follows from Lemma 9 and Proposition 7. Therefore, $\Lambda_a = \Lambda + O(4)$. ■

E Details of the Regression Approach

This section discusses econometric details associated with estimating our key equation (11), which relates inefficient price dispersion to suboptimal inflation at the product level. In our baseline empirical approach, we estimate equation (11) at the level of finely disaggregated expenditure items, exploiting variation across products within the item. Our sample contains more than 1000 expenditure items, so that obtain a large number of estimates of the coefficient of interest c in equation (11).

We use a two-step estimation approach, because neither the left-hand side variable nor the right hand-side variables in equation (11) can be directly observed. This section presents this approach and discusses how first-stage estimation errors affect second-stage regression outcomes. In particular, it shows that first-stage error biases the estimates of the coefficient c towards zero, i.e., towards finding no marginal effect of suboptimal inflation on inefficient price dispersion.

Our first-stage estimation consists of a seemingly unrelated regression (SUR) system that contains two equations. The left-hand side variable in equation (11) can be estimated using the residuals of relative-price regressions of the form

$$\ln p_{jzt} = \ln a_{jz} - (\ln b_{jz}) \cdot t + u_{jzt} \tag{46}$$

where j denotes the product and $z \in \{1, ...Z\}$ the expenditure item under consideration, with Z being the total number of expenditure items in our sample.[59]

Estimation of the right-hand side variables in equation (11) would require estimating the average inflation rate, $\ln \Pi_z$, and the product specific optimal inflation rate, $\ln \Pi^*_{jz}$. Having two first-stage estimates on the right-hand side of equation (11) is, however, unattractive on econometric grounds.[60] A more parsimonious way to proceed is to estimate instead directly the gap between the item-level and product-specific optimal inflation rate ($\ln \Pi_z/\Pi^*_{jz}$) in the first stage. This can be achieved by adding the price level equation

$$\ln P_t = \ln P_0 + \ln \Pi \cdot t$$

to equation (4). Adding the item-level subindex z, we obtain for every product *another* first-stage regression of the form

$$\ln P_{jzt} = \ln \widetilde{a}_{jz} + \left(\ln \Pi_z/\Pi^*_{jz}\right) \cdot t + \widetilde{u}_{jzt} \tag{47}$$

where P_{jzt} denotes the *nominal* product price. Equation (47) shows that the time trend in the nominal price of the product directly identifies the gap between item-level inflation and the product-specific optimal inflation rate. Equations (46) and (47) jointly make up our first-stage seemingly unrelated regression (SUR) system.

Since the SUR system (46)-(47 does not feature exclusion restrictions, OLS estimation is identical to GLS estimation, despite the presence of correlated residuals. OLS estimation delivers an unbiased estimate of the gap $\ln \Pi_z/\Pi^*_{jz}$ and an unbiased estimate of the residual variance of interest,

$$\widehat{Var}(u_{jzt}) = \frac{1}{T_{jz} - 2} \sum_t \left(\widehat{u}_{jzt}\right)^2,$$

[59] To simplify notation, the previous sections have suppressed the item index z.

[60] It requires discussing, amongst other things, the covariance in the estimation errors of these two right-hand side variables.

where T_{jz} denotes the number of price observations for product j in item z.

The first-stage estimates for each product j within expenditure item z can then be used to estimate the second-stage equation

$$\widehat{Var}(u_{jzt}) = v_z + c_z \cdot (\ln \widehat{\Pi_z/\Pi_{jz}^*})^2 + \varepsilon_{jz} \tag{48}$$

using OLS estimation. This delivers an estimate of c_z for each each expenditure item $z = 1, \ldots, Z$. The error term ε_{jz} in equation (48) absorbs measurement error of the left-hand side variable, as discussed below, as well as the higher-order approximation errors implied by menu-cost models, see equation (15).

While the first-stage estimates $\widehat{Var}(u_{jzt})$ and $\ln \widehat{\Pi_z/\Pi_{jz}^*}$ are unbiased, they are contaminated by sampling error. Sampling error is an important concern because the product price time series underlying the first-stage system can be relatively short. Fortunately, the effect of the first-stage sampling error consists solely of biasing the estimate of c_z towards zero, as we now show next.

To illustrate this point, we assume that the first-stage residuals are normally distributed. (The more general case with non-normal errors is discussed in appendix E.1 below.) When estimating the SUR system (46)-(47), the estimation error in $\ln \widehat{\Pi_z/\Pi_{jz}^*}$ is orthogonal to the estimation error in the residuals $\{\widehat{u}_{jzt}\}$, by construction of the OLS estimate. With normality, both estimation errors are also independent of each other. Therefore, the estimation error in $\widehat{Var}(u_{jzt})$ on the l.h.s. of equation (48) is independent of the estimation error in $(\ln \widehat{\Pi_z/\Pi_{jz}^*})^2$ on the r.h.s. of the equation, because both variables are nonlinear transformations of independent random variables.

First-stage estimation error on the l.h.s. of equation (48) thus takes the form of classical measurement error: it does not generate any bias in the second-stage estimates of c_z, instead gets absorbed by the regression residual ε_{jz}. However, first-stage estimation error in $(\ln \widehat{\Pi_z/\Pi_{jz}^*})^2$ biases the second-stage estimate of c_z towards zero. This is so because measurement error in $(\ln \widehat{\Pi_z/\Pi_{jz}^*})^2$ generates a classic attenuation effect. In addition, estimation error in $\ln \widehat{\Pi_z/\Pi_{jz}^*}$ raises the expected value of $(\ln \widehat{\Pi_z/\Pi_{jz}^*})^2$, which generates a further bias towards zero.

Our second-stage estimates for c_z thus provides a *lower bound* of the true marginal effect of suboptimal inflation on price distortions. Since we are interested in rejecting the null hypothesis of inflation *not* creating inefficient price dispersion, $H_0 : c_z = 0$, the bias is working against our main finding.

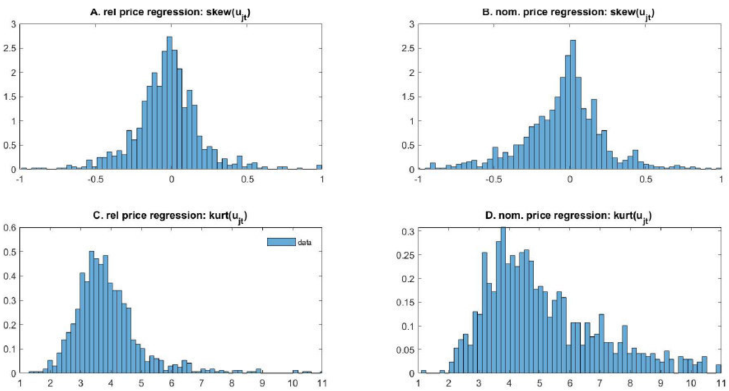

Figure 18: Skewness and kurtosis of the first-stage regression residuals

Finally, to insure that our results are not driven by outliers, e.g., ones associated with errors in price collection, we eliminate within each expenditure item all products falling into the top 5% of the distribution of residual variances $\widehat{Var}(u_{jzt})$ and the top 5% of estimated inflation gaps $(\ln \widehat{\Pi_z/\Pi^*_{jz}})^2$ when running our second-stage regression.

E.1 General Case with Non-Normal First-Stage Residuals

Figure 18 reports the skewness and kurtosis of the first-stage regression residuals of equation (46) (left-hand side panels) and equation (47) (right-hand side panels) across the considered expenditure items.[61] The top panels show that skewness is centered around zero and relatively tightly so, in line with the zero skewness of the normal distribution. For kurtosis, shown in the lower panels of figure 18, the situation looks different. Kurtosis values often lie above the value of 3 implied by a normal distribution.

We now show that quite similar arguments as for the normal case apply to our second-stage estimates of c_z when first-stage residuals fail to be normal. In fact, to insure that there is at most a downward bias in the second-stage estimate of c_z, it is sufficient to insure that the estimation error in the l.h.s. variable $\widehat{Var}(u_{jzt})$ in equation (48) is orthogonal to (rather than independent of) the estimation error in the r.h.s. regressor $(\ln \widehat{\Pi_z/\Pi^*_{jz}})^2$.

Recall that the errors in $(\ln \widehat{\Pi_z/\Pi^*_{jz}})$ and $\{\widehat{u}_{jzt}\}$ are orthogonal by

[61] The measures use outlier trimmed residuals by considering the 2.5%-97.5% quantile of the residual distribution.

construction. A violation of orthogonality between $(\ln \widehat{\Pi_z/\Pi^*_{jz}})^2$ and $\widehat{Var}(u_{jzt})$ can thus only arise because these variables are nonlinear rather than linear functions of $\ln \widehat{\Pi_z/\Pi^*_{jz}}$ and $\{\widehat{u}_{jzt}\}$, respectively. This illustrates that violations of orthogonality conditions are somewhat unlikely to emerge on a priori grounds, even in the absence of normality.

We show below that orthogonality of the estimation errors in $(\ln \widehat{\Pi_z/\Pi^*_{jz}})^2$ and $\widehat{Var}(u_{jzt})$ holds whenever the residuals satisfy

$$Cov[\left((0,1)\left(X'X\right)^{-1}X'u(0,1)'\right)^2, (1,0)'u'Mu(1,0)|X] = 0, \tag{49}$$

where

$$X' \equiv \begin{pmatrix} 1\,1\,1\ldots \\ 0\,1\,2\ldots \end{pmatrix} \tag{50}$$

is the matrix of first-stage regressors and M the matrix defined in (51) below. Condition (49) is a condition on the true residuals u, which is satisfied in the special case with normal errors. Condition (49) holds by construction when replacing the true residuals u by the estimated OLS or GLS residuals $\widehat{u}$, thus cannot be tested empirically using the regression residuals.[62]

To understand why condition (49) insures that the same outcome is obtained as with normality, consider our first-stage regression system, which takes the form of a seemingly unrelated regression (SUR) system:

$$\underbrace{Y}_{Tx2} = \underbrace{X}_{Tx2}\underbrace{\beta}_{2x2} + \underbrace{u}_{Tx2},$$

where X denotes the (deterministic) regressors defined in (50) and Y the stacked vector of the left-hand side variables (p_{jzt}, P_{jzt}) in equations (46) and (47). Letting u_t denote the residuals at date t and u the stacked residual vector, we have $E[u_t] = 0$ and

$$Var(u_t) = \begin{pmatrix} v_{11}^2\, v_{12} \\ v_{12}\, v_{22}^2 \end{pmatrix}.$$

[62] Using the notation introduced below, this follows from the fact that

$$\begin{aligned} &\left(X'V^{-1}X\right)^{-1}X'V^{-1}\widehat{u} \\ &= \left(X'V^{-1}X\right)^{-1}X'V(I - X\left(X'V^{-1}X\right)^{-1}X'V^{-1})Y \\ &= 0. \end{aligned}$$

Since the SUR system does not feature exclusions restrictions, OLS estimation is identical to GLS estimation. In particular, the OLS/GLS estimate $\widehat{\beta}$ of β is given by

$$\widehat{\beta} \equiv (X'X)^{-1} X'Y$$

and the regression residuals by

$$\underbrace{\widehat{u}}_{Tx2} = MY = Mu \text{ where } M \equiv (I - X(X'X)^{-1} X') \tag{51}$$

We have

$$\begin{aligned} E[\widehat{u}'\widehat{u}|X] &= E[\underbrace{u'}_{2xT}\underbrace{M'M}_{TxT}\underbrace{u}_{Tx2}|X] \\ &= E[\underbrace{u'}_{2xT}\underbrace{M}_{TxT}\underbrace{u}_{Tx2}|X] \\ &= tr(M)E[u'u|X] \\ &= \frac{1}{T-2}\begin{pmatrix} v_{11}^2 & v_{12} \\ v_{12} & v_{22}^2 \end{pmatrix}, \end{aligned}$$

An unbiased estimate of the residual variance v_{11}^2 is thus given by

$$\widehat{v_{11}^2} \equiv \frac{(1,0)'\widehat{u}'\widehat{u}(1,0)}{T-2}. \tag{52}$$

The estimation errors in the variables used in the second-stage regression, i.e., of $\left((0,1)\left(\widehat{\beta}-\beta\right)(0,1)'\right)^2$ and $\left(\widehat{v_{11}^2} - v_{11}^2\right)$, are orthogonal if and only if

$$\begin{aligned} & E[\left(\left((0,1)\left(\widehat{\beta}-\beta\right)(0,1)'\right)^2\right)\left(\widehat{v_{11}^2} - v_{11}^2\right)|X] \stackrel{!}{=} 0 \\ \Leftrightarrow\ & E[\left((0,1)(X'X)^{-1}X'u(0,1)'\right)^2\left(\frac{(1,0)'\widehat{u}'\widehat{u}(1,0)}{T-2} - v_{11}^2\right)|X] \stackrel{!}{=} 0 \\ \Leftrightarrow\ & E[\left((0,1)(X'X)^{-1}X'u(0,1)'\right)^2\left(\frac{(1,0)'u'u(1,0)}{T-2} - v_{11}^2\right)|X] \stackrel{!}{=} 0 \end{aligned}$$

The last equality holds if and only if

$$\begin{aligned} & E[\left((0,1)(X'X)^{-1}X'u(0,1)'\right)^2\frac{(1,0)'u'Mu(1,0)}{T-2}|X] \\ & = E[\left((0,1)(X'X)^{-1}X'u(0,1)'\right)^2 v_{11}^2|X], \end{aligned}$$

which is the case if and only if condition 49 holds, as $E[\frac{(1,0)'uM'Mu(1,0)}{tr(M'M)}] = v_{11}^2$.

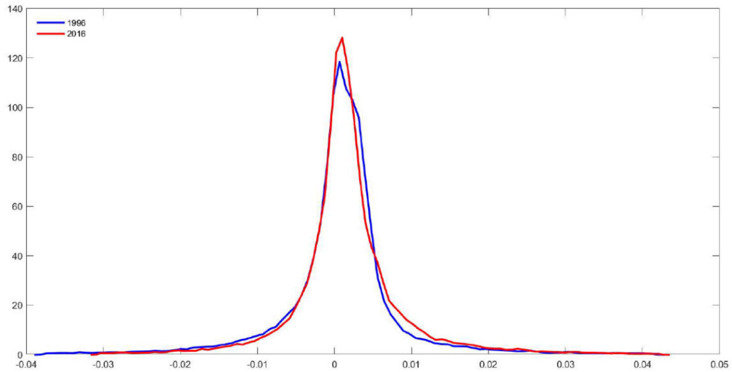

Figure 19: Distribution of product-specific optimal inflation rates Π^*_{jz} in 1996 and 2016 (monthly rates, unweighted)

F Cross Sectional Distribution of Product-Specific Optimal Inflation Rates over Time

Figure 19 depicts the cross-sectional distribution of product-specific optimal inflation rates Π^*_{jz} across all products and all items in the first and last year of our sample (1996 and 2016). It shows that this distribution is remarkably stable over the considered 20 year period.

G Proof of Proposition 3

From equation (21) we get

$$\begin{aligned} Var^j\left(\ln p_{jzt}\right) &= Var^j\left(\ln p^*_{jz} - \ln \Pi^*_{jz} \cdot t\right) + Var^j\left(u_{jzt}\right) \\ &+ Cov^j(\ln p^*_{jz}, u_{jzt}) \\ &- t \cdot Cov^j(\ln \Pi^*_{jz}, u_{jzt}). \end{aligned}$$

We next show that $Cov^j(\ln p^*_{jz}, u_{jzt}) = Cov^j(\ln \Pi^*_{jz}, u_{jzt}) = 0:$

$$\begin{aligned} Cov^j(\ln p^*_{jz}, u_{jzt}) &= E^j[\ln p^*_{jz} u_{jzt}] - E^j[\ln p^*_{jz}]\underbrace{E^j[u_{jzt}]}_{=0} \\ &= E^j[E^j[\ln p^*_{jz} u_{jzt} | p^*_{jz}]] \\ &= E^j[\ln p^*_{jz} \underbrace{E^j[u_{jzt} | p^*_{jz}]}_{=0}] \\ &= 0. \end{aligned}$$

Similarly:

$$
\begin{aligned}
Cov^j(\ln \Pi^*_{jz}, u_{jzt}) &= E^j[\ln \Pi^*_{jz} u_{jzt}] - E^j[\ln \Pi^*_{jz}]\underbrace{E^j[u_{jzt}]}_{=0} \\
&= E^j[E^j[\ln \Pi^*_{jz} u_{jzt} | \Pi^*_{jz}]] \\
&= E^j[\ln \Pi^*_{jz} \underbrace{E^j[u_{jzt} | \Pi^*_{jz}]}_{=0}] \\
&= 0.
\end{aligned}
$$

It thus only remains to compute the cross-sectional variance of residuals, $Var^j(u_{jzt})$. These residuals are described by a mixture distribution in which one first draws the relative price trend $\Pi_z^{*(i)}$ with probability m_{zi}. Subsequently, we draw corresponding residuals u_{jzt}. Since the residuals are independent across j, the cross-variance of residuals for any given $\Pi_z^{*(i)}$ is equal to their variance over time, as given in equation (22). Therefore, the variance of the mixture distribution is given by equation (24).