

# Rationing Under Sticky Prices

Tom Holden

Deutsche Bundesbank

Paper and slides available at <https://www.tholden.org/>.

The views expressed in this paper are those of the author and do not represent the views of the Deutsche Bundesbank, the Eurosystem or its staff.

# Motivation

- Covid, the Suez Canal blockage and the war in Ukraine all led to widespread **stockouts** and **delivery delays**.
  - Many supermarket shelves were empty during the pandemic.
  - In 2022, many new cars were subject to delivery delays of at least a year.
- Stockouts and delivery delays are forms of **rationing**. They are ultimately a choice of the supplier.
  - MC is never infinite. If you're prepared to pay a high enough amount for a production input, you can always obtain it.
- If prices were flexible, they would have increased proportionally to the increase in MC  $\Rightarrow$  no stockouts!
  - But with sticky prices, firms ration demand to avoid selling below MC.
- Rationing is **also common in normal times**.
  - Over 10% of all consumer goods in the US were out of stock pre-Covid (Cavallo & Kryvtsov 2023).

# Sticky prices inevitably lead to rationing

- If a firm cannot adjust its nominal price, then its **real price will decline** over time at the rate of inflation.
- A lower real price implies **higher demand** for its good. Higher demand means **higher marginal costs**.
- Eventually, its marginal costs (rising) will be greater than its price (falling) if it continues to meet all demand.
- No firm wants to sell at a price below marginal cost. Instead, it should stop producing, **rationing demand**.
- Yet essentially all the prior sticky price literature (Calvo or menu cost) assumes that firms always meet all demand.

# This paper

- What are the macroeconomic implications of allowing firms to ration?
- I allow for random rationing in a continuous time NK model with endogenous price rigidity, and find:
  1. Rationing generates a convex Phillips curve.
  2. Rationing massively reduces the welfare costs of positive trend inflation.
  3. True output falls following “expansionary” monetary shocks (but measured output increases).
- Basic mechanism:

High inflation  $\Rightarrow$  low markups  $\Rightarrow$  high rationing  $\Rightarrow$  fewer varieties consumed  $\Rightarrow$  downward pressure on output.

# Does rationing matter in practice?

*“Mark-ups are 10%, inflation is 2%, prices are updated at least once per year, real prices will not hit marginal cost.”*

- But: firm demand:  $y \propto \left(\frac{p}{P}\right)^{-\epsilon}$  ( $\epsilon \approx 10$ ) and marginal costs:  $mc \propto y^{\frac{\alpha}{1-\alpha}}$  ( $\alpha \approx \frac{3}{5}$ ), so  $mc \propto \left(\frac{p}{P}\right)^{-\epsilon \frac{\alpha}{1-\alpha}} \approx \left(\frac{p}{P}\right)^{-15}$ .
  - In the short run, some labour and intermediate inputs are fixed ( $\approx \frac{2}{5}$  at annual freq. (Abraham et al. 2024))  $\Rightarrow \alpha \approx \frac{3}{5}$ .
- So: A 2% fall in real prices increases marginal costs by 30%. Good-bye mark-ups! Hello rationing!
- This calculation understates firms' reasons to ration:
  - Firms face high frequency demand fluctuations. Mark-ups are much lower at times of high demand.
  - Inflation can be much higher than 2%. It was around 7% post-Covid!
  - Demand is growing over time due to aggregate income growth. A 2% increase in aggregate demand increases MC by 3%.
  - Marginal costs are also rising over time if not all capital depreciation can be fixed quickly.

# Price adjustment

- Why don't firms just change prices, rather than rationing?
  - Firms can benefit from tolerating  $P < MC$  if they anticipate lower MC (mean reversion?) or lower menu costs in future.
  - Modern menu cost models rely heavily on random menu costs and/or free price change opportunities.
  - By revealed preference, firms that can ration make higher profits than firms that cannot. Under rationing, profits always  $> 0$ .
  - Since profits are higher when rationing is allowed, lower menu (etc.) costs are needed to justify the observed price stickiness.
- I will take a tractable approach to state-dependent pricing broadly following Blanco et al. (2024).
  - Firms will be owned by conglomerates. Conglomerates choose the rate of price adjustment, not which firms adjust.
  - Provides aggregate state dependence.
  - Matches flat adjustment hazard rate found by Klenow & Kryvtsov (2008), Nakamura & Steinsson (2008), Klenow & Malin (2010).

# Prior literature

- Early:
  - Barro & Grossman (1971), Drèze (1975), Svensson (1984), Corsetti & Pesenti (2005) (restrict shocks to ensure no rationing).
- NK rationing models (all with sales-capped rationing):
  - Under sticky wages: Huo & Ríos-Rull (2020), Gerke et al. (2023): Infinite dimensional state, numerical.
  - Under sticky prices: Hahn (2022): Only steady state results. No dynamics. No idiosyncratic shocks.
- Alternatives capturing some of the effects of rationing:
  - Michaillat & Saez (2015), Liu (2025): search and matching, smoothing out the Barro & Grossman (1971) model.
  - Kharroubi and Smets (2024): shutdown if profits negative at demanded quantity.
- Stockouts in inventory models:
  - Alessandria, Kaboski & Midrigan (2010), Kryvtsov & Midrigan (2013), Bilal (2016).
  - In these papers, firms always meet demand if they have stock available, even if marginal value of that stock  $>$  price.

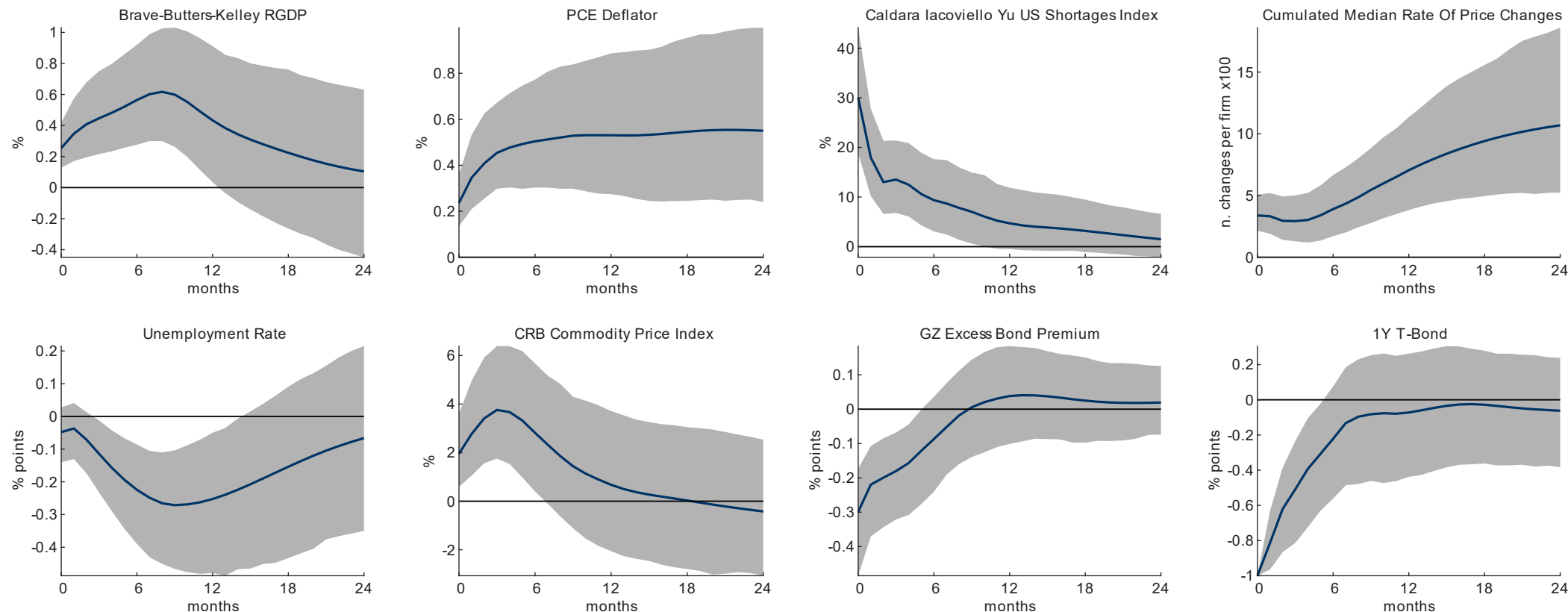
Evidence



# Empirical evidence for rationing

- Cavallo & Kryvtsov (2023) find that around 11% of all US consumer goods were out of stock (=rationing) in 2019.
- In 2022 (January-July), this number was around 23%. In line with my story: high inflation  $\Rightarrow$  high rationing.
  - Cavallo & Kryvtsov (2023) stress causality in the opposite direction. (Stockouts lead to inflation.)
- **I'll show 1:** Expansionary monetary shocks increase a measure of rationing and other shortages.
- **I'll show 2:** Increases in price adjustment rates are followed by reductions in stockouts.
- **I'll show 3:** Quantities sold are concave in price age, in line with goods with old prices being rationed.
- **I'll show 4:** Rationing helps match the convexity of the Phillips curve (Forbes, Gagnon & Collins 2022).
- **I'll show 5:** Rationing helps match the fast response of prices to cleanly identified monetary policy shocks.
  - "Clean" monetary shock papers: Miranda-Agrippino & Ricco (2021), Bauer & Swanson (2023).

# Impulse response to an expansionary monetary shock



- Estimated following Figure 3 of Miranda-Agrippino & Ricco (2021) (95% bands) but:
  - with Brave-Butters-Kelley RGDP (Brave, Cole & Kelley 2019; Brave, Butters & Kelley 2019) not IP, with PCEPI not CPI,
  - with the Caldara-Iacoviello-Yu US Shortages Index (Caldara, Iacoviello & Yu 2025),
  - and with the cumulation of the median rate of price changes, excluding sales (Montag & Villar 2025).

# Price-adjustment rates and stockouts

- Data from replication package of Cavallo & Kryvtsov (2023).

- Specification:

$$y_{i,j,t} = \beta x_{i,j,t-1} + \alpha_{i,j} + \gamma_{i,t} + \delta_{j,t} + \varepsilon_{i,j,t}$$

- $i$  indexes 7 countries (Canada, China, France, Germany, Japan, Spain, USA).
  - $j$  indexes 37 product categories (3-digit).       $t$  indexes 11 months in 2019.
  - $y_{i,j,t}$  is the **average stockout rate** (for a country, category, month), in percent.
  - $x_{i,j,t}$  is  $100 \times$  the **average annualized rate of price adjustment** (for a country, category, month).
  - Standard errors 3-way clustered. Clusters indexed by  $(i, j)$ ,  $(i, t)$ ,  $(j, t)$ .
- Result:  $\hat{\beta} = -0.155$ , std. err. 0.044, p-value  $< 0.001$ . Increased price flexibility reduces stockouts.

# Average output over the life of a price: Setup

- Data: Dominick's Finer Foods (1989-1994).
- 21,474,126 observations after dropping for each product and store:
  - First/last price, any price  $\neq$  cumulative max price, one week after each price change/missing, any price older than four years, one observation due to differencing.
- Specification, estimated via FGLS:

$$\frac{y_{i,j,k,l,t} - y_{i,j,k,l,t-1}}{\bar{y}_{i,j,k,l}} = \beta_{A(i,j,k,l,t)} + \alpha_{i,j,k} + \gamma_{i,j,t} + \delta_{i,k,t} + \sigma_{i,j,A(i,j,k,l,t)}^{(1)} \sigma_{i,j,k}^{(2)} \sigma_{i,j,t}^{(3)} \sigma_{i,k,t}^{(4)} \varepsilon_{i,j,t}$$

$i$  indexes narrow categories (86)

$k$  indexes products (10,166)

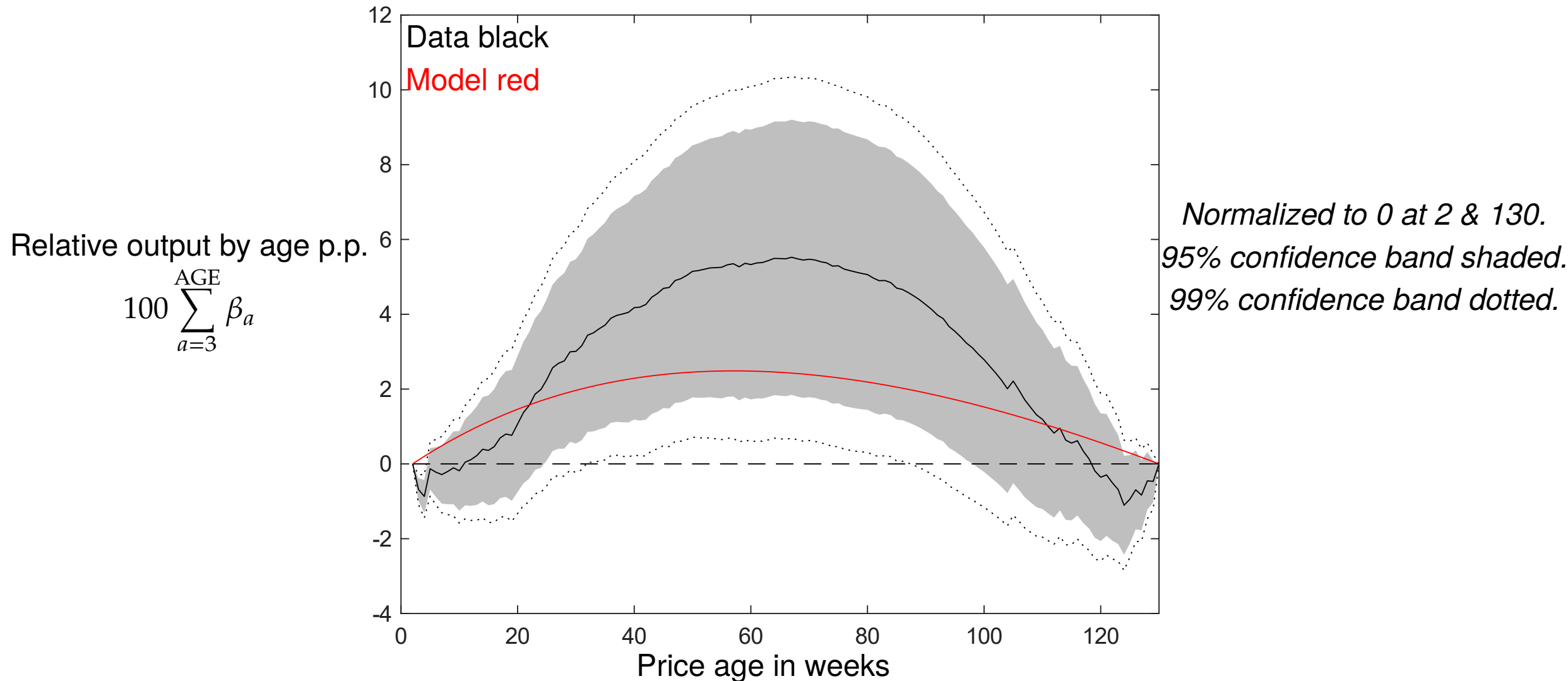
$t$  indexes weeks (398)

$j$  indexes stores (93)

$l$  indexes prices (947,660)

- $A(i,j,k,l,t)$  is the “age” of the  $i,j,k,l$  price at  $t$ .  $y_{i,j,k,l,t}$  gives units sold.  $\bar{y}_{i,j,k,l}$  is average of  $y_{i,j,k,l,t}$  over the life of the price.
- Standard errors 4-way clustered. Clusters indexed by:  $(i,j,A(i,j,k,l,t))$ ,  $(i,j,k)$ ,  $(i,j,t)$ ,  $(i,k,t)$ .

# Average output over the life of a price: Results



# The model

# Setup

- The model is in continuous time, with no aggregate uncertainty, just MIT shocks.
- Assume firm price change opportunities arrive at rate  $\lambda_t > 0$ .
- The time  $t$  density of firms that last updated at  $\tau$  is  $\lambda_\tau e^{-\int_\tau^t \lambda_v dv}$ . Note  $\int_{-\infty}^t \lambda_\tau e^{-\int_\tau^t \lambda_v dv} d\tau = 1$ .
- Index firms (and products) by the time they last updated their price,  $\tau$ , and by their demand shock  $\zeta$ .
- Firm output:  $y_{\zeta,\tau,t}$ .
- $g(\zeta)$  is the PDF of the demand shock, which is independent across time and firms.
- For tractability, I assume  $g(\zeta) = \theta \zeta^{\theta-1}$  where  $\theta > 0$  (so  $\zeta \sim \text{Beta}(\theta, 1)$ ). Mean  $\zeta$ :  $\frac{\theta}{\theta+1}$ . Variance  $\zeta$ :  $\frac{\theta}{(\theta+1)^2(\theta+2)}$ .

# Rationing and aggregation

- $\psi \in [0,1]$  denotes a purchaser-good-time-specific shock controlling if a given purchaser can buy a given good.
- $y_{\psi,\zeta,\tau,t}$ : sales to buyer with shock  $\psi$ , at  $t$ , of good produced by firm that updated price at  $\tau$ , with demand shock  $\zeta$ .
- Specialize to  $\psi$  uniform,  $y_{\psi,\zeta,\tau,t} = \begin{cases} y_{\zeta,\tau,t}^* & \psi \leq \bar{\psi}_{\zeta,\tau,t} \\ 0 & \psi > \bar{\psi}_{\zeta,\tau,t} \end{cases} \cdot y_{\zeta,\tau,t}^*$  sales when not rationed.
- $\bar{\psi}_{\zeta,\tau,t}$  is the probability a consumer is not rationed at the firm. Adjusts to ensure:  $y_{\zeta,\tau,t} = \int_0^1 y_{\psi,\zeta,\tau,t} d\psi = y_{\zeta,\tau,t}^* \bar{\psi}_{\zeta,\tau,t}$ .
- The aggregate good is produced from intermediates by a perfectly competitive industry with technology:

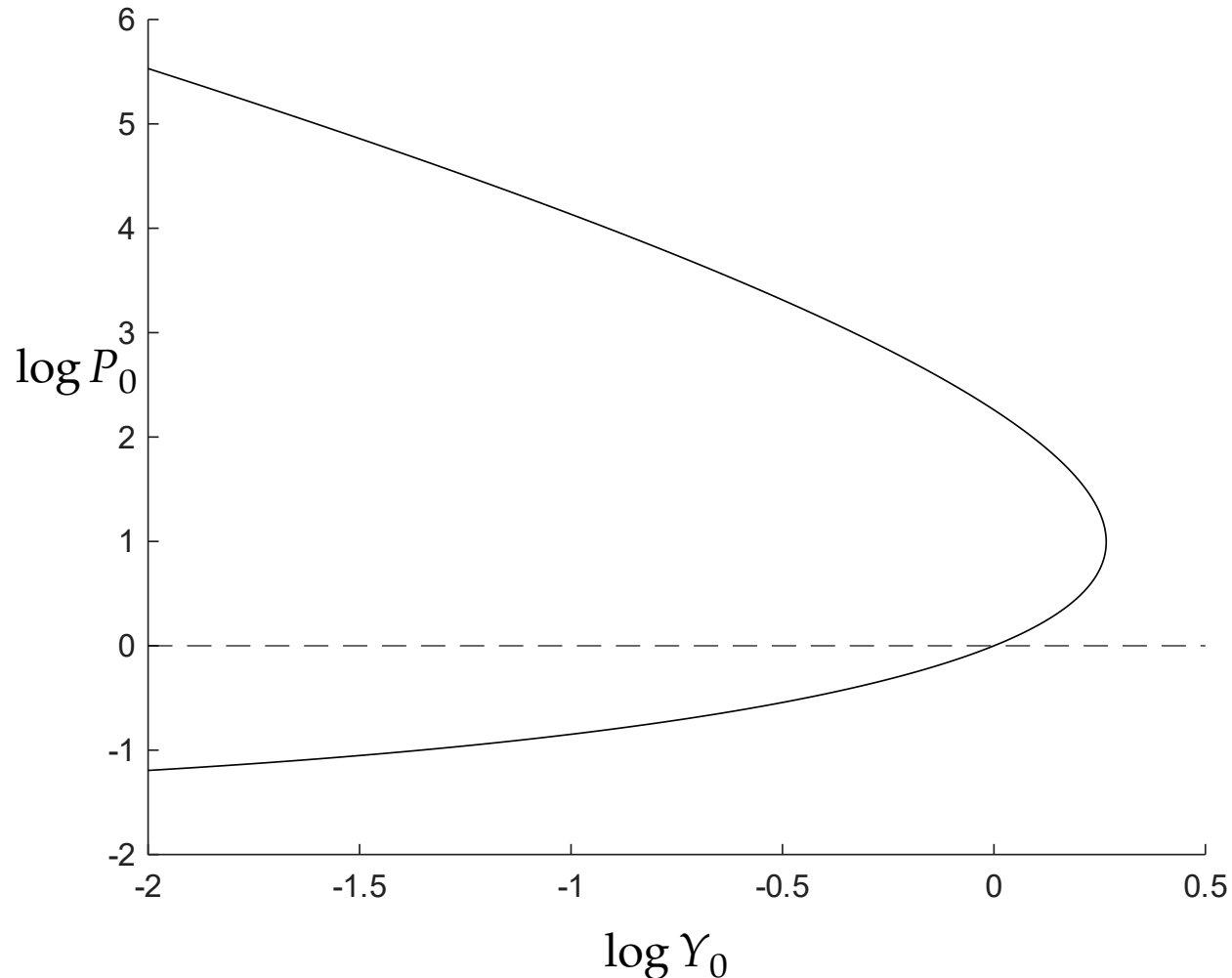
$$Y_t = D^{-\frac{\epsilon}{\epsilon-1}} \left[ \int_{-\infty}^t \lambda_\tau e^{-\int_\tau^t \lambda_v dv} \int_0^1 \zeta g(\zeta) \int_0^1 y_{\psi,\zeta,\tau,t}^{\frac{\epsilon-1}{\epsilon}} d\psi d\zeta d\tau \right]^{\frac{\epsilon}{\epsilon-1}}, \quad D = \frac{\theta}{\theta+1}$$



# Firm production

- The FOC of the aggregators imply demand must satisfy:  $y_{\zeta,\tau,t} \leq y_{\zeta,\tau,t}^* := \left(\frac{D p_\tau}{\zeta P_t}\right)^{-\epsilon} Y_t$ .
- Firms produce using the production technology:  $y_{\zeta,\tau,t} = (A_t l_{\zeta,\tau,t})^{1-\alpha}$ . Real wage is  $W_t$ .
- Firm flow real production profits are guaranteed to be positive for small enough  $l_{\zeta,\tau,t}$ .
- Optimal production: There is a quantity  $\bar{\zeta}_{\tau,t} := D \left(\frac{p_\tau}{P_t}\right)^{1+\frac{1-\alpha}{\epsilon\alpha}} \left(\frac{1-\alpha}{W_t/A_t}\right)^{\frac{1-\alpha}{\epsilon\alpha}} Y_t^{-\frac{1}{\epsilon}} > 0$  such that:
  - If  $\zeta < \bar{\zeta}_{\tau,t}$ , there is no rationing, so:  $y_{\zeta,\tau,t} = \left(\frac{D p_\tau}{\zeta P_t}\right)^{-\epsilon} Y_t$ .
  - If  $\zeta > \bar{\zeta}_{\tau,t}$ , there is rationing, so:  $y_{\zeta,\tau,t} = \left(\frac{p_\tau}{P_t} \frac{1-\alpha}{W_t/A_t}\right)^{\frac{1-\alpha}{\alpha}}$ .
  - $\bar{\psi}_{\zeta,\tau,t} = \min \left\{ 1, \left(\frac{\bar{\zeta}_{\tau,t}}{\zeta}\right)^\epsilon \right\}$ . High  $\bar{\zeta}_{\tau,t}$ , less likely to be rationed.

# Plotting the short-run Phillips curve (true price/output)



- Assume:  $P_t = \exp(\pi t)$  for  $t < 0$ .
- And:  $P_t = P_0 \exp(\pi t)$  for  $t \geq 0$ .
- So, prices jump at time 0.
- Graphs plot possible  $(Y_0, P_0)$  (% , relative to s.s.).
- Solid line is short-run PC allowing rationing.
- Dashed line is short-run PC without rationing.
- Independent of price setting! From aggregation.
- Full calibration will be given shortly.

# Price change opportunity arrival rate choice

- If long-run inflation were higher, then prices would be changed more frequently.
  - Aggregate state dependence is necessary for reasonable comparative static results.
  - I broadly follow Blanco et al. (2024) in modelling an endogenous rate of price change opportunities.
- All firms are owned by conglomerates. Each conglomerate owns countably many firms.
- Each conglomerate chooses the price adjustment rate  $\lambda_t$  for the firms it owns (the same rate for all firms).
  - The conglomerate maximizes its firms' total profit, minus a cost of  $\frac{\kappa_1}{1+\kappa_2} (\max\{0, \lambda_t - \underline{\lambda}\})^{1+\kappa_2}$  labour units.
- The conglomerate cannot control which particular firms update at any point in time, only the total quantity.
  - Surprisingly consistent with price micro data, which finds hazard rates are flat in price age (Klenow & Malin 2010).
- Optimal:
  - MC of increasing  $\lambda_t$  = expected profits of firm with new price (over price life) – expected profits of current firms (over price life).

# Households and monetary policy

- In period  $\tau$ , households maximize:  $\int_{\tau}^{\infty} e^{-\int_{\tau}^t \rho_v dv} \left[ \log Y_t - \Psi_t \frac{1}{1+\nu} \left( L_t + \frac{\kappa_1}{1+\kappa_2} (\lambda_t - \underline{\lambda})^{1+\kappa_2} \right)^{1+\nu} \right] dt$ .
- They face the budget constraint:  $Y_t + \frac{\dot{B}_t^{(i)}}{P_t} + \dot{B}_t^{(r)} = W_t \left( L_t + \frac{\kappa_1}{1+\kappa_2} (\lambda_t - \underline{\lambda})^{1+\kappa_2} \right) + i_t \frac{B_t^{(i)}}{P_t} + r_t B_t^{(r)} + T_t$ .
  - $B_t^{(i)}$  nominal bonds.  $B_t^{(r)}$  real bonds.  $Y_t$  output = consumption, at price  $P_t$ .  $W_t$  wage.  $L_t$  labour.  $T_t$  profits from owning firms.
- FOCs imply  $\Psi_t \left( L_t + \frac{\kappa_1}{1+\kappa_2} (\lambda_t - \underline{\lambda})^{1+\kappa_2} \right)^{\nu} = \frac{W_t}{Y_t}$ ,  $r_t = \rho_t + \frac{\dot{Y}_t}{Y_t}$ ,  $i_t = r_t + \pi_t$ , where  $\pi_t = \frac{\dot{P}_t}{P_t}$ .
- Monetary policy sets  $i_t = r_t + \pi_t^* + \phi(\pi_t - \pi_t^*)$  with  $\phi > 1$  and  $\pi_t^*$  an exogenous target (Holden 2024).
- From Fisher equation,  $r_t + \pi_t = i_t = r_t + \pi_t^* + \phi(\pi_t - \pi_t^*)$ , so  $\pi_t = \pi_t^*$  for all  $t$ . Inflation is effectively exogenous.

# Other aggregates

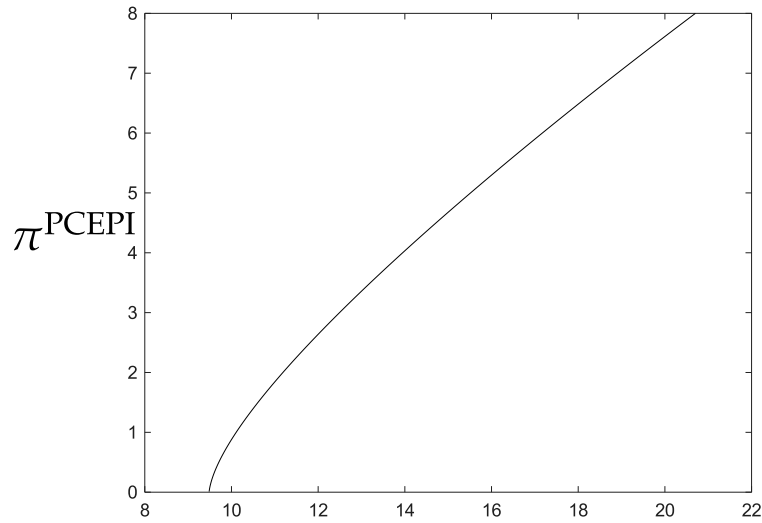
- Average probability that a buyer from a particular firm receives order:  $\bar{\psi}_{\tau,t} := \int_0^1 \bar{\psi}_{\zeta,\tau,t} g(\zeta) d\zeta = \frac{\theta \bar{\zeta}_{\tau,t}^\epsilon - \epsilon \bar{\zeta}_{\tau,t}^\theta}{\theta - \epsilon}$ .
- Average probability of receiving order across all firms:  $\bar{\psi}_t := \int_{-\infty}^t \lambda_\tau e^{-\int_\tau^t \lambda_v dv} \bar{\psi}_{\tau,t} d\tau$ .
  - Equal weighted for comparability with Cavallo & Kryvtsov (2023) evidence.
- Model PCEPI, following BLS imputation procedure (unobserved price changes assumed equal average change):
  - $\frac{1}{\Delta} (\log P_t^{\text{PCEPI}} - \log P_{t-\Delta}^{\text{PCEPI}}) = \frac{1}{\Delta} \int_{-\infty}^{t-\Delta} \lambda_\tau e^{-\int_\tau^{t-\Delta} \lambda_v dv} \left[ \bar{\psi}_{\tau,t-\Delta} \lambda_t \Delta \left[ \bar{\psi}_{t,t} \log \frac{p_t}{p_\tau} + (1 - \bar{\psi}_{t,t}) \log \frac{P_t^{\text{PCEPI}}}{P_{t-\Delta}^{\text{PCEPI}}} \right] + \bar{\psi}_{\tau,t-\Delta} (1 - \lambda_t \Delta) \left[ \bar{\psi}_{\tau,t} 0 + (1 - \bar{\psi}_{\tau,t}) \log \frac{P_t^{\text{PCEPI}}}{P_{t-\Delta}^{\text{PCEPI}}} \right] + (1 - \bar{\psi}_{\tau,t-\Delta}) \log \frac{P_t^{\text{PCEPI}}}{P_{t-\Delta}^{\text{PCEPI}}} \right] d\tau$
  - So:  $\pi_t^{\text{PCEPI}} := \frac{d \log P_t^{\text{PCEPI}}}{dt} = \lambda_t \bar{\psi}_{t,t} \frac{\int_{-\infty}^t \lambda_\tau e^{-\int_\tau^t \lambda_v dv} \bar{\psi}_{\tau,t}^2 \frac{1}{\bar{\psi}_{\tau,t}} \log \frac{p_t}{p_\tau} d\tau}{\int_{-\infty}^t \lambda_\tau e^{-\int_\tau^t \lambda_v dv} \bar{\psi}_{\tau,t}^2 d\tau}$ .

# Parameterization / calibration

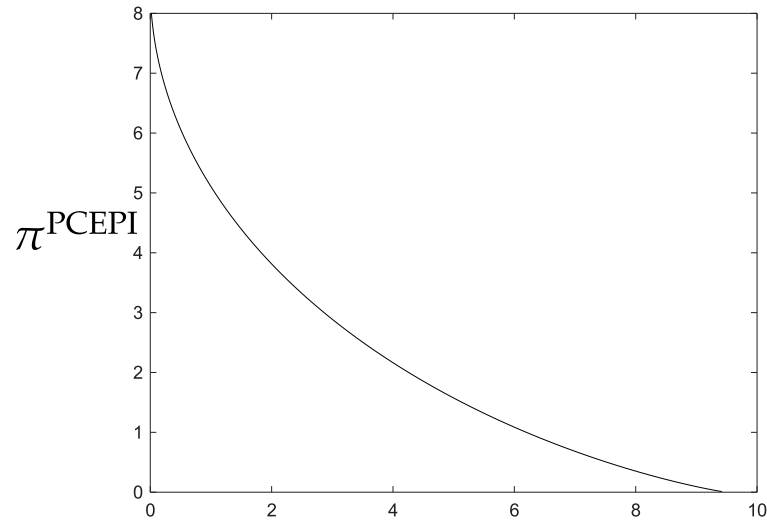
- $\rho = 2\%$ .  $\pi^{\text{PCEPI}} := 2\%$  unless otherwise stated. With rationing, requires  $\pi^* := 2.04\%$ , without  $\pi^* := 2.00\%$ .
- $\epsilon := 10$ ,  $\nu := 2$ , Smets & Wouters (2007).  $\alpha := \frac{3}{5}$ , Abraham et al. (2024).
- $\theta := 27$ . Matching 11% stockouts in 2019 from Cavallo & Kryvtsov (2023) to  $1 - \bar{\psi}$ .  $\zeta$  mean 0.96.  $\zeta$  s.d. 0.03.
- $\underline{\lambda} := 0.73$ , minimum annual median price adjustment rate in the Montag & Villar (2025) data.
- With rationing,  $\kappa_1 = 0.016$  and  $\kappa_2 = 3.75$ , without rationing  $\kappa_1 = 0.105$  and  $\kappa_2 = 2.06$ . Imply  $\lambda = 1.48$ .
  - Matching time series mean of the median rate of price adjustment from the Montag & Villar (2025) data,  $\lambda = 1.48$ .
  - And matching  $\int_0^{\frac{1}{4}} (\lambda_t - \lambda) dt / \int_0^{\frac{1}{4}} (\pi_t^{\text{PCEPI}} - \pi^{\text{PCEPI}}) dt$  following a monetary policy shock to 8.2 as estimated from Figure 1.
  - With rationing: 0.1% of labour is used for price adjustment. Without: 2.0% of labour is used for price adjustment.
  - Rationing reduces the price adjustment frictions needed to match the data!

# Results

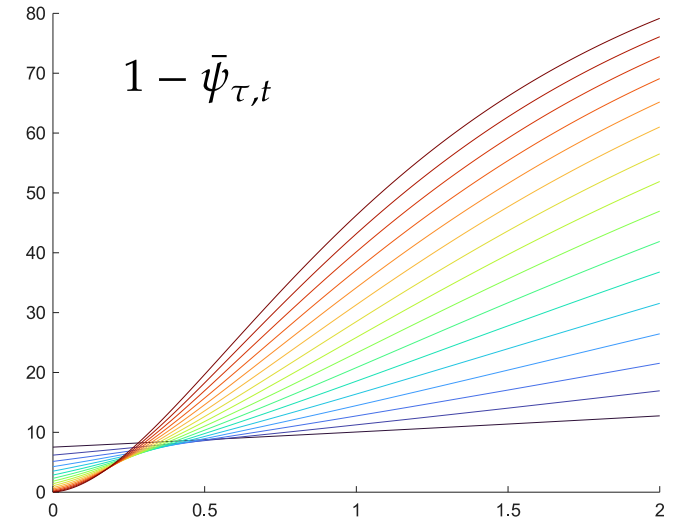
# Stockouts and rationing as a function of inflation



Average stockout level,  $1 - \bar{\psi}$  (percent).



Stockout rate at firms with new prices,  $1 - \bar{\psi}_{t,t}$  (percent).



Stockout levels (percent) as a function of price age (years), with varying steady state inflation levels.

Dark blue corresponds to 0.5% inflation.

Dark red corresponds to 8% inflation.

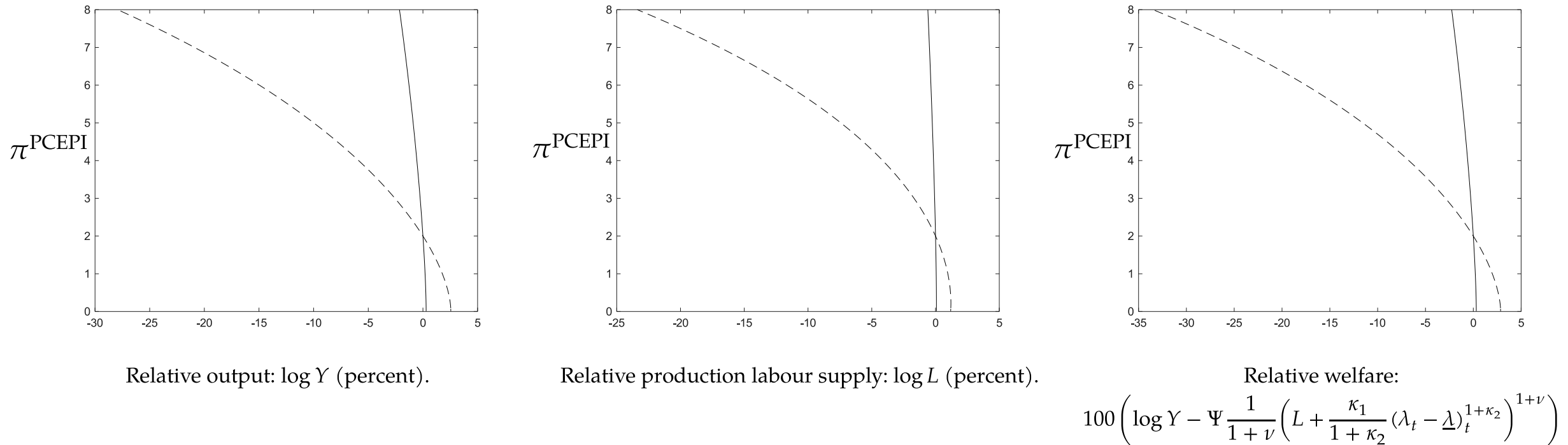
Figure 6: Stockouts and rationing as a function of PCEPI inflation (percent).

When inflation is high rationing is high. High inflation quickly erodes mark-ups.

Firms with new prices ration less with high inflation as they set high initial mark-ups.



# Output and welfare as a function of inflation



**Figure 7: Output and welfare as a function of inflation (percent).**

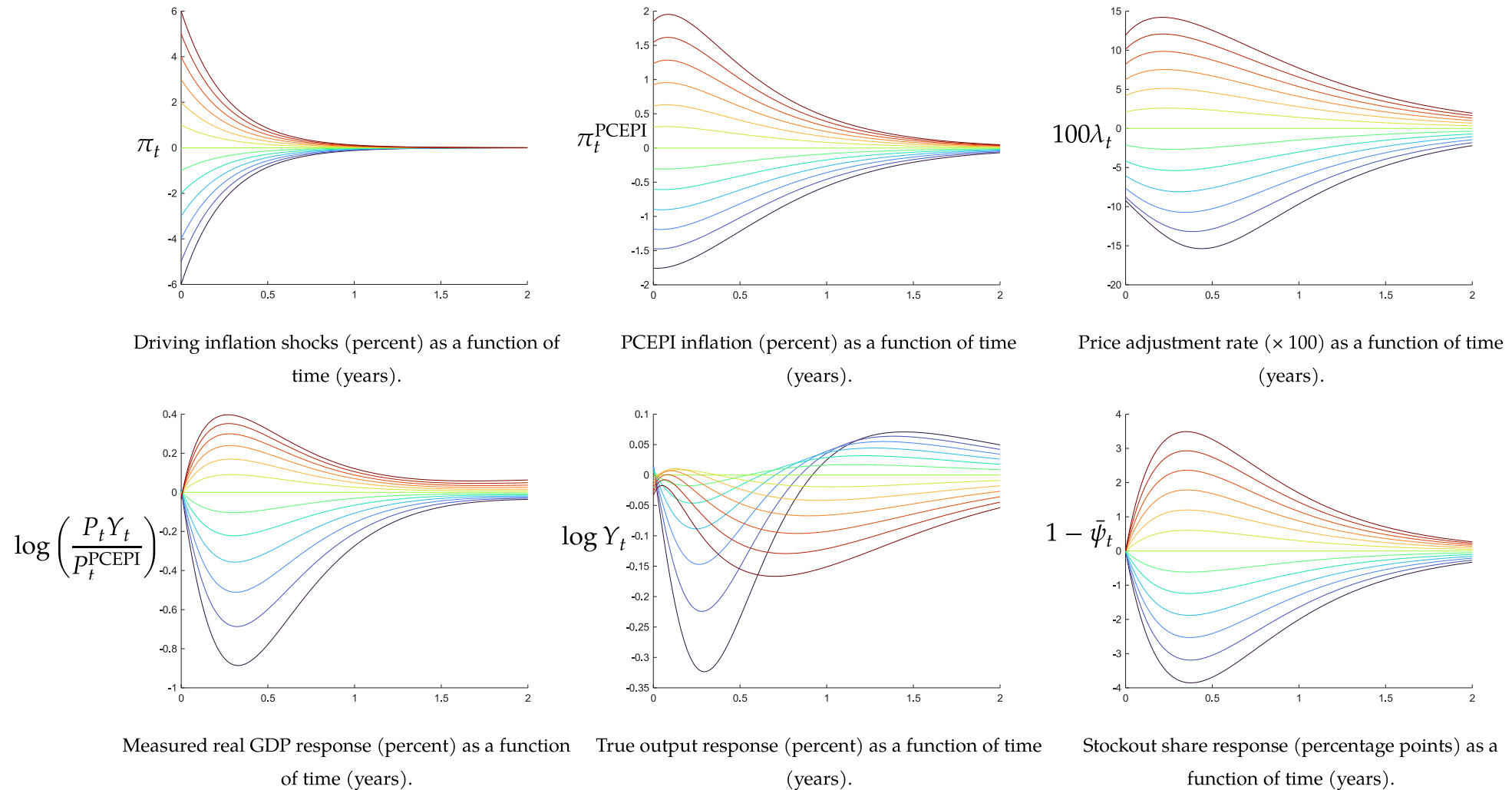
Black solid lines are the model with rationing. Black dashed lines are the model without rationing.

All plots are normalized to hit 0% on the horizontal axis when  $\pi^{\text{PCEPI}} = 2\%$ .

Welfare costs of inflation are much lower under rationing.

Firms with old prices are no longer making losses on every unit sold. Less labour is used in price adjustment.

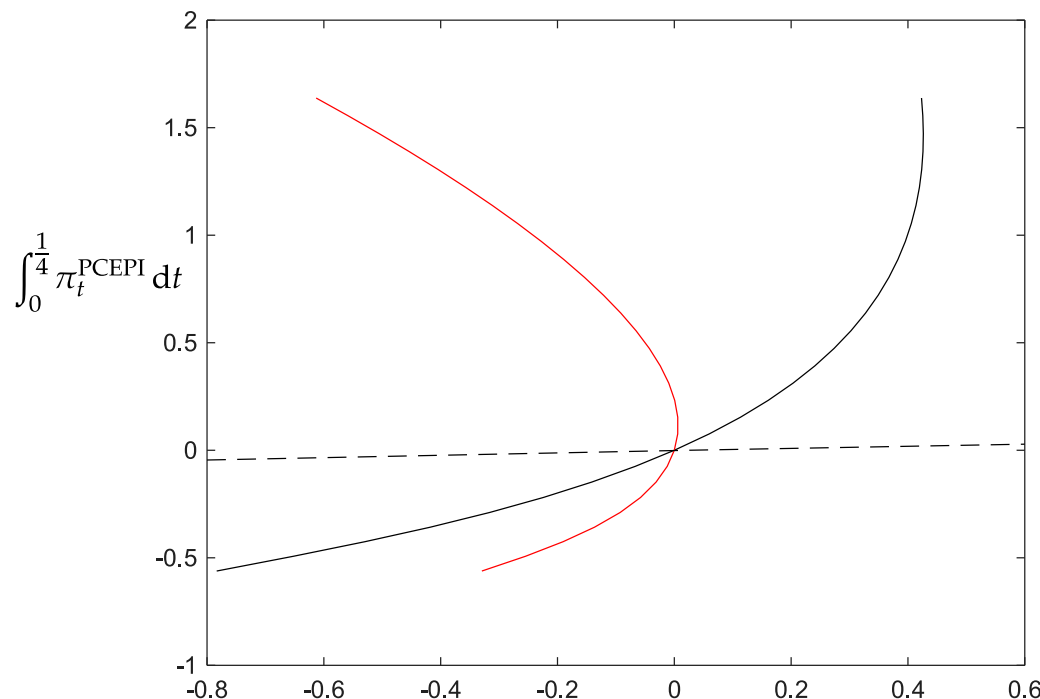
# IRFs to $\pi^*$ shocks (persistence matched to VAR)



**Figure 10: Impulse responses to monetary shocks, with rationing.**

Colours are consistent across subplots. All responses are relative to the no-shock counterfactual.

# The three-month Phillips curve



- The black line (measured output with rationing) matches the PC slope of 1.2 derived from applying the same calculations to the VAR IRFs I showed previously.

- Note that expansionary monetary policy reduces true output but increases measured output.

Black solid line: measured cumulated real GDP,  $\log \int_0^{\frac{1}{4}} \frac{P_t Y_t}{P_t^{\text{PCEPI}}} dt$ , with rationing.

Red solid line: cumulated true output,  $\log \int_0^{\frac{1}{4}} Y_t dt$ , with rationing.

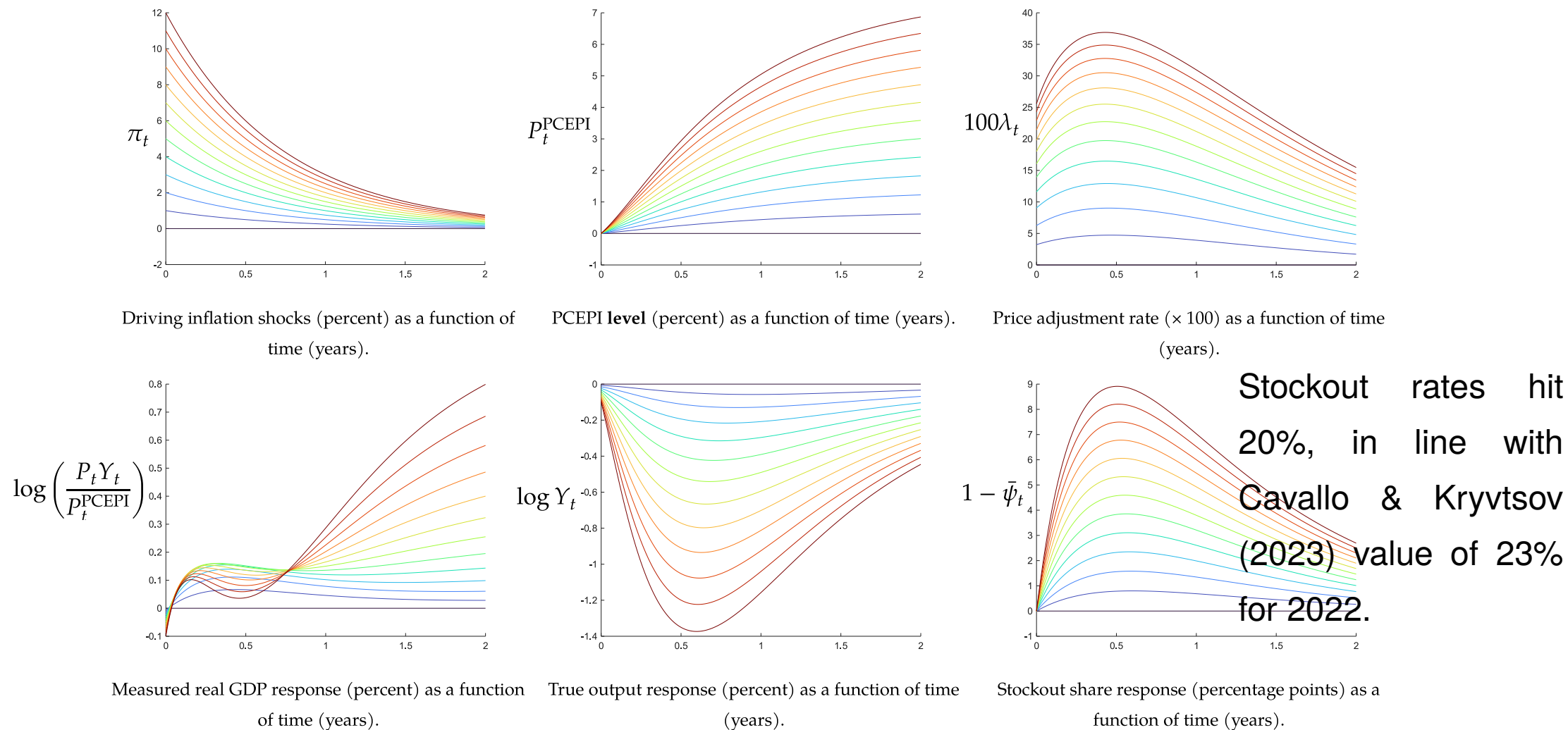
Black dashed line: measured cumulated real GDP,  $\log \int_0^{\frac{1}{4}} \frac{P_t Y_t}{P_t^{\text{PCEPI}}} dt$ , without rationing.

- The PCEPI deflator does not capture gains from variety, so it misses the losses coming from rationing reducing the measure of consumed varieties.

**Figure 9: The three-month Phillips curve with (solid lines) and without (dashed lines) rationing.**

All variables in percent. All variables are relative to the no-shock counterfactual.

# IRFs to $\pi^*$ shocks (matching 7% post-Covid inflation)



**Figure 11: Impulse responses to more persistent monetary shocks, with rationing.**

Colours are consistent across subplots. All responses are relative to the no-shock counterfactual.

# Conclusion

- The standard assumption that firms always satisfy all demand is **not innocuous**.
- Allowing rationing produces a model that **fits the data** better and performs more reasonably in extreme conditions.
- Allowing rationing drastically **reduces the welfare costs** of positive trend inflation.
- But when rationing is allowed, **true output declines** following “expansionary” monetary shocks.
- Monetary policy may be **less useful for stabilisation** than we previously thought.
- Extensions in final paper: quantity-capped rationing, consumer distaste for rationed varieties, firm specific capital, partially fixed intermediaries, long-run growth.
- I am interested to hear thoughts on other essential extensions, or crucial empirical results to establish.

Extra slides

# Random rationing vs sales-capped rationing

- Two potential models of rationing:
  - Sales-capped rationing: The firm caps the quantity it sells to any given consumer.
  - Random rationing: Some consumers get lucky and buy their entire order. Others go home with nothing.
- While we saw some sales caps during Covid, random rationing seems more natural.
  - It also fits the aggregate data better. (Sales-capped rationing generates an excessively steep Phillips curve.)
  - While with tight consumer storage constraints, if goods are semi-durable and consumers shop frequently, the result of random rationing can look a lot like quantity-capped rationing, without such constraints the result is random rationing again.
- With random rationing, changes in rationing change measure of varieties consumed. Consumers love variety!

# State variables and the short-run Phillips curve

- All of the model's state variables are of the form:  $X_{j,t} := \int_{-\infty}^t \lambda_{\tau} e^{-\int_{\tau}^t \lambda_v dv} p_{\tau}^{\chi_j} d\tau$ , so:  $\dot{X}_{j,t} = \lambda_t [p_t^{\chi_j} - X_{j,t}]$  (for  $j \in \mathbb{Z}$ ).
- The definition of total labour demand is:  $L_t := \int_{-\infty}^t \lambda_{\tau} e^{-\int_{\tau}^t \lambda_v dv} \int_0^1 l_{\zeta,\tau,t} g(\zeta) d\zeta d\tau$ .
- This implies an equilibrium condition relating  $L_t, A_t, \widehat{W}_t, Y_t, P_t, X_{1,t}$  &  $X_{2,t}$ , with  $\chi_1 := \theta + \frac{1}{\alpha} + \frac{\theta}{\epsilon} \frac{1-\alpha}{\alpha}$ ,  $\chi_2 := \frac{1}{\alpha}$ .
- The definition of aggregate output implies an equilibrium condition relating  $\widehat{W}_t, Y_t, P_t, X_{1,t}$  &  $X_{2,t}$ .
- Combined with the household labour FOC, these two equations give a short-run Phillips curve, holding states fixed.
- Without rationing, the equivalent first equation relates  $L_t, A_t, \widehat{W}_t, Y_t, P_t$  &  $X_{-1,t}$  with  $\chi_{-1} := -\frac{\epsilon}{1-\alpha}$ .
- And the second just relates  $P_t$  &  $X_{-2,t}$  with  $\chi_{-2} := -(\epsilon - 1)$ .
  - Thus, if  $X_{-2,t}$  is fixed,  $P_t$  is fixed. The short-run Phillips curve is horizontal in the NK model without rationing!



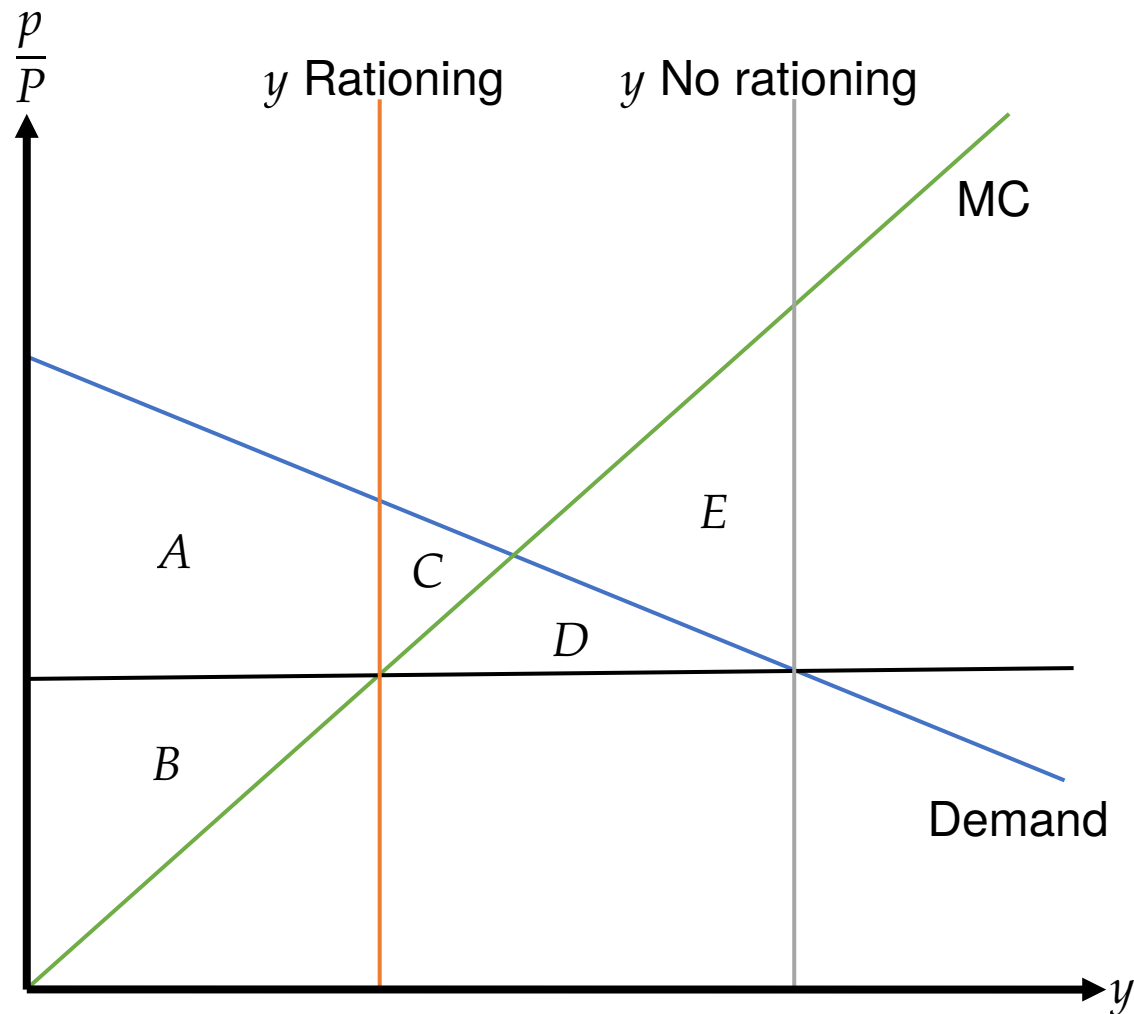
# Instability without rationing

- I stationarize  $X_{j,t}$  by defining  $\hat{X}_{j,t} := \frac{X_{j,t}}{P_t^{\chi_j}}$ . And I define:  $\hat{p}_t := \frac{p_t}{P_t}$ . Then:  $\dot{\hat{X}}_{j,t} = \lambda_t \hat{p}_t^{\chi_j} - (\lambda_t + \chi_j \pi_t) \hat{X}_{j,t}$ .
- So:  $\lambda_t + \chi_j \pi_t$  determines the stability of  $\hat{X}_{j,t}$ . It is stable if  $\lambda_t + \chi_j \pi_t > 0$ .
- For the model with rationing,  $\chi_1 = \theta + \frac{1}{\alpha} + \frac{\theta}{\epsilon} \frac{1-\alpha}{\alpha} > 0$  and  $\chi_2 = \frac{1}{\alpha} > 0$ . Stability guaranteed!
- For the model without rationing,  $\chi_{-1} = -\frac{\epsilon}{1-\alpha} < 0$  and  $\chi_{-2} = -(\epsilon - 1) < 0$ .
- If  $\epsilon$ ,  $\alpha$  or  $\pi_t$  are large enough, then  $\lambda_t + \chi_{-1} \pi_t < 0$  or  $\lambda_t + \chi_{-2} \pi_t < 0$ . Potential instability!

# New prices

- For  $j \in \mathbb{N}$ , define:  $z_{j,\tau} := \int_{\tau}^{\infty} e^{-\int_{\tau}^t (\lambda_v + r_v) dv} D^{\omega_{j,1}} \widehat{W}_t^{\omega_{j,2}} Y_t^{\omega_{j,3}} P_t^{\omega_{j,4}} dt$ , so  $\dot{z}_{j,\tau} = -D^{\omega_{j,1}} \widehat{W}_t^{\omega_{j,2}} Y_{\tau}^{\omega_{j,3}} P_{\tau}^{\omega_{j,4}} + (\lambda_{\tau} + r_{\tau}) z_{j,\tau}$ .
- Allowing rationing, updating firms optimally set:  $p_{\tau}^{\theta + \frac{\theta(1-\alpha)}{\epsilon}} \propto \frac{z_{2,\tau}}{z_{1,\tau}}$ .
  - Where:  $\omega_{1,1} := \theta$ ,  $\omega_{1,2} := -\frac{\theta+\epsilon}{\epsilon} \frac{1-\alpha}{\alpha}$ ,  $\omega_{1,3} := -\frac{\theta}{\epsilon}$ ,  $\omega_{1,4} := -\chi_1$ ,  $\omega_{2,1} := 0$ ,  $\omega_{2,2} := -\frac{1-\alpha}{\alpha}$ ,  $\omega_{2,3} := 0$ ,  $\omega_{2,4} := -\chi_2$ .
- Without rationing, updating firms optimally set:  $p_{\tau}^{1+\epsilon \frac{\alpha}{1-\alpha}} \propto \frac{z_{-2,\tau}}{z_{-1,\tau}}$ .
  - Where:  $\omega_{-1,1} := 0$ ,  $\omega_{-1,2} := -\epsilon$ ,  $\omega_{-1,3} := 1$ ,  $\omega_{-1,4} := -\chi_{-2}$ ,  $\omega_{-2,1} := -\frac{\epsilon}{1-\alpha}$ ,  $\omega_{-2,2} := 1$ ,  $\omega_{-2,3} := \frac{1}{1-\alpha}$ ,  $\omega_{-2,4} := -\chi_{-1}$ .
- I stationarize by defining:  $\hat{z}_{j,t} := \frac{z_{j,t}}{P_t^{\omega_{j,4}}}$ . Again, with rationing the  $\hat{z}_{j,t}$  are (backwards) stable, but not without.

# The microeconomics of rationing vs excess production



- Without rationing: CS is  $A + C + D$ . PS is  $B - D - E$ .
- Without rationing: Welfare is  $A + B + C - E$ .
- With rationing: CS is  $A$ . PS is  $B$ . Welfare is  $A + B$ .
- Welfare is higher with rationing when  $E > C$ .
- Plausible as demand ( $\propto y^{-\frac{1}{\epsilon}}$ ) is flatter than MC ( $\propto y^{\frac{\alpha}{1-\alpha}}$ ).
- The economy with rationing should be less distorted!

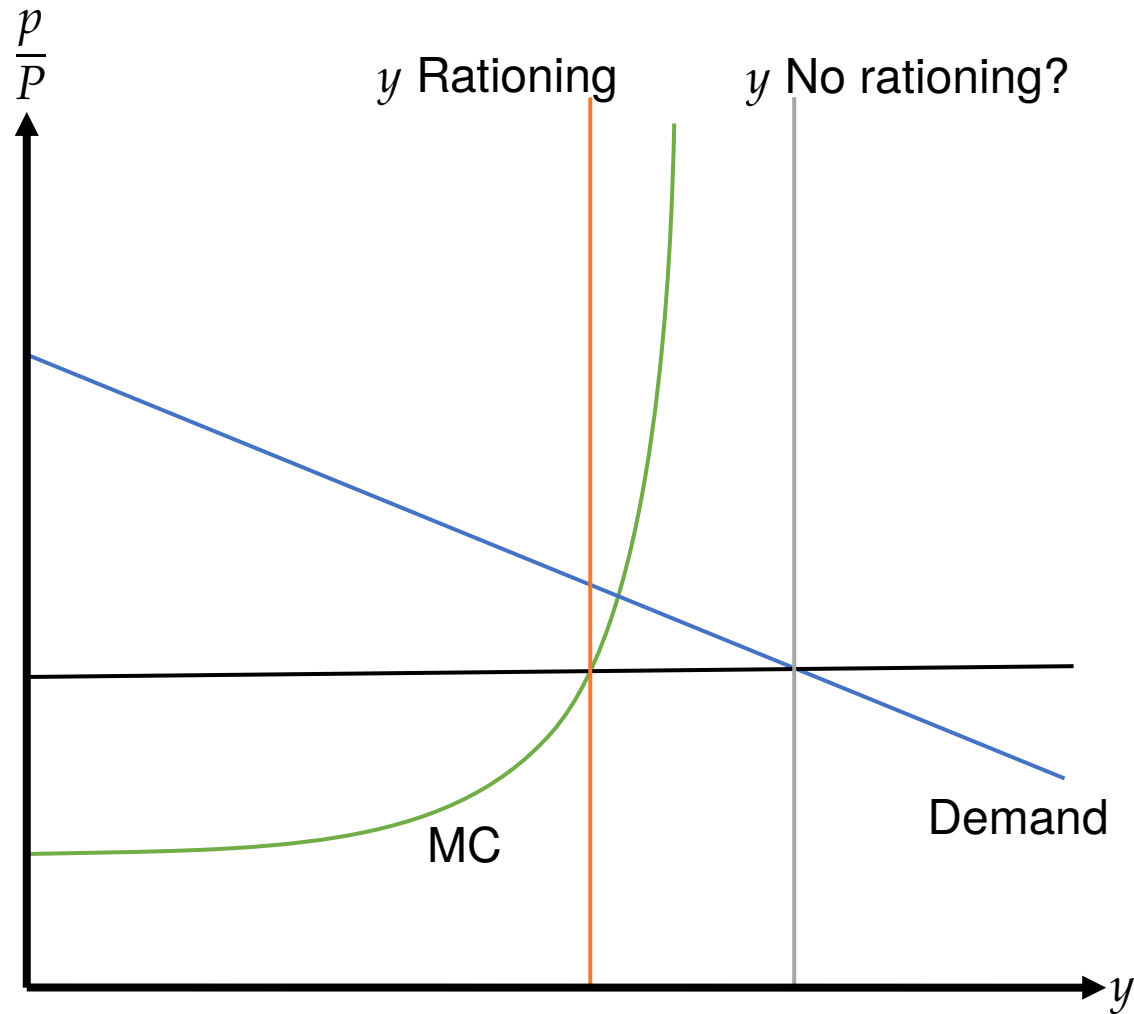
# The macroeconomics of rationing vs excess production

- If too much is produced by some firms (with old prices), other firms face higher marginal costs, so produce less.
  - Demand is shifted from undistorted firms (with new prices) to distorted ones (with old prices).
  - Bad!
- 
- If demand is rationed for some goods (with old prices), other firms face lower marginal costs, so produce more.
  - Demand is shifted from distorted firms (with old prices) to undistorted ones (with new prices).
  - Good!

# Strange properties of the Calvo model

- The Calvo model has some deeply strange properties (Holden, Marsal & Rabitsch 2024).
  - It implies a hard upper bound on steady-state inflation. With standard parameters, this is 5% to 10%.
  - Inflation above this level reduces the output *growth rate* not just the output level, due to ever growing price dispersion.
  - Under standard monetary rules, temporary high inflation can push the economy to this growing price dispersion path.
- These strange properties are tightly linked to the losses made by firms forced to sell at prices below marginal cost.
- When rationing is allowed, these strange properties disappear. Additional motivation for looking at it.

# Sticky prices with near vertical marginal costs

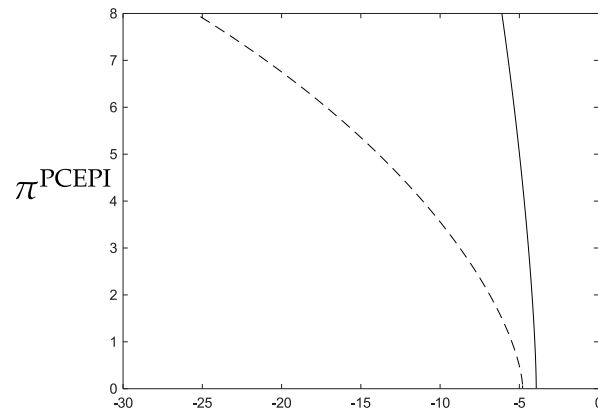


- Holding macro quantities fixed, there is no way to have equilibrium in this (micro) market without rationing.
- Total cost to produce “ $y$  No rationing” is infinite.
- So, what happens?
- As we climb the green line, an ever-increasing share of the economy’s productive capacity goes to this market.
- $\Rightarrow$  Aggregate output falls.
- $\Rightarrow$  Lower demand for this good at any price.
- Macro quantities move to clear this micro market!
- Is this really plausible???
- Rationing seems more reasonable.

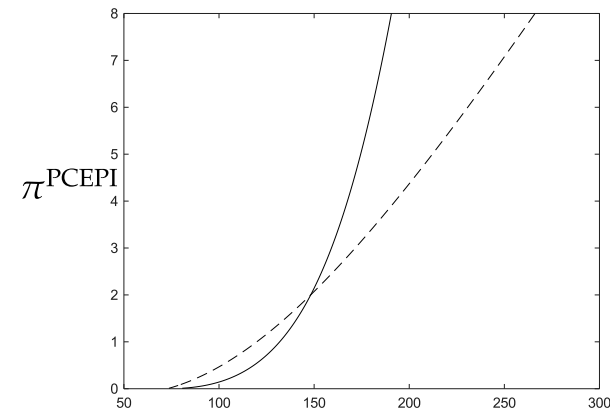
# The quasi flexible and fully flexible price cases

- The limit as  $\lambda_t \rightarrow \infty$  is not fully flexible prices, as for any  $\lambda_t$ , firms face all possible  $\zeta$  before changing price.
- Instead, the limit is quasi flexible prices, which maximize  $o_{\tau,t} := \int_0^1 o_{\zeta,\tau,t} g(\zeta) d\zeta$ .
- If  $\frac{\epsilon}{\epsilon-1} \frac{\theta+\epsilon}{\theta+\frac{\epsilon}{1-\alpha}} \leq 1$  then even quasi-flex-price firms ration with positive probability (for all  $t$ ), meaning  $\bar{\zeta}_{\tau,t} < 1$ .
  - This condition will hold in my calibration. It would be violated if  $\alpha$  was very small, or  $\theta$  was very large.
- A hypothetical fully flexible price firm would choose its price to maximize  $o_{\zeta,\tau,t}$ .
- Optimal choice is:  $\left(\frac{p_{\zeta,\tau,t}}{P_t}\right)^{1+\epsilon\frac{\alpha}{1-\alpha}} = \frac{\epsilon}{\epsilon-1} \left(\zeta \frac{\theta+1}{\theta}\right)^{\epsilon\frac{\alpha}{1-\alpha}} \frac{\widehat{W}_t}{1-\alpha} Y_t^{\frac{\alpha}{1-\alpha}}$ .
- Note that this is increasing in  $\zeta$ , while the price of a sticky or quasi-flex-price firm is not increasing in  $\zeta$ .
- Rationing reduces quantities for high  $\zeta$ , like in the fully flex price case!

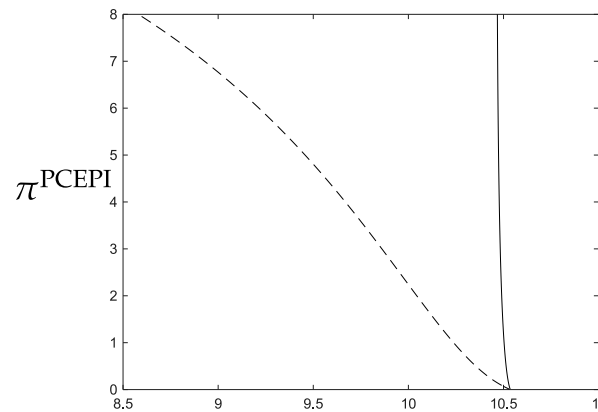
# Other consequences of varying inflation



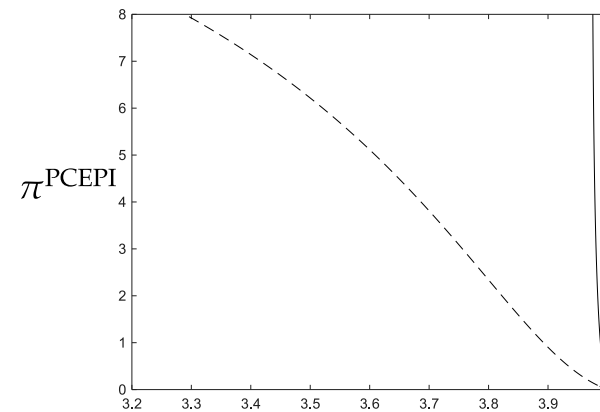
TFP loss from efficient benchmark:  $\log \frac{Y}{Y^{\text{SF}}}$  (percent).



Price adjustment rate (percent):  $100\lambda$ .



Aggregate mark-ups:  $\log \frac{(1-\alpha)Y}{W/L}$  (percent).



Excess firm profits shares:  $\frac{Q}{Y} - \alpha$  (percent).

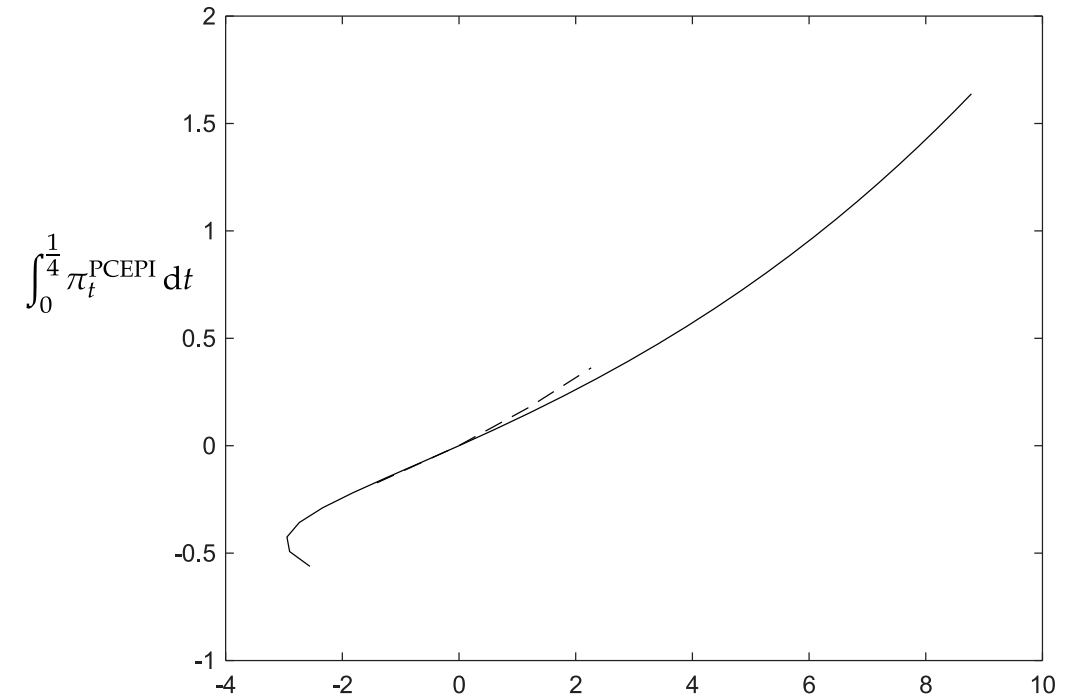
**Figure 8: Other consequences of varying inflation (percent).**

Black solid lines are the model with rationing. Black dashed lines are the model without rationing.



# The three-month adjustment rate Phillips curve

- The slope of these lines was a calibration target. (It matches the slope derived from applying the same calculations to the VAR IRFs I showed previously.)
- Small contractionary shocks reduce price adjustment rates, as real prices will be eroded by trend inflation anyway.
- Following large contractionary shocks though, price adjustment rates increase as firms benefit from cutting prices.



Black solid line: cumulated price adjustment rate,  $100 \int_0^{\frac{1}{4}} \lambda_t dt$ , with rationing.

Black dashed line: cumulated price adjustment rate,  $100 \int_0^{\frac{1}{4}} \lambda_t dt$ , without rationing.

**Figure 9: The three-month Phillips curve with (solid lines) and without (dashed lines) rationing.**

All variables in percent. All variables are relative to the no-shock counterfactual.

# References

- Abraham, Filip, Yannick Bormans, Jozef Konings & Werner Roeger. 2024. 'Price-Cost Margins, Fixed Costs and Excess Profits'. *The Economic Journal*: ueae037.
- Alessandria, George, Joseph P. Kaboski & Virgiliu Midrigan. 2010. 'Inventories, Lumpy Trade, and Large Devaluations'. *American Economic Review* 100 (5): 2304–2339.
- Barro, Robert J. 1977. 'Long-Term Contracting, Sticky Prices, and Monetary Policy'. *Journal of Monetary Economics* 3 (3): 305–316.
- Bauer, Michael D. & Eric T. Swanson. 2023. 'A Reassessment of Monetary Policy Surprises and High-Frequency Identification'. *NBER Macroeconomics Annual* 37: 87–155.
- Bils, Mark. 2016. 'Deducing Markups from Stockout Behavior'. *Research in Economics* 70 (2): 320–331.
- Blanco, Andrés, Corina Boar, Callum J. Jones & Virgiliu Midrigan. 2024. 'The Inflation Accelerator'. Working Paper. Working Paper Series. National Bureau of Economic Research.
- Brave, Scott A., R. Andrew Butters & David Kelley. 2019. 'A New “Big Data” Index of U.S. Economic Activity'. *Economic Perspectives* (1): 1–30.
- Brave, Scott A., Ross Cole & David Kelley. 2019. 'A “Big Data” View of the U.S. Economy: Introducing the Brave-Butters-Kelley Indexes'. *Chicago Fed Letter*.
- Caldara, Dario, Matteo Iacoviello & David Yu. 2025. 'Measuring Shortages since 1900'.

- Cavallo, Alberto & Oleksiy Kryvtsov. 2023. 'What Can Stockouts Tell Us about Inflation? Evidence from Online Micro Data'. *Journal of International Economics* 146. NBER International Seminar on Macroeconomics 2022: 103769.
- Corsetti, Giancarlo & Paolo Pesenti. 2005. 'International Dimensions of Optimal Monetary Policy'. *Journal of Monetary Economics* 52 (2): 281–305.
- Drèze, Jacques H. 1975. 'Existence of an Exchange Equilibrium under Price Rigidities'. *International Economic Review* 16 (2): 301–320.
- Forbes, Kristin J., Joseph E. Gagnon & Christopher G. Collins. 2022. 'Low Inflation Bends the Phillips Curve around the World'. *Economia* 45 (89): 52–72.
- Gerke, Rafael, Sebastian Giesen, Matija Lozej & Joost Röttger. 2023. 'On Household Labour Supply in Sticky-Wage HANK Models'. SSRN Scholarly Paper. Rochester, NY.
- Hahn, Volker. 2022. 'Price Dispersion and the Costs of Inflation'. *Journal of Money, Credit and Banking* 54 (2–3): 459–491.
- Holden, Tom D. 2024. *Robust Real Rate Rules*. Working Paper. Kiel, Hamburg: ZBW – Leibniz Information Centre for Economics.
- Holden, Tom D., Ales Marsal & Katrin Rabitsch. 2024. 'From Linear to Nonlinear: Rethinking Inflation Dynamics in the Calvo Pricing Mechanism'.
- Huo, Zhen & José-Víctor Ríos-Rull. 2020. 'Sticky Wage Models and Labor Supply Constraints'. *American Economic Journal: Macroeconomics* 12 (3): 284–318.

- Klenow, Peter J. & Oleksiy Kryvtsov. 2008. 'State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?'. *The Quarterly Journal of Economics* 123 (3): 863–904.
- Klenow, Peter J. & Benjamin A. Malin. 2010. 'Microeconomic Evidence on Price-Setting☆'. In *Handbook of Monetary Economics*, edited by Benjamin M. Friedman & Michael Woodford, 3:231–284. Elsevier.
- Kryvtsov, Oleksiy & Virgiliu Midrigan. 2013. 'Inventories, Markups, and Real Rigidities in Menu Cost Models'. *The Review of Economic Studies* 80 (1): 249–276.
- Miranda-Agrippino, Silvia & Giovanni Ricco. 2021. 'The Transmission of Monetary Policy Shocks'. *American Economic Journal: Macroeconomics* 13 (3): 74–107.
- Montag, Hugh & Daniel Villar. 2025. 'Post-Pandemic Price Flexibility in the U.S.: Evidence and Implications for Price Setting Models'.
- Nakamura, Emi & Jón Steinsson. 2008. 'Five Facts about Prices: A Reevaluation of Menu Cost Models\*'. *The Quarterly Journal of Economics* 123 (4): 1415–1464.
- Posch, Olaf. 2018. 'Resurrecting the New-Keynesian Model: (Un)Conventional Policy and the Taylor Rule'. *CESifo Working Paper Series*. CESifo Working Paper Series.
- Posch, Olaf, Juan F. Rubio-Ramírez & Jesús Fernández-Villaverde. 2011. 'Solving the New Keynesian Model in Continuous Time'. *2011 Meeting Papers*. 2011 Meeting Papers.
- Smets, Frank & Rafael Wouters. 2007. 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach'. *American Economic Review* 97 (3): 586–606.
- Svensson, Lars EO. 1984. *Sticky Goods Prices, Flexible Asset Prices, and Optimum Monetary Policy*. IIES.