Ramsey optimal inflation with heterogeneous firms *

Anton Nakov Henning Weber

ECB, CEPR, and Bundesbank

ECB

19 Nov 2025

*The views expressed here are those of the authors only and do not necessarily coincide with the views of the Bundesbank, ECB, or the Eurosystem.

Introduction

- Golosov & Lucas (2007) started a literature highlighting the key role of large *idiosyncratic shocks* for firms' price setting and inflation dynamics
- Sticky-price models with large idiosyncratic shocks are now widely used for *positive analysis* of monetary non-neutrality and shock transmission
- But *normative analysis* in such models, e.g., on the optimal inflation rate, is scarce
- "Folk wisdom" suggests that standard New Keynesian (NK) logic is immune to firm-level shocks
- *This paper:* Revisit NK prescription of zero inflation target in light of firm-level heterogeneity from the perspective of a Ramsey planner
- *Finding:* Zero inflation is no longer optimal once forward-looking pricing is combined with *mean-reverting* firm-level shocks

Core mechanism: price distortion with mean-reverting shocks

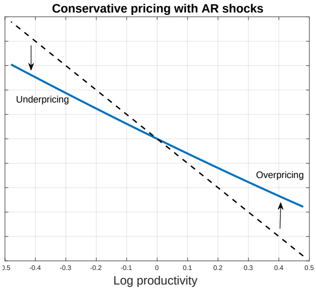
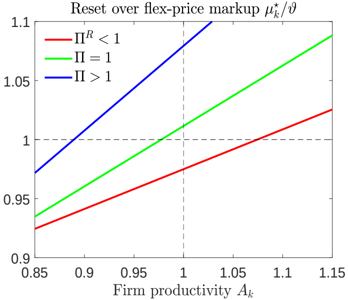


Figure: Pricing "conservatively" with mean-reverting productivity

Main findings

- Firms markups systematically deviate from flex-price benchmark because firms anticipate future productivity path



- *Deflation* reduces overpricing (reducing reset markup) of high-productivity firms
- *Inflation* reduces underpricing (raising reset markup) of low-productivity firms
- Because high-productivity firms bear a lion share of aggregate demand, and are more sensitive to inflation, a moderate *negative inflation* rate emerges as socially optimal.

Main findings

- Characterize Ramsey inflation analytically in stylized sticky-price models (Taylor, Calvo)
- Optimal inflation $\Pi^\star = -1.75\%$ in calibrated state-dependent pricing model
- Sizeable economic costs of deviating from optimal inflation

Productivity loss per period (bp)	$\Pi = +2\%$	$\Pi = +4\%$
Calvo pricing	34	90
State-dep pricing (incl menu costs)	15	31

- We do not interpret this to imply that the inflation target literally should be negative given all the motives for a positive target omitted from our analysis
- Rather, we provide a new argument against raising targets above 2% as some suggest

Related work

- Optimal inflation in sticky price models with non-stationary idiosyncratic “quality” shocks
  - Blanco (2021): Optimal target for a Taylor rule in a model with ZLB
  - Karadi, Nakov, Nuno, Pasten & Thaler (2024): Ramsey optimal stabilization policy
- Welfare costs of inflation in sticky price models with idiosyncratic shocks (Burstein & Hellwig 2008, Nakamura, Steinsson, Sun & Villar 2018, Midrigan, Blanco, Boar & Jones 2024, Cavallo, Lippi & Miyahara 2024, Jenuwine & Tielens 2025, Adam, Alexandrov & Weber 2025)
- Mean-reverting idiosyncratic shocks are generic feature in literature on monetary non-neutrality (Golosov & Lucas 2007, Nakamura & Steinsson 2008, Midrigan 2011, Karadi & Reiff 2019, Dotsey & Wolman 2020)

Sticky-price model with firm-level productivity shocks

- Representative household with discount rate $\beta \in (0, 1)$
- Aggregate output Y_t is CES composite with substitution elasticity θ
- At date t, firm $j \in [0, 1]$ employs labor to produce output: $Y_{jt} = A_{jt} L_{jt}$
- Firm productivity $A_{jt} > 0$ is stationary Markov process
- Firm sells output at price P_{jt} under monopolistic competition
- Price setting cases: flexible prices, Taylor, Calvo, random menu costs
- Flexible-price markup with sales subsidy $\tau \geq 0$

$$\vartheta \equiv \frac{\theta}{\theta - 1} \frac{1}{1 + \tau}$$

Definitions

- Markup of firm j with nominal wage W_t

$$\mu_{jt} \equiv \frac{P_{jt}}{W_t / A_{jt}}$$

- Product demand $Y_{jt} / Y_t = (P_{jt} / P_t)^{-\theta}$

- Aggregate markup is a weighted harmonic mean of firm markups

$$\mu_t \equiv \Big(\int_0^1 \underbrace{(P_{jt} / P_t)^{1-\theta}}_{\text{expenditure share}} \mu_{jt}^{-1} \, \mathrm{dj} \Big)^{-1}$$

- Aggregate productivity (inverse price dispersion) A_t is endogenous

- Gross inflation $\Pi_t \equiv P_t / P_{t-1}$ with price level P_t

Markup dispersion implies across-firm misallocation

- Markup dispersion $\mu_{jt} \neq \mu_t$ acts like a wedge between relative prices and productivities

$$\frac{P_{jt}}{P_t} = \frac{\mu_{jt}}{\mu_t} \frac{A_t}{A_{jt}}$$

  $\Longrightarrow$ inefficient allocation of resources across firms

- *Flexible-price markup* ϑ is constant despite time varying productivity A_{jt}

  $\Longrightarrow$ no markup dispersion, $\mu_{jt} = \vartheta = \mu_t$, hence no misallocation

- *Sticky-price markups* co-vary with firm productivity:

  High firm productivity cuts marginal cost and raises markup $\Longrightarrow$ misallocation

How does inflation affect markup dispersion?

- Two-period Taylor contracts and *iid* firm-level productivity shock with two states A_k, $k = 1, 2$
- Markup of firm with reset price $P^\star_{kt}$ and productivity A_k

$$\mu^\star_k \equiv \frac{P^\star_{kt}}{W_t} A_k$$

- Expected non-reset markup denoting firm's expected productivity by $A^\circ \equiv \left(E_t[1/A_{kt+1}]\right)^{-1}$

$$\mu^\circ_k \equiv \frac{P^\star_{kt}}{W_{t+1}} A^\circ$$

- Taking the ratio and with $\frac{W_{t+1}}{W_t} = \frac{P_{t+1}}{P_t}$

$$\frac{\mu^\star_k}{\mu^\circ_k} = \frac{P_{t+1}}{P_t} \frac{A_k}{A^\circ}$$

How does inflation affect markup dispersion?

- If we were to require zero markup dispersion over price spell, the $\mu_k^\star = \mu_k^\circ$ then

$$\frac{P_{t+1}}{P_t} = \frac{A^\circ}{A_k}$$

  Above inflation rate aligns firm's relative price dynamics with expected productivity evolution.

- Negative inflation for highly productive firms (anticipating lower productivity in the future, $A^\circ < A_k$) lowers incentives to overprice
- And positive inflation for less productive firms, $A^\circ > A_k$
- *Tradeoff*: Any chosen inflation rate cannot eliminate markup dispersion both for the high and the low productive firms
- *Question:* How does Ramsey planner weigh these firms when setting one inflation for all?

Proposition

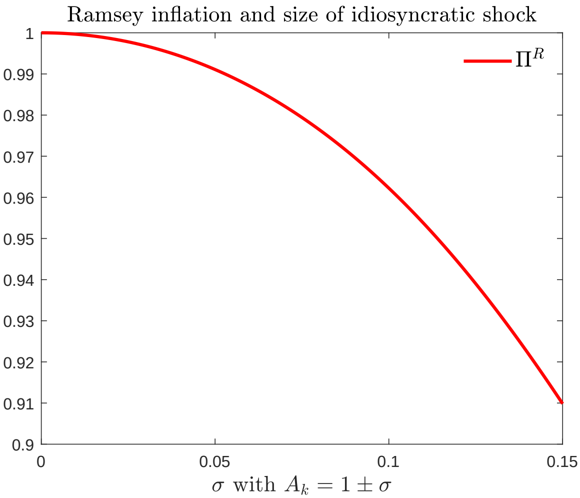
In the limit with $\beta \to 1$, gross inflation in the Ramsey steady state is given by

$$\Pi^R = \sum_{k=1,2} \left(\frac{(\mu_k^\star / A_k)^{-\theta}}{\sum_{h=1,2} (\mu_h^\star / A_h)^{-\theta}} \right) \cdot (A^\circ / A_k),$$

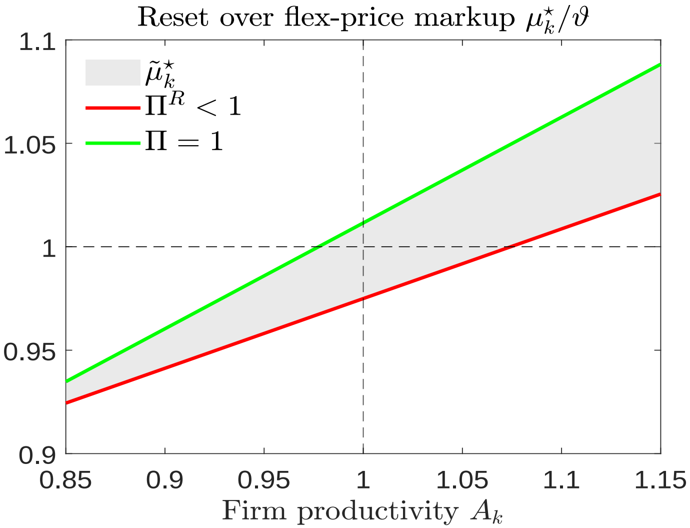
where output share of reset firms in state k depends on reset markup $\mu_k^\star(\Pi^R)$.

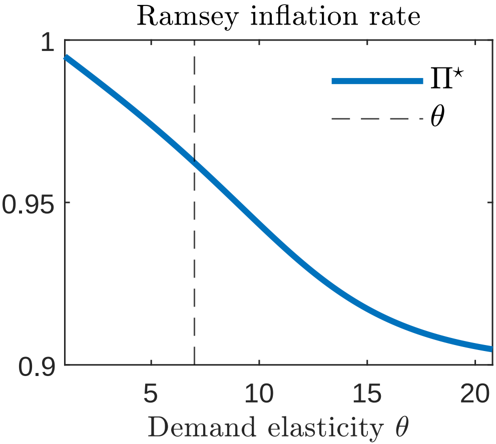
- Planner weighs more the path of productive reset firms
  1. They absorb more aggregate demand since they sell at a lower price
  2. They are also more sensitive to markup dispersion than less productive firms

$\Longrightarrow$ Overall, Ramsey inflation is always negative with heterogeneous sticky-price firms!

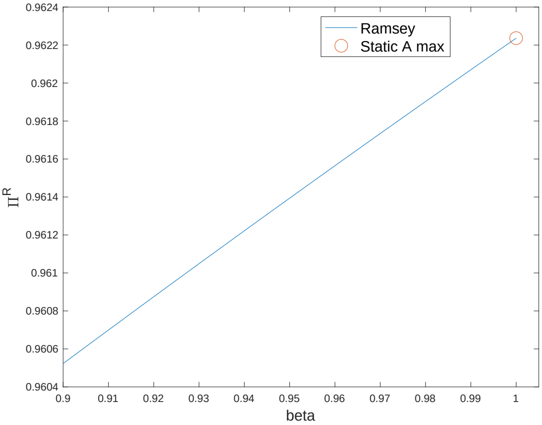

Ramsey inflation and size of idiosyncratic shock
1
0.99
0.98
0.97
0.96
0.95
0.94
0.93
0.92
0.91
0.9
Π^R
0
0.05
0.1
0.15
σ with $A_k = 1 \pm \sigma$

Ramsey inflation Π^R minimizes but does not eliminate markup dispersion




Ramsey inflation rate
1
0.95
0.9
5
10
15
20
Demand elasticity θ
$\Pi^\star$
θ

Productivity maximization gives upper bound on Ramsey inflation



We prove analytically that with $\beta \to 1$ Ramsey SS inflation maximizes steady-state productivity A

Ramsey problem with Taylor contracts and *iid* shocks

$$\max \sum_{t=0}^{\infty} \beta^t U\left(\frac{A_t}{\mu_t}, \frac{1}{\mu_t}\right) \text{ subject to}$$

$$
\begin{aligned}
\frac{1}{\vartheta} &= s_{t+1}\frac{1}{\mu^\star_{kt}} + (1-s_{t+1})\frac{1}{\mu^\circ_{kt+1}}, \quad k=1,2 \\
s_{t+1} &= (1+\beta\Pi_{t+1}^{\theta-1})^{-1} \\
\frac{\mu^\star_{kt}}{\mu_t}\frac{A_t}{A_k} &= \Pi_{t+1}\cdot\left(\frac{\mu^\circ_{kt+1}}{\mu_{t+1}}\frac{A_{t+1}}{A^\circ}\right), \quad k=1,2 \\
\frac{1}{\mu_t} &= \sum_{k=1,2}\left[\frac{1}{4}\left(\frac{\mu^\star_{kt}}{\mu_t}\frac{A_t}{A_k}\right)^{1-\theta}\frac{1}{\mu^\star_{kt}} + \frac{1}{4}\left(\frac{\mu^\circ_{kt}}{\mu_t}\frac{A_t}{A^\circ}\right)^{1-\theta}\frac{1}{\mu^\circ_{kt}}\right] \\
1 &= \sum_{k=1,2}\left[\frac{1}{4}\left(\frac{\mu^\star_{kt}}{\mu_t}\frac{A_t}{A_k}\right)^{1-\theta} + \frac{1}{4}\left(\frac{\mu^\circ_{kt}}{\mu_t}\frac{A_t}{A^\circ}\right)^{1-\theta}\right]
\end{aligned}
$$

Calvo pricing and general process for productivity shock

- Firm productivity A_{kt} is stationary Markov process with K states
- Reset markup of firm with productivity A_k

$$\mu_k^\star = \frac{P_{kt}^\star}{W_t} A_k$$

- Expected non-reset markup of this firm at price spell $i > 0$

$$\mu_{ki}^\star = \frac{P_{kt}^\star}{W_{t+i}} A_{k|i}$$

with expected productivity at spell i starting from state k given by $A_{k|i} = \left(E_t[1/A_{kt+i}]\right)^{-1}$

- Expected markup path related to expected productivity path according to

$$\frac{\mu_k^\star}{\mu_{ki}^\star} = \frac{P_{t+i}}{P_t} \frac{A_k}{A_{k|i}} \quad \text{for each price spell } i = 1, 2, \ldots$$

Calvo model with persistent firm level shocks

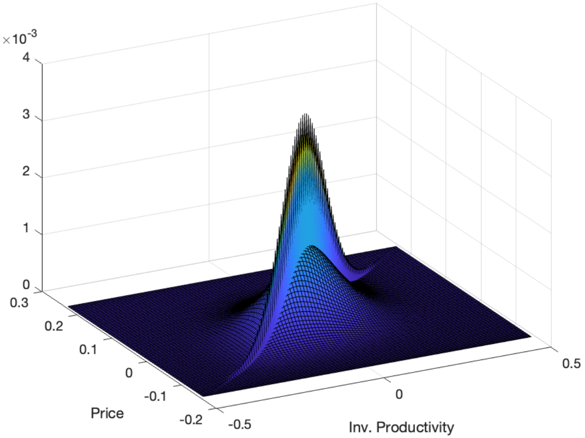


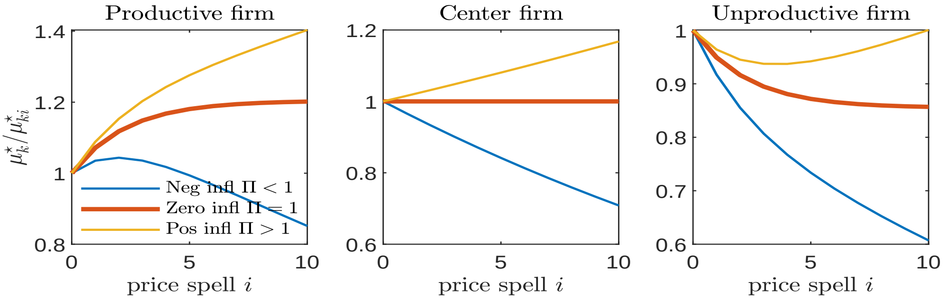
Figure: Stationary distribution of firms

Calvo pricing: Unavoidable markup dispersion over price spell

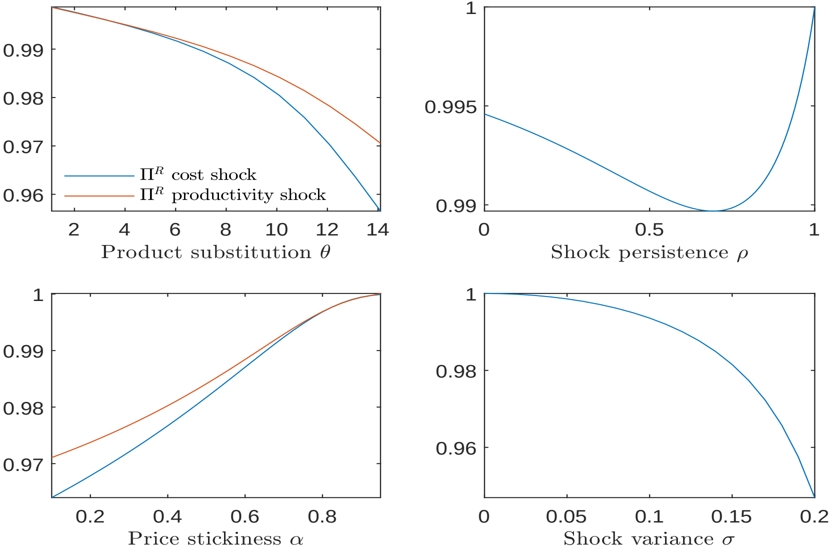
- Condition for constant markups $\mu_k^\star = \mu_{ki}^\star$ over price spell generally fails

$$1 \neq \frac{P_{t+i}}{P_t} \frac{A_k}{A_{k|i}}$$

- P_{t+i} diverges whereas $A_{k|i}$ converges $\implies$ unavoidable markup dispersion



Ramsey inflation is negative with Calvo pricing and general shock process



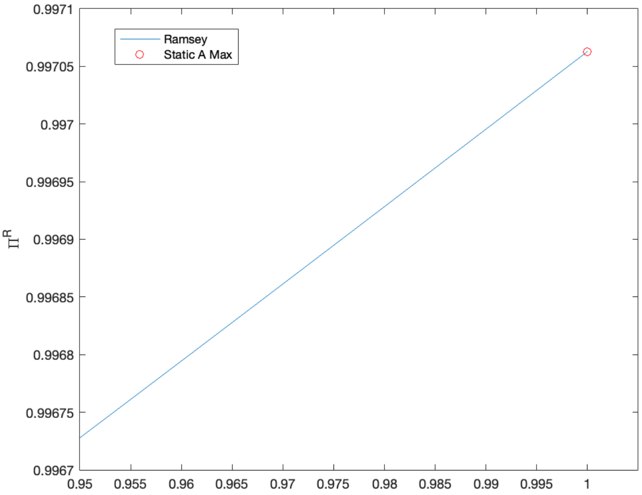
Ramsey problem with Calvo pricing and general shock

- To solve Ramsey problem with $\beta < 1$, aggregate Calvo model with idiosyncratic shocks analytically
- Define idiosyncratic *cost shock* as inverse productivity $Z_{kt} \equiv 1/A_{kt}$
- Z_{kt} follows a stationary Markov process with centro-symmetric transition matrix F determined by shock persistence $\rho < 1$ as in Rouwenhorst (1995)
- For example, $K = 3$ yields

$$F \quad = \quad \frac{1}{2}\begin{bmatrix} \frac{(1+\rho)^2}{2} & \frac{1-\rho^2}{2} & \frac{(1-\rho)^2}{2} \\ 1-\rho^2 & 1+\rho^2 & 1-\rho^2 \\ \frac{(1-\rho)^2}{2} & \frac{1-\rho^2}{2} & \frac{(1+\rho)^2}{2} \end{bmatrix}$$

- Analytical aggregation essential to verify accuracy of numerical solution approach for SDP model

Ramsey inflation in $\beta \to 1$ limit represents upper bound



Steady-state Ramsey inflation with $\beta \to 1$ maximizes steady-state productivity A

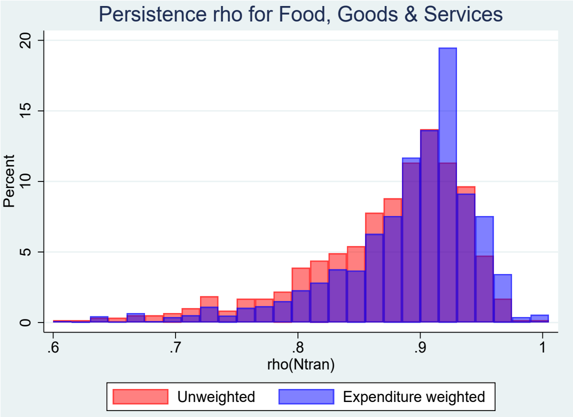
Calibrating ρ with micro price data from German CPI

- With Calvo, de-meaned reset price related to *inverse* firm productivity shock

$$r_{jt} \equiv \frac{P^{\star}_{jt}}{E_j(P^{\star}_{jt})} - 1 = \left(\frac{1-\alpha\Pi^{\theta}}{1-\alpha\rho\Pi^{\theta}}\right) \cdot (Z_{jt} - E(Z_{jt}))$$

- At date t, sort r_{jt}'s into bins implied by quantiles of ergodic distribution ξ
- Track transitions of r_{jt} between bins over time
- Observed transition freq $\hat{f}_{ij}$ yield estimate of transition matrix, $\widehat{F} = (\hat{f}_{ij})$
- Recover $\hat{\rho}$ from $\widehat{F}$ using vector of standardized (inverse) productivities $h(K)$

$$\hat{\rho} = \frac{h'\widehat{F}h}{h'h}$$



- Baseline sample for Jan 2015 to Dec 2021
  - replaces prices on sale by sale-adjusted prices from DESTATIS
  - drops imputed prices and product substitutions
  - covers 73% of official basket and 635 COICOPs
  - contains 4.5 million r_{jt} transitions (3244 in median COICOP)

Sectoral breakdown

	Baseline sample		Sample incl substitutions	
Persistence $\hat{\rho}$	Mean	SD	Mean	SD
Food	.872	.052	.868	.053
Goods	.905	.050	.897	.073
Services	.888	.086	.878	.091
Energy	.736	.113	.763	.103
Housing	.964	.019	.939	.048
Total	.893	.092	.884	.088

# of r_{jt} transitions	Sum	Median	Mean	Min	Max
Food	1293449	3157	6916.84	367	65310
Goods	2653732	4514.5	8397.886	10	85146
Services	131634	471	1278	9	26199
Energy	425709	1424.5	23650.5	225	145999
Housing	11243	691	1022.091	96	4524
Total	4515767	3244	7111.444	9	145999

Calibrated state-dependent pricing model

How does state-dependent pricing (SDP) affect optimal inflation?

- Random menu costs are *iid* and uniform over $[0, \bar{\xi}]$, as in Boar et al (2025)
- Adjustment probability function with probability of "free" price change $\bar{\lambda}$

$$\lambda(p, a) = \bar{\lambda} + (1 - \bar{\lambda}) \cdot \min\left(\frac{V^{adj}(a) - V^{nadj}(p, a)}{\bar{\xi} w}, 1\right)$$

- p denotes log relative price of firm with log productivity a which evolves as

$$a_{jt} \quad = \quad \rho a_{jt-1} + \sigma \epsilon_{jt}, \quad \epsilon_{jt} \sim \mathrm{N}(0, 1)$$

and is discretized to finite-state Markov process as in Rouwenhorst (1995)

Monthly calibration for Euro Area micro price statistics

Description	Parameter	Value	Source
Persistence productivity	ρ	0.9	Monthly German CPI data
Discount factor	β	$1.02^{-1/12}$	Annual real interest of 2%
Elasticity of substitution	θ	7	Golosov and Lucas (2007)

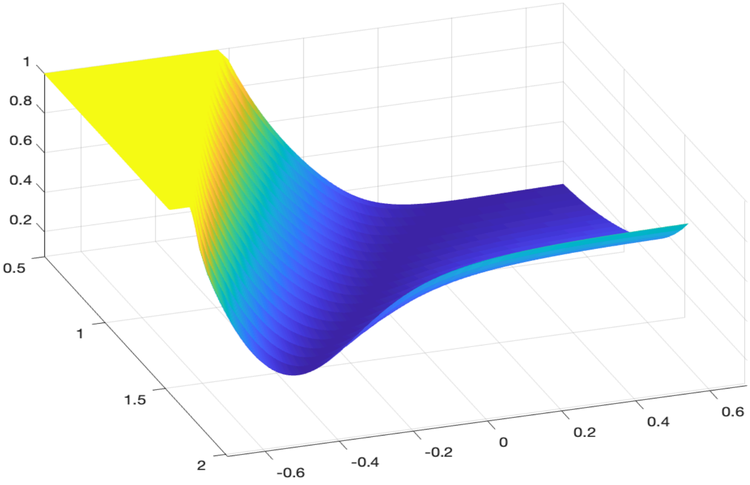
Targeted Moment	Data	Model
Frequency Δp (%)	9.5	9.5
Size Δp (%)	9.3	9.3
Kurtosis Δp	3.2	3.2

Source: Karadi et al (2023) (Euro Area 4)

Description	Parameter	Calibration
Maximum menu cost	$\bar{\xi}$	1.6319
Probability of free change	$\bar{\lambda}$	0.0746
Std dev of productivity shock	σ	0.1158

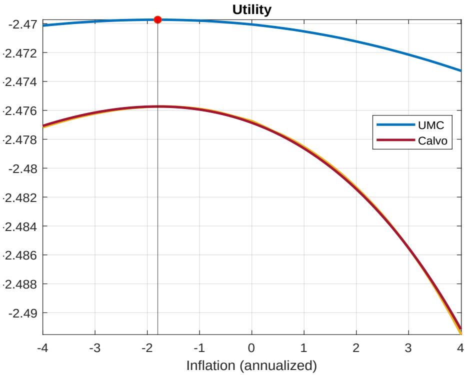
At $\Pi = 2\%$, calibrated menu costs absorb $\approx 0.3\%$ of hours worked or $\approx 2\%$ of average profits

Adjustment probability function



Negative inflation maximizes steady-state utility with SDP

- We cannot solve Ramsey problem for SDP model with uniform menu costs numerically yet
- Prior results suggest Ramsey inflation close to utility-maximizing steady-state optimal inflation
- Optimal inflation in calibrated SDP model: $\Pi^{\star} = -1.75\%$ per year
- Similar rate maximizes utility in Calvo (using same shock process calibration)



Sizable economic costs of deviating from optimal inflation

Productivity loss per period (bp)	$\Pi = +2\%$	$\Pi = +4\%$
Calvo pricing	34	90
State-dep pricing (incl menu costs)	15	31

Conclusions

- Models consistent with observed micro-price heterogeneity imply sizable misallocation at zero inflation and a role for monetary policy to reduce it
- Price stickiness combined with idiosyncratic productivity shocks distorts relative prices thereby reducing aggregate productivity
- This distortion provides a new motive for choosing lower $\pi^\star$ even when abstracting from Friedman rule considerations
- Our analysis abstract from known motives for choosing higher $\pi^\star$ such as downward wage rigidity, ZLB or trends in efficient relative prices
- Results warn against raising inflation target above 2%

Discretized firm-level shocks

Lemma

Let firm-level shocks $Z_{jt} = 1 + \hat{z}'\xi_{jt}$, *with* $E(Z_{jt}) = 1$, $Z_{jt} > 0$, *state* ξ_{jt} *equal to* e_j *in state* j *and* $\hat{z} = [\hat{z}_1 \ldots \hat{z}_j \ldots \hat{z}_K]'$, *discretize a continuous, symmetric AR(1) process with white-noise residual, persistence* $|\rho| < 1$ *and standard deviation* σ_z *as in Rouwenhorst (1995) and Kopecky & Suen (2010). Then,*

(i) *transition matrix* F *is centro-symmetric and determined by persistence* ρ

(ii) *the conditional expectation at horizon* s *is given by*

$$E_t[\hat{z}'\xi_{jt+s}] = \hat{z}_j\rho^s,$$

where $\hat{z}_j$ *is determined by* σ_z *and the number of states* K*.*

back

Ramsey problem with Calvo pricing and general shock

$$\max \sum_{t=0}^{\infty} \beta^t U\left(\frac{A_t}{\mu_t}, \frac{1}{\mu_t}\right) \text{ subject to initial conditions and}$$

$$\mu^\star_{kt,t} = \mu^\star_{kt,t+i}\left(\frac{A_{t+i}}{\mu_{t+i}}\frac{\mu_t}{A_t}\right)\frac{P_{t+i}}{P_t}A_k E_t\left[\frac{1}{A_{kt+i}}\right], \quad k = 1, \dots K$$

$$\vartheta^{-1} = \sum_{i=0}^{\infty} s_{t,t+i}(\mu^\star_{kt,t+i})^{-1}, \quad k = 1, \dots K$$

$$s_{t,t+i} = \frac{(\alpha\beta)^i\left(P_{t+i}/P_t\right)^{\theta-1}}{\sum_{h=0}^{\infty}(\alpha\beta)^h\left(P_{t+h}/P_t\right)^{\theta-1}}$$

$$\frac{1}{\mu_t} = \sum_{k=1}^{K}\xi_k\left\{(1-\alpha)\sum_{i=0}^{\infty}\alpha^i\left(\frac{P^\star_{kt-i}}{P_t}\right)^{1-\theta}\frac{1}{\mu^\star_{kt-i,t}}\right\}$$

$$1 = \sum_{k=1}^{K}\xi_k\left\{(1-\alpha)\sum_{i=0}^{\infty}\alpha^i\left(\frac{P^\star_{kt-i}}{P_t}\right)^{1-\theta}\right\}$$

$$\sum_{t=0}^{\infty} \beta^t \left(\ln \frac{A_t}{\mu_t} - \frac{1}{\mu_t} \right)$$

$$\frac{\mu_{kt}^{\star}}{\vartheta}(1+\hat{z}_k) = \frac{N_t^{co} + \hat{z}_k N_t^{id}}{D_t} \;, \quad k = 1, \ldots K$$

$$N_t^{co} = 1 + \alpha\beta \left(\frac{A_{t+1}}{\mu_{t+1}} \frac{\mu_t}{A_t} \right) \Pi_{t+1}^{\theta} N_{t+1}^{co}$$

$$N_t^{id} = 1 + \alpha\rho\beta \left(\frac{A_{t+1}}{\mu_{t+1}} \frac{\mu_t}{A_t} \right) \Pi_{t+1}^{\theta} N_{t+1}^{id}$$

$$D_t = 1 + \alpha\beta \Pi_{t+1}^{\theta-1} D_{t+1}$$

$$\frac{1}{\mu_t} = H_t^{co} + H_t^{id}$$

$$H_t^{co} = (1-\alpha) \sum_{k=1}^{K} \xi_k \left(\frac{P_{kt}^{\star}}{P_t} \right)^{1-\theta} \frac{1}{\mu_{kt}^{\star}} \frac{1}{1+\hat{z}_k} + \alpha \Pi_t^{\theta} \left(\frac{A_t}{\mu_t} \frac{\mu_{t-1}}{A_{t-1}} \right) H_{t-1}^{co}$$

$$H_t^{id} = (1-\alpha) \sum_{k=1}^{K} \xi_k \left(\frac{P_{kt}^{\star}}{P_t} \right)^{1-\theta} \frac{1}{\mu_{kt}^{\star}} \frac{\hat{z}_k}{1+\hat{z}_k} + \rho\alpha \Pi_t^{\theta} \left(\frac{A_t}{\mu_t} \frac{\mu_{t-1}}{A_{t-1}} \right) H_{t-1}^{id}$$

$$\frac{P_{kt}^{\star}}{P_t} = \frac{\mu_{kt}^{\star}}{\mu_t} A_t \bar{z}(1+\hat{z}_k) \;, \quad k = 1, \ldots K$$

$$1 = (1-\alpha) \sum_{k=1}^{K} \xi_k \left(\frac{P_{kt}^{\star}}{P_t} \right)^{1-\theta} + \alpha \Pi_t^{\theta-1}$$

Recovering persistence $\hat{\rho}(s)$ from transition matrix $\widehat{F^s}$

- Track transitions of r_{jt} between bins over time
  - Of all r_{jt} in bin j, select those with same forward spell s of reset price
  - Compute observed transition frequencies $\hat{f}_{ij}^{(s)}$
  - Estimate s-period transition matrix, $\widehat{F^s} = (\hat{f}_{ij}^{(s)})$ (Discard degenerate $\widehat{F^s}$)
- Define vector h with elements $h_j = 2 \cdot \left(\frac{j-1}{K-1}\right) - 1, \quad$ for $j = 1 \dots K$
- Recover scalar $\hat{\rho}(s)$ from matrix $\widehat{F^s}$ according to

$$\hat{\rho}(s) = \left(\frac{h'\widehat{F^s}h}{h'h}\right)^{\frac{1}{s}} \text{ with } s \geq 1$$

- Aggregate $\hat{\rho}(s)$ over all forward spells s according to

$$\hat{\rho} = \sum_s g(s)\hat{\rho}(s),$$

with $g(s)$ denoting share of transitions of reset prices with spell s.

Calibrating ρ using micro price data from German CPI

Persistence $\hat{\rho}$	Mean	SD
$\hat{\rho}(\mathit{tran})$	0.893	0.092
$\hat{\rho}(\mathit{freq})$	0.894	0.087
$\hat{\rho}(\mathit{uwgh})$	0.910	0.079

Note: Stats weighted by expd shares at 10-digit COICOP level

Downward Nominal Wage Rigidity

- Add common exogenous productivity shock: $Y_{jt} = A_t^c A_{jt} L_{jt}$
- Definitions of aggregate markup and aggregate productivity imply

$$P_t = \mu_t \frac{W_t}{A_t^c A_t}$$

- Obtain $\mu_t = \vartheta$ with $\beta \to 1$ and Taylor contracts or Calvo pricing
- Treat aggregate variables as constant; taking growth rates and using $\Pi^W = W_t / W_{t-1}$ and $a^c = A_t^c / A_{t-1}^c$ yields

$$\Pi = \Pi^W / a^c$$

- Inflation Π can still be negative with downwardly rigid nominal wages $\Pi^W \geq 1$ and exogenous growth $a^c > 1$
- Amano, Moran, Murchison & Rennison 2009: NK model with sticky prices and wages