Priors and the Slope of the Phillips Curve[*]

Callum Jones[†] Mariano Kulish[‡] Juan Pablo Nicolini[§]

February, 2021

Abstract

The slope of the Phillips curve in New Keynesian models is difficult to estimate using aggregate data. We show that, in a Bayesian estimation, the priors placed on the parameters governing nominal rigidities significantly influences posterior estimates and thus inferences about the importance of nominal rigidities. Conversely, we show that priors play a negligible role in a New Keynesian model estimated using state-level data. An estimation with state-level data exploits a relatively large panel dataset and removes the influence of endogenous monetary policy.

Keywords: Slope of the Phillips curve, priors, Bayesian estimation, state-level data.

JEL classifications: E52, E58.

[*]The views expressed herein are those of the authors and not necessarily those of the Board of Governors of the Federal Reserve System, the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

[†]Federal Reserve Board. callum.j.jones@frb.gov.

[‡]School of Economics, University of Sydney. mariano.kulish@sydney.edu.au.

[§]Federal Reserve Bank of Minneapolis and Universidad Di Tella. juanpa@minneapolisfed.org.

1 Introduction

The slope of the Phillips curve in New Keynesian models is notoriously difficult to estimate. For example, Ireland (2004) estimates a New Keynesian model by maximum likelihood, but calibrates the price adjustment cost parameter after attempts that led to implausible large nominal rigidities. Along the same line, Schorfheide (2008) states that "*the identification of the Phillips curve coefficients is tenuous and no consensus about its slope and the importance of lagged inflation has emerged from the empirical studies.*"

This problem of identification may be one of the reasons why the Bayesian approach to estimation became widespread in the DSGE literature. In part, through the use of priors, researchers were able to bring information from other sources. For example, in the case of the slope of the Phillips curve, data on the frequency of price adjustment was typically pointed to when setting priors on the Calvo parameters. But in part, priors add curvature to the posterior along dimensions where the likelihood may be very flat. So while the Bayesian approach has provided a constructive way to move forward in estimating DSGE models, issues of identification become blurred in practice. A natural way to partially address this problem is to explore the sensitivity of the results to the assumed priors. This is the purpose of this paper.

We follow the approach of Del Negro and Schorfheide (2008) to assess quantitatively the impact of priors on inferences of the Phillips Curve slopes. Del Negro and Schorfheide (2008) show that medium scale New Keynesian models estimated on post-1982 US aggregate data cannot discriminate between models that differ in the quantitative importance of nominal rigidities. As a result, they argue priors play a crucial role.

Our contribution is to estimate the same model, but using state-level data, and show that, in contrast, priors on nominal rigidities have a negligible impact on posterior inferences. A main advantage of using state-level data, as we argue in a companion paper, Fitzgerald et al. (2020), is that when monetary policy is endogenous relying on aggregate data to uncover the slope of the Phillips curve can be problematic. Relying on deviations of state-level data from the aggregate, however, circumvents the problem that arises with endogenous monetary policy.[1] A second advantage is that the panel nature of the state-level data brings in more observations.

We structure our analysis in two main steps. In our first step, that we call the *direct effect* of priors, we explore how the posterior distributions of the Calvo parameters depend on the priors

[1] With an estimated structural model, Fitzgerald et al. (2020) find that estimates of the slope of Phillips curve which are found to be unstable when estimated of aggregate data, are found to be stable when estimated of state-level data for the United States.

assumed for those parameters. However, this is not the only way in which priors can affect the posterior distributions of the Calvo parameters. In our second step, that we call the *indirect effect* of priors, we study how the posterior distributions of the Calvo parameters depend on the priors assumed for the policy parameters and for the parameters governing the distribution of the shocks, while keeping the priors on the Calvo parameters fixed. In considering this effect, using state-level data also has an advantage. The reason is that the model, in deviations from the aggregate, can be estimated independently of the policy rule, so there is no need to take a stand on the priors regarding the policy rule parameters.

Our main findings are these. In studying the direct effect, we show that, in line with the results in Del Negro and Schorfheide (2008), the posterior distributions for the Calvo parameters are very sensitive to the assumed prior for those parameters. In contrast, the posteriors of the same parameters, estimated using state-level data, are notably robust to the assumed priors. In studying the indirect effect, we find that with aggregate data, priors over the policy rule parameters and for the parameters governing the shocks, matter in general for posterior inferences on the Calvo parameters. There are, however, important differences in the results when we compare a small scale with a medium scale model. With state-level data, there is no indirect effect from priors on policy parameters, since the Calvo parameters can be estimated separately, so those priors are inconsequential. The state-level estimates are quite robust to changing priors on the parameters governing the shocks.

State-level data helps in identification in two non-mutually exclusive ways: first since it uses state-level variation to estimate the Calvo parameters, they can be estimated removing any interaction with endogenous policy. Second, because it brings in more information allowing the likelihood to dominate the prior in posterior inferences.

In the next section we discuss the model and our empirical strategy. Section 3 presents the results. Section 4 then sets up a simple example that shows analytically why moving to a state-level model helps to estimate the slope of the Phillips curve. The simple example shows that priors are potentially very powerful in models which are not identified in a frequentist sense. Priors can be thought to select among the infinitely many solutions. But the example also shows how inferences of the slope of the Phillips curve are intertwined with the policy rule parameters in an aggregate level model. The analytical results help interpret our findings with the quantitative estimated models. It is not meant to be an exhaustive analysis of identification of forward looking structural models. For further analysis on identification in structural models, see Canova and Sala (2009), Iskrev (2010), Komunjer and Ng (2011) and Müller (2012).

2 Estimation of a Small Scale New Keynesian model

In this section we estimate a small scale New Keynesian model using both aggregate and state level data and for three alternative specifications for the priors. In doing so, we study the robustness of our estimates to the priors used and, more importantly, how this robustness depends on the type of data used for estimation.

The model we estimate in this section is a variation of the standard three-equation New Keynesian model. We adapt that model to a series of geographically separated units in which local shocks can move local pricing and employment decisions that are different than those for the country as a whole. To do so, we do need to extend that basic popular model to allow for tradable and non-tradable goods. This is the only deviation from the standard textbook example of the New Keynesian model with price and wage frictions. We explain the model in detail below.

2.1 The Model

We use the model described in Fitzgerald et al. (2020). The economy consists of a continuum of ex ante identical islands which together form a monetary union and trade with one another. Consumers on each island derive utility from the consumption of a final good and from leisure:

$$\max \mathbb{E}_0 \sum_{t=0}^{\infty} \beta_t(s) \left[\log(c_t(s)) - \frac{\eta_t(s)}{1+\nu} n_t(s)^{1+\nu} \right],$$

where s indexes the island, $c_t(s)$ is consumption, $n_t(s)$ is labor supplied, $\beta_t(s)$ is a preference shock, and $\eta_t(s)$ is a labor disutility shock. Both shocks follow AR(1) processes:

$$\log \beta_t(s) = (1-\rho_\beta)\log \beta + \rho_\beta \log \beta_{t-1}(s) + \sigma_\beta \varepsilon_t^\beta(s)$$
$$\log \eta_t(s) = (1-\rho_\eta)\log \eta + \rho_\eta \log \eta_{t-1}(s) + \sigma_\eta \varepsilon_t^\eta(s)$$

The production technology is as follows. The final good $y_t(s)$ is assembled using inputs of non-traded goods $y_t^N(s)$ and traded goods $y_t^M(s,j)$ imported from island j:

$$y_t(s) = \left(\omega^{\frac{1}{\sigma}} y_t^N(s)^{\frac{\sigma-1}{\sigma}} + (1-\omega)^{\frac{1}{\sigma}} \left(\int_0^1 y_t^M(s,j)^{\frac{\kappa-1}{\kappa}} \mathrm{d}j \right)^{\frac{\kappa}{\kappa-1}\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

where ω determines the share of non-traded goods, σ is the elasticity of substitution between non-traded and traded goods, and κ is the elasticity of substitution across varieties of traded goods. Letting $p_t^N(s)$ and $p_t^M(s)$ denote the inputs' corresponding prices, the price of the final good on an island is

$$p_t(s) = \left(\omega p_t^N(s)^{1-\sigma} + (1-\omega) \left(\int_0^1 p_t^M(j)^{1-\kappa} \mathrm{d}j \right)^{\frac{1-\sigma}{1-\kappa}} \right)^{\frac{1}{1-\sigma}}.$$

Non-traded goods and traded export goods $y_t^X(s)$ on each island are CES composites of varieties k of differentiated intermediate inputs with an elasticity of substitution ϑ:

$$y_t^N(s) = \left(\int_0^1 y_t^N(s,k)^{\frac{\vartheta-1}{\vartheta}} \mathrm{d}k \right)^{\frac{\vartheta}{\vartheta-1}}$$
$$y_t^X(s) = \left(\int_0^1 y_t^X(s,k)^{\frac{\vartheta-1}{\vartheta}} \mathrm{d}k \right)^{\frac{\vartheta}{\vartheta-1}}.$$

The production of the varieties of non-traded goods and the varieties of traded exports on each island is linear in labor:

$$y_t^N(s,k) = z_t^N(s) n_t^N(s,k)$$
$$y_t^X(s,k) = z_t^X(s) n_t^X(s,k),$$

where $z_t^N(s)$ and $z_t^X(s)$ are productivity shocks that follow AR(1) processes:

$$\log z_t^N(s) = (1-\rho_z^N) \log z^N + \rho_z^N \log z_{t-1}^N(s) + \sigma_z^N \varepsilon_{z,t}^N(s)$$
$$\log z_t^X(s) = (1-\rho_z^X) \log z^X + \rho_z^X \log z_{t-1}^X(s) + \sigma_z^X \varepsilon_{z,t}^X(s).$$

Nominal frictions affect this economy in a standard way. Individual producers of tradable and non-tradable intermediate goods are subject to Calvo price adjustment frictions – parameterized by λ_p, the probability that a firm cannot reset its price in a given period. The evolution of non-tradeable prices thus evolves according to:

$$\hat{p}_t^N(s) = \lambda_p \hat{p}_{t-1}^N(s) + (1-\lambda_p) \hat{p}_t^{N\star}(s),$$

where variables with hats denote their log-deviation from steady-state and $\hat{p}_t^{N\star}(s)$ is the optimal price that firms set in the case they are able to adjust prices. This optimal price is forward-looking and solves:

$$\hat{p}_t^{N\star}(s) = \beta \lambda_p \mathbb{E}_t \hat{p}_{t+1}^{N\star}(s) + (1-\beta\lambda_p)(\hat{w}_t(s) - \hat{z}_t^N(s)),$$

where $\hat{w}_t(s)$ is nominal wages. These two equations yield the price Phillips curve for non-tradeable goods:

$$\hat{\pi}_t^N(s) = \beta \mathbb{E}_t \hat{\pi}_{t+1}^N(s) + \frac{(1-\beta\lambda_p)(1-\lambda_p)}{\lambda_p}\left(\hat{w}_t(s) - \hat{p}_t^N(s) - \hat{z}_t^N(s)\right),$$

so that the slope of the price Phillips curve in non-tradeable inflation is $\frac{(1-\beta\lambda_p)(1-\lambda_p)}{\lambda_p}$. An analogous argument applies to the slope of the Phillips curve in tradeable inflation.

Labor is immobile across states and is aggregated using a CES aggregator with an elasticity of substitution across labor varieties of ψ. Individual households supply differentiated varieties of labor that are subject to Calvo wage adjustment frictions – parameterized by λ_w, the probability that a labor variety cannot reset its wage in a given period. The evolution of wages evolves according to:

$$\hat{w}_t(s) = \lambda_w \hat{w}_{t-1}(s) + (1-\lambda_w)\hat{w}_t^\star(s),$$

where $\hat{w}_t^\star(s)$ is the optimal wage that unions set in the case they are able to adjust wages, which evolves according to:

$$\hat{w}_t^\star(s) = \beta\lambda_w \mathbb{E}_t \hat{w}_{t+1}^\star(s) + \frac{(1-\lambda_w\beta)}{(1+\psi\nu)}\left(-\hat{\mu}_t(s) + \psi\nu\hat{w}_t(s) + \log\eta_t(s) - \log\eta + \nu\hat{n}_t(s)\right),$$

where $\hat{\mu}_t$ is the shadow value of wealth.

Applying the same logic used in deriving the price inflation Phillips curve to this optimal wage setting equation, we can derive the following wage inflation Phillips curve:

$$\hat{\pi}_t^w = \beta\mathbb{E}_t\hat{\pi}_{t+1}^w + \frac{(1-\beta\lambda_w)(1-\lambda_w)}{\lambda_w}\frac{1}{(1+\psi\nu)}\left(-\hat{\mu}_t(s) - \hat{w}_t(s) + \log\eta_t(s) - \log\eta + \nu\hat{n}_t(s)\right).$$

Notice that under this specification, the slope of the wage inflation Phillips curve is proportional to $\frac{(1-\beta\lambda_w)(1-\lambda_w)}{\lambda_w}$, the analogous term to the price Phillips curve. This term highlights the non-linear relationship between the Calvo parameters that we estimate and the slopes of the corresponding Phillips curves. This non-linearity implies that 0.05 difference in the Calvo parameter implies much larger changes in the corresponding slope when the Calvo parameter is 0.90 than when it is 0.60, values that we estimate below.

At the aggregate level, monetary policy is set using a Taylor rule when the ZLB does not bind. The nominal interest rate i_t responds to its lag with weight α_i; deviations of inflation π_t from target $\bar{\pi}$ with weight α_π; deviations of output y_t from the flexible-price level of output y_t^F, with weight α_y; and the growth rate of the output gap with weight α_x:

$$1 + i_t = (1+i_{t-1})^{\alpha_i}\left[(1+\bar{i})\left(\frac{\pi_t}{\bar{\pi}}\right)^{\alpha_\pi}\left(\frac{y_t}{y_t^F}\right)^{\alpha_y}\right]^{1-\alpha_r}\left(\frac{y_t}{y_{t-1}}/\frac{y_t^F}{y_{t-1}^F}\right)^{\alpha_x}\exp(\varepsilon_t^i).$$

When the zero lower bound binds, the nominal interest rate equals zero.

Finally, at the aggregate level, in addition to shocks to the interest rate rule ε_t^i, we also include shocks to the aggregate price Phillips curve (via standard markup shocks), which follow an AR(1) process:

$$\log \xi_t = (1 - \rho_\xi) \log \xi + \rho_\xi \log \xi_{t-1} + \sigma_\xi \varepsilon_t^\xi.$$

2.2 Estimation Strategy

Method. To capture the period of zero nominal interest rates, we use a piecewise linear approximation as proposed in Jones (2017), Kulish, Morley and Robinson (2017), and Guerrieri and Iacoviello (2015). Under this approximation, the reduced form solution of our model has a time-varying VAR representation:

$$\mathbf{x}_t = \mathbf{J}_t + \mathbf{Q}_t \mathbf{x}_{t-1} + \mathbf{G}_t \epsilon_t,$$

where $\mathbf{x}_t$ collects the state and aggregate endogenous variables and ϵ_t collects the state and aggregate shocks. The time-varying coefficient matrices $\mathbf{J}_t$, $\mathbf{Q}_t$, and $\mathbf{G}_t$, arise because of the non-linearities induced by the ZLB.

Following Jones, Midrigan and Philippon (2018), we separate the state-level variables from the aggregate variables. We decompose the vector of variables for each island s, expressed in log-deviations from the steady state as $\mathbf{x}_t(s)$, into a component due to state s's dependence on its own history $\mathbf{x}_{t-1}(s)$ and its own shocks $\epsilon_t(s)$ and a component encoding the state-level dependence on aggregate variables:

$$\mathbf{x}_t(s) = \underbrace{\mathbf{Q}\mathbf{x}_{t-1}(s) + \mathbf{G}\epsilon_t(s)}_{\text{state-level component}} + \underbrace{\tilde{\mathbf{J}}_t + \tilde{\mathbf{Q}}_t \mathbf{x}^*_{t-1} + \tilde{\mathbf{G}}_t \epsilon^*_t}_{\text{aggregate component}}. \tag{1}$$

The coefficient matrices that appear in the aggregate component, $\tilde{\mathbf{J}}_t$, $\tilde{\mathbf{Q}}_t$, and $\tilde{\mathbf{G}}_t$, are time-varying because of the non-linearities induced by the ZLB. The vector $\mathbf{x}^*_t$ which contains the aggregate variables evolves as:

$$\mathbf{x}^*_t = \mathbf{J}^*_t + \mathbf{Q}^*_t \mathbf{x}^*_{t-1} + \mathbf{G}^*_t \epsilon^*_t. \tag{2}$$

Here, ϵ^*_t are the aggregate shocks.

Next, let $\bar{\mathbf{x}}^*_t = \int \mathbf{x}_t(s) \mathrm{d}s$ denote the economy-wide average of the island-level variables. The vector $\bar{\mathbf{x}}^*_t$ is a subset of the set of aggregate variables $\mathbf{x}^*_t$ (which, for example, can include the nominal interest rate). Given this structure of our model, the deviation of island-level variables from their economy-wide averages, $\hat{\mathbf{x}}_t(s) = \mathbf{x}_t(s) - \bar{\mathbf{x}}_t^*$, is a time-invariant function of island-level variables alone:

$$\hat{\mathbf{x}}_t(s) = \mathbf{Q}\hat{\mathbf{x}}_{t-1}(s) + \mathbf{G}\epsilon_t(s), \tag{3}$$

where we use the assumption that island-level shocks have zero mean in the aggregate, that is, $\int \epsilon_t(s)\mathrm{d}s = 0$. We make explicit also that a key assumption we make in (1) in order to arrive at (3) is that the parameters across states are the same (that is, that the coefficient matrices $\mathbf{Q}$ and $\mathbf{G}$ for the state-level components are not state-specific).

The use of deviations of state-level observables from aggregates in estimation removes the dependence of state-level outcomes on aggregate variables, so that the nominal interest rate drops out from the reduced form. Equation (3) therefore circumvents the problem of having to rely on aggregate data to estimate the Phillips curve in the presence of endogenous and possibly time-varying policy at the aggregate level, as discussed in detail in Fitzgerald et al. (2020).[2]

Practically, the use of equations (2) and (3) to estimate the model involves first expressing each state's observable variable as a deviation from its aggregate counterpart by subtracting time effects for each year and each variable. It also involves subtracting a state-specific fixed effect and time trend for each observable, since in the model, all islands are ex ante identical.

We estimate the model using state-level data, following the strategy just described. With the purpose of comparing results, we also estimate the model using aggregate data. In doing so, we jointly estimate the structural parameters and the policy rule.

In all cases, we use Bayesian methods to estimate the model's structural parameters. To construct the posterior distribution, as the island-level shocks in (3) are independent and do not affect aggregate outcomes, we can write the likelihood of the model as the product of each individual state's likelihood, computed from (3). When we estimate the model using aggregate data, we use equation (2) to compute the aggregate likelihood. Additional details on the likelihood function construction are provided in the Appendix.

Calibrated and Estimated Parameters. We calibrate a subset of parameters to the same values as in Jones, Midrigan and Philippon (2018). The discount factor is chosen to match an annual real interest rate of 2%. The elasticity of substitution of labor varieties, ψ, is calibrated to a value of 21, as in Christiano et al. (2005). The inverse labor supply elasticity ν is set to 2. At the

[2] Another advantage of representation (3) is that we can overcome the curse of dimensionality associated with all 51 states' dependence on the time-varying aggregate structure, which would otherwise make our estimation with state-level data computationally infeasible.

state-level, we follow the trade literature in setting the parameters governing the tradeable and non-tradeable sectors. The Appendix contains all details of the calibrated parameters.

The remaining parameters of the model are estimated. We reiterate that in our estimations using state-level data, there is no need to estimate the parameters of the Taylor Rule. This is because in our estimation strategy, the nominal interest rate drops out in equation (3).

We thus estimate the following parameters at the aggregate-level: those governing price stickiness λ_p; wage stickiness λ_w; the persistences and standard deviations of the aggregate shocks; and the parameters of the monetary policy rule α_i, α_π, α_y, and α_x. At the state-level, we estimate λ_p and λ_w, and the persistences and standard deviations of the state-level shocks.

Data For our benchmark exercise, we use United States data for the period 1977-2017. The year 1977 is the first one for which we have state-level data. This allows us to use the same period in both the aggregate data and the state-level data estimations, so any potential difference between the two cannot be attributed to the time period.[3]

We use aggregate data on employment, output, wages, inflation, and the Fed Funds rate.[4] To pin down expectations of the nominal interest rate during the ZLB period, we use the sequence of expected lower bound durations between 2009 and 2015 from the Blue Chip Financial Forecasts survey (for 2009 to 2010) and the New York Federal Reserve's Survey of Primary Dealers (for 2011 to 2015) (see Kulish, Morley and Robinson, 2017). At the state-level, we use a panel of employment, nominal output, wages, and inflation in the cross section of 51 US states. The state-level data forms an unbalanced mixed-frequency panel, as most variables are observed for all states at an annual frequency but inflation is observed at a biannual frequency for about half of the states.

In the appendix we show the results to be similar when using a sample size starting in 1966, which is the first year of the sample used to estimate the medium scale model in the next section. That is also the first year of the sample typically used in estimating medium scale models, as the one we analyze below. Our benchmark period includes the building up and bursting of the financial crisis, plus the ensuing years of policy at its effective zero bound. In the online appendix we present the results if one ends the sample in 2004 or if we allow for credit shocks, which, as documented in Mian and Sufi (2011), had a substantial heterogeneous impact across US regions. As we show there, the overall effects are barely changed by these modifications.

[3] For a throughout analysis of the stability of the Calvo parameters over time, see Fitzgerald et al. (2020).

[4] See Appendix E for details of data availability and how we construct our series.

3 Results

3.1 Using Aggregate Data

We first study the results using aggregate data and separate our discussion into, first, the direct effect of changing the priors for the Calvo parameters on the Calvo estimates, and second, the indirect effect of changing the priors for the other parameters on the Calvo estimates.

3.1.1 The Direct Effect

In Table 1 we show the results using aggregate data. We use three different priors. In all cases, we center the prior for both Calvo parameters at 0.5. We then allow for different confidence on that value. In the top panel of the Figure, we show the results when using a uniform distribution on the interval [0,1]. The panel in the middle shows the results when using a Beta distribution with parameters $(0.5, 0.1)$, which is the specification used in Smets and Wouters (2007) (SW). Finally, the bottom panel reports results when using a tighter Beta distribution, with parameters $(0.5, 0.05)$.

Figure 1 shows the posterior distributions for the two Calvo parameters for each of the three specifications. The results, consistent with the ones reported in Del Negro and Schorfheide (2008) are quite striking: as the confidence in the prior, centered around 0.5 weakens, the posterior moves closer to the upper bound of the parameter, both for the Calvo price and wage parameters.

The implications regarding the corresponding slopes of the Phillips curve are also large. These are reported in Table 1. Moving from the SW priors – a Beta distribution with parameters $(0.5, 0.1)$ as depicted in the middle panel of Table 1 – to uniform priors, reduces the slope of the price Phillips curve by two thirds and the slope of the wage Phillips curve by half.

3.1.2 The Indirect Effect

Precision of priors. In describing the direct effect, we estimated the model assuming the same priors as SW for the policy rule parameters and for the moments of the distributions governing the shocks.[5] We now compare those results with the ones in which we maintain the priors on

[5]For the policy rule parameters, SW use the following: for α_i, a Beta distribution centered around 0.75 with a standard deviation of 0.1; for α_π, a Normal distribution centered at 1.5 with a standard deviation of 0.125; for α_x, a Normal distribution centered at 0.125 with a standard deviation of 0.05; and for α_y, a Normal distribution centered at 0.125 with a standard deviation of 0.05. For the autoregressive shock processes, SW use Beta distributions centered at 0.5 with a standard deviation of 0.2 for the persistences, and an Inverse Gamma distribution with a

Table 1: Estimated Parameters and Slopes, Aggregate Data, 1977 to 2017

	Calvo Estimates			Phillips Curve Slopes*		
	Mode	5%	95%	Mode	5%	95%
	Uniform Distribution Priors, $U(0,1)$					
Prices (λ_p)	**0.94**	0.92	0.98	**0.003**	0.001	0.009
Wages (λ_w)	**0.88**	0.84	0.91	**0.017**	0.009	0.030
	Beta Distribution Priors, $Beta(0.5, 0.1)$					
Prices (λ_p)	**0.90**	0.88	0.93	**0.010**	0.006	0.019
Wages (λ_w)	**0.85**	0.81	0.88	**0.027**	0.017	0.046
	Beta Distribution Priors, $Beta(0.5, 0.05)$					
Prices (λ_p)	**0.84**	0.81	0.86	**0.032**	0.023	0.047
Wages (λ_w)	**0.78**	0.75	0.81	**0.065**	0.046	0.089

$\star$: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
$\star$: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

Figure 1: Posterior Distributions Calvo Parameters, Aggregate Data, 1977 to 2017

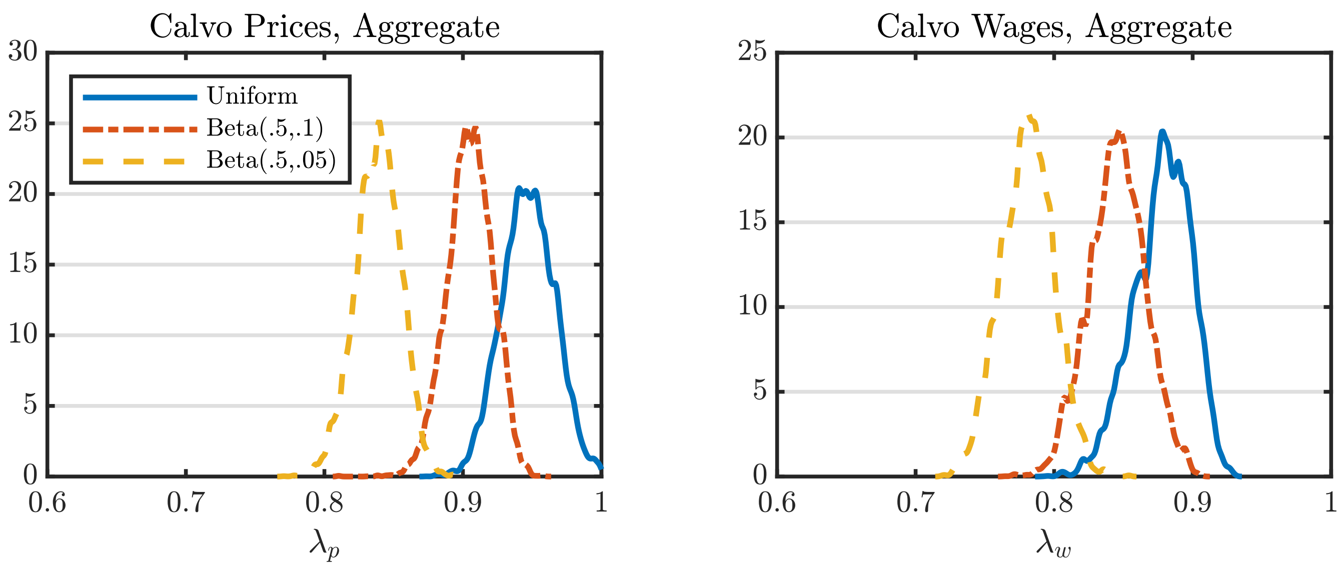


Table 2: Estimated Parameters Across Non-Calvo Priors, Aggregate Data, 1977 to 2017

	SW Priors on All			Uniform on TR Only			Uniform on All Else		
	Mode	5%	95%	Mode	5%	95%	Mode	5%	95%
					Parameter Estimates				
λ_p	**0.90**	0.88	0.93	**0.90**	0.87	0.93	**0.90**	0.87	0.92
λ_w	**0.85**	0.81	0.88	**0.85**	0.81	0.88	**0.76**	0.71	0.84
α_i	**0.81**	0.77	0.84	**0.83**	0.77	0.87	**0.84**	0.78	0.88
α_p	**1.61**	1.47	1.79	**2.07**	1.71	2.61	**2.20**	1.84	2.99
α_x	**0.28**	0.24	0.33	**0.42**	0.34	0.55	**0.41**	0.33	0.54
α_y	**0.12**	0.10	0.17	**0.12**	0.07	0.19	**0.13**	0.09	0.21
					Phillips Curve Slopes*				
Prices	**0.012**	0.006	0.017	**0.012**	0.006	0.020	**0.012**	0.007	0.020
Wages	**0.027**	0.017	0.046	**0.027**	0.017	0.046	**0.077**	0.031	0.120

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

the Calvo parameters as in SW, but set uniform priors for the policy rule parameters and the moments of the distributions of the shocks.[6]

The results are reported in Table 2. The first column presents the results for the Calvo parameters as well as for the policy rule parameters, when using SW priors for all. The second column results correspond to the estimation when using uniform priors on the Taylor rule parameters only, while the third column reports the case in which uniform priors are used for the Taylor rule parameters as well as for the parameters governing the distribution of the shocks.

For the Calvo price parameter, there is remarkably little indirect effect: the mode of the posterior is the same, independently of the priors used for the policy parameters or for the moments of the shocks. However, there are important differences for the Calvo wage parameter. The mode of the posterior goes down from 0.85 to 0.76.[7] However, this is only the case when the confidence on the priors of the moments of the shocks is relaxed. Allowing for uniform priors on the policy parameters does not change the estimated value, in spite the fact that the values for the policy parameter do change somewhat. Note also that the estimated policy parameters

mean of 0.1 and a standard deviation of 2.0 for the standard deviations of the shocks.

[6]For details, see the Online Appendix.

[7]Due to the very non-linear relationship between the Calvo parameter and the slope of the corresponding Phillips curve when the $\lambda's$ are close to their extreme points, this apparently small difference in the λ's translate into very large differences in the slopes. For a detailed discussion of this see Fitzgerald et al. (2020).

Table 3: Estimated Parameters Across Non-Calvo Prior Means, Aggregate Data, 1977 to 2017

	Baseline SW Priors			α_p Centered at 5		
	Mode	5%	95%	Mode	5%	95%
λ_p	**0.90**	0.88	0.93	**0.90**	0.88	0.93
λ_w	**0.85**	0.81	0.88	**0.84**	0.81	0.89
α_i	**0.81**	0.77	0.84	**0.94**	0.93	0.95
α_p	**1.61**	1.47	1.79	**4.89**	4.71	5.12
α_x	**0.28**	0.24	0.33	**0.36**	0.32	0.41
α_y	**0.12**	0.10	0.17	**0.19**	0.11	0.25

	α_y Centered at 1			α_x Centered at 1		
	Mode	5%	95%	Mode	5%	95%
λ_p	**0.90**	0.88	0.93	**0.90**	0.87	0.92
λ_w	**0.86**	0.83	0.90	**0.86**	0.83	0.89
α_i	**0.97**	0.96	0.98	**0.70**	0.62	0.79
α_p	**1.56**	1.35	1.76	**1.67**	1.50	1.87
α_x	**0.35**	0.30	0.39	**0.94**	0.86	1.02
α_y	**0.96**	0.87	1.04	**0.09**	0.05	0.15

barely change when using uniform distributions for the moments governing the shocks - only α_p changes, but by a very small magnitude relative to the case in which uniform priors were used for the policy parameters. This implies that the relevant changes are not the variation of the policy parameters. The next analysis, where we change the mean value for the priors of the policy parameters, confirms this notion.

Mean of priors. We now repeat the estimation, but changing the mean of the prior of the coefficients of the Taylor rule that govern the response of the policy rate to inflation, the output gap and the growth of output. In all cases, we use the same priors as in SW, with the same precision, but we center the parameter at alternative values. The results are reported in Table 3. The upper left panel reproduces the results of estimating the model with SW priors, as reported in Table 2 above. The other panels show the results when we change, one at a time, the means of the prior of α_p from 1.5 to 5, α_y from 0.125 to 1, and α_x from 0.125 to 1. As it can be seen from the Table, the change in the prior does have a substantial impact on the estimated value for the corresponding parameter. However, they have very little impact on the estimated value for the other policy parameters. And, more remarkable, there is barely any effect on the posterior distribution of the Calvo parameters.

3.1.3 Discussion of Results Using Aggregate Data

Overall, we see the results of this subsection as consistent with the notion, documented in the literature, that estimates of small scale New Keynesian models deliver very high degrees of price and wage rigidities. Obviously, tight priors around lower values does deliver lower posteriors, but, as it becomes clear in using uniform priors, the data prefers quite strong price rigidities.

The results of the indirect effect strongly reinforce that notion: the data so strongly prefers high price and wage rigidity, that varying the priors on policy parameters does not change the conclusion. This is the case when we allow for less precise priors, but also, and quite strongly, when we center the priors of the policy parameters at different levels.

An important exception appears once we allow for less precise priors on the parameters governing the law of motion for the shocks, in which case we estimate substantially lower rigidity in wages.

3.2 Using State-level Data

The results are quite the opposite in using state-level data to identify the Calvo parameters. As before, we separately analyze the direct effect from the indirect effect.

3.2.1 The Direct Effect

We estimated the model using state-level data for the same three alternative priors for the Calvo parameter, namely, a beta distribution with parameters $(0.5, 0.1)$ as is standard in the literature, a beta distribution with parameters $(0.5, 0.05)$ and a uniform distribution on the unit interval $[0, 1]$. In all three cases, we used beta priors for the moments governing the distributions of the shocks, as it is customary in the literature. Recall that in this case we do not need to jointly estimate the parameters of the policy rule.

The results show a notable robustness of the posterior distribution with respect to the assumed priors. The mode and the 5% and 95% percentiles are reported, for each case, in Table 4. The three posterior distributions for the two Calvo parameters are reported on Figure 2. As it is clear from the Table, the results are totally insensitive with respect to the priors. We estimate the Calvo price parameter to be around 0.6 and the Calvo wage parameter to be around 0.4 quite consistently. The corresponding values for the slopes of the Phillips curves are presented in Table 4.

Table 4: Estimated Parameters and Slopes, State-Level Data, 1977 to 2017

	Calvo Estimates			Phillips Curve Slopes*		
	Mode	5%	95%	Mode	5%	95%
Uniform Distribution Priors, $U(0,1)$						
Prices (λ_p)	**0.59**	0.57	0.60	**0.29**	0.27	0.33
Wages (λ_w)	**0.41**	0.39	0.42	**0.87**	0.79	0.97
Beta Distribution Priors, $Beta(0.5, 0.1)$						
Prices (λ_p)	**0.59**	0.57	0.60	**0.30**	0.27	0.33
Wages (λ_w)	**0.41**	0.39	0.43	**0.85**	0.77	0.96
Beta Distribution Priors, $Beta(0.5, 0.05)$						
Prices (λ_p)	**0.59**	0.57	0.60	**0.30**	0.27	0.34
Wages (λ_w)	**0.41**	0.39	0.43	**0.86**	0.75	0.94

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

Figure 2: Posterior Distributions Calvo Parameters, State Data, 1977 to 2017

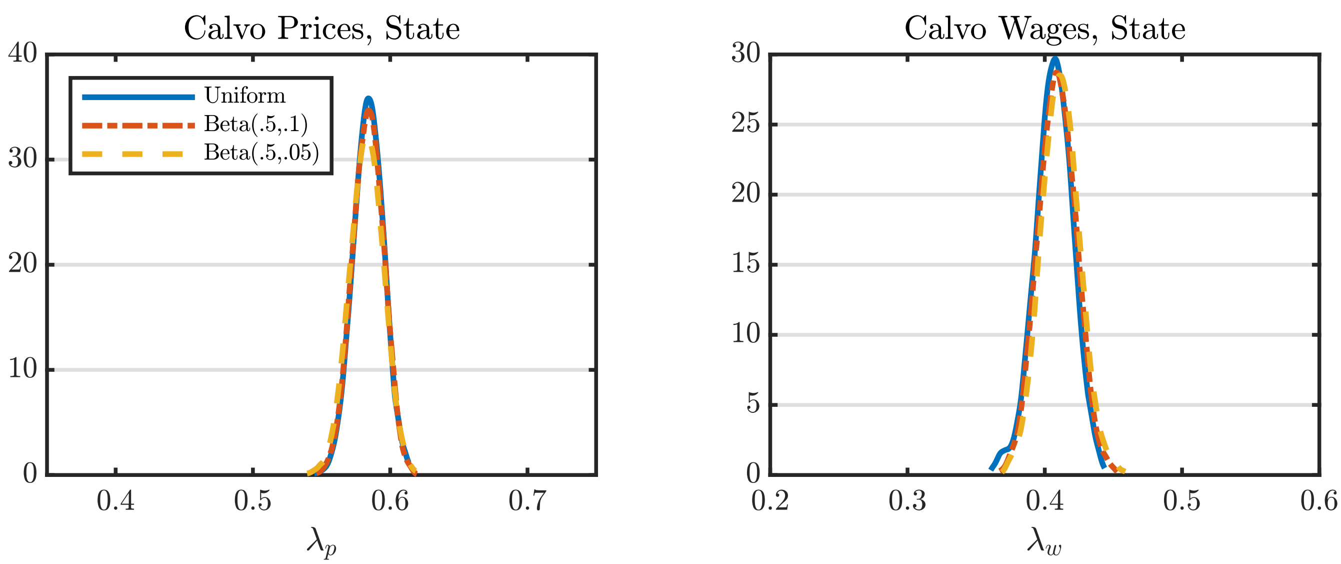


Table 5: Estimated Parameters Across Non-Calvo Priors, State Data, 1977 to 2017

	SW Priors on All			Uniform on Variances			Uniform on All Else		
	Mode	5%	95%	Mode	5%	95%	Mode	5%	95%
				Calvo Estimates					
λ_p	**0.59**	0.57	0.60	**0.59**	0.57	0.60	**0.57**	0.55	0.58
λ_w	**0.41**	0.39	0.43	**0.42**	0.40	0.44	**0.38**	0.36	0.39
				Phillips Curve Slopes*					
Prices	**0.295**	0.272	0.331	**0.294**	0.267	0.324	**0.330**	0.301	0.370
Wages	**0.850**	0.768	0.956	**0.785**	0.719	0.897	**1.006**	0.940	1.112

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$

*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

3.2.2 The Indirect Effect

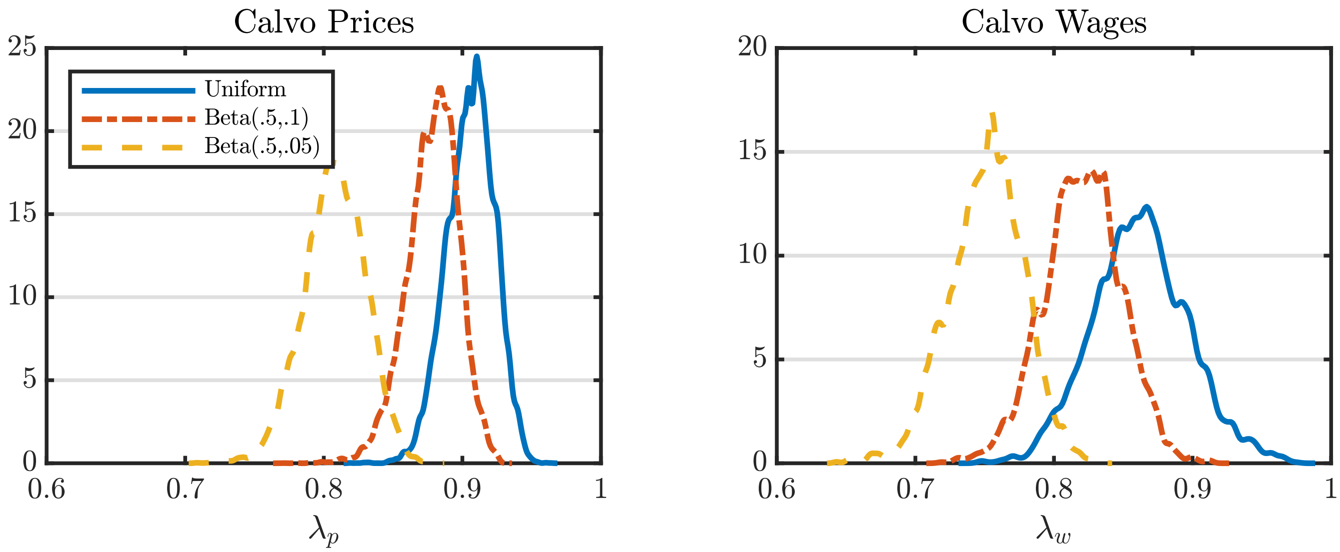
We now repeat the estimation, but assuming uniform distributions for the moments of the variance-covariance matrix of the structural shocks. The results are reported in Table 5. The first panel of the Table reproduces the results using SW priors for all parameters. The second panel shows the effect of using uniform priors for the variances and the third panel the results of using uniform priors for the variances and the persistence parameters. As the Table makes clear, the effect of alternative priors on the estimated values is quite small. These differences translate into very minor differences of the slopes of the Phillips curves, that remain around 0.3 for prices, and lie between 0.8 and 1.0 for wages.

3.3 The Role of Priors in a Medium-Scale New Keynesian Model

In this sub-section we explore the role of priors for both the direct and the indirect effect in estimating a medium-scale model, as the one developed in Smets and Wouters (2007) (SW). The equations of the model are provided in Appendix F.

This model uses data that is not available at the state-level, so we provide estimates using aggregate data only. As we cannot compare it with state-level data estimation, we use 1966 as the starting year of the sample, as in SW. We discuss both the direct and the indirect effect.

Figure 3: Posterior Distributions, Aggregate Data, Medium-Scale Model, 1966 to 2017



3.3.1 The Direct Effect

In Table 6 we show the results for the three different priors we used before. In Figure 3 we plot the posterior distributions for the two Calvo parameters for each of the three specifications. The results are similar to the ones obtained with the small-scale model: as the confidence in the center of the prior weakens, the posterior moves closer to the upper bound of the parameter, both for the Calvo price and wage parameters.

The estimated Calvo parameters are a just a few points lower than for the small-scale model. The way they depend on the priors affects the posterior is notably similar. As it was the case for the small scale model, the estimated values are substantially higher than the ones obtained with state-level data.

3.3.2 The Indirect Effect

Precision of priors. As before, we compare the benchmark estimation where all priors are set as in Smets and Wouters (2007) to the case in which we maintain the priors on the Calvo parameters as in SW, but set uniform priors for the policy rule parameters and the moments of the distributions of the shocks.

The results of the estimation when using SW priors are presented in the first column of Table 6. In the second column we report the results when using the same SW priors for the Calvo parameters and for the moments governing the distribution of shocks, but use uniform priors

Table 6: Estimated Parameters, Aggregate Data, Medium-Scale Model, 1966 to 2017

	Mode	5%	95%	Mode	5%	95%	Mode	5%	95%
				Direct Effect, Across Priors on Calvos					
	Uniform			*Beta*(0.5,0.1)			*Beta*(0.5,0.05)		
λ_p	**0.91**	0.88	0.93	**0.88**	0.85	0.91	**0.80**	0.77	0.84
λ_w	**0.87**	0.81	0.92	**0.83**	0.77	0.87	**0.76**	0.71	0.79
α_i	**0.78**	0.59	0.90	**0.75**	0.59	0.89	**0.74**	0.58	0.90
α_p	**1.55**	1.40	1.71	**1.61**	1.45	1.78	**1.79**	1.61	1.94
α_x	**0.23**	0.19	0.28	**0.23**	0.18	0.27	**0.19**	0.15	0.23
α_y	**0.11**	0.08	0.14	**0.12**	0.09	0.14	**0.12**	0.10	0.15
				Direct Effect, Phillips Curve Slopes*					
Prices	**0.009**	0.006	0.017	**0.017**	0.009	0.027	**0.051**	0.031	0.070
Wages	**0.020**	0.007	0.046	**0.036**	0.020	0.070	**0.077**	0.057	0.120
				Indirect Effect, Across Priors on Other Parameters					
	SW Priors on All			Uniform on TR Only			Uniform on All Else		
λ_p	**0.88**	0.85	0.91	**0.87**	0.84	0.90	**0.88**	0.83	0.91
λ_w	**0.83**	0.77	0.87	**0.81**	0.77	0.86	**0.82**	0.77	0.86
α_i	**0.75**	0.59	0.89	**0.40**	0.05	0.92	**0.43**	0.04	0.83
α_p	**1.61**	1.45	1.78	**1.86**	1.57	2.19	**1.91**	1.65	2.26
α_x	**0.23**	0.18	0.27	**0.27**	0.22	0.33	**0.27**	0.22	0.33
α_y	**0.12**	0.09	0.14	**0.13**	0.10	0.17	**0.12**	0.10	0.16
				Indirect Effect, Phillips Curve Slopes*					
Prices	**0.017**	0.009	0.027	**0.020**	0.012	0.031	**0.017**	0.009	0.036
Wages	**0.036**	0.020	0.070	**0.046**	0.023	0.070	**0.040**	0.023	0.070

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

for the policy rule parameters. Finally, the third column reports estimates using SW priors on the Calvo parameters, while using uniform priors for the parameters of the Taylor rule and for the moments governing the evolution of the shocks. Interestingly enough, the estimated values for both Calvo parameters are remarkably robust to this indirect effect. There are some variations on the estimated values for the Calvo parameter, but the differences are very small in all cases. This is so even though the estimated values for the policy parameters do vary quite a bit in some cases.

Mean of priors. As a final exercise, we now study the robustness of the estimated posterior distributions of the Calvo parameters when we change the mean of the priors of the policy parameters. As we did in the case of the small scale model, we change, one at a time, the means of the prior of α_p from 1.5 to 5, α_y from 0.125 to 1, and α_x from 0.125 to 1.

The results are reported in Table 7. In this case, the effect is larger than in the small scale model, particularly so in setting a higher value for the response of the interest rate to inflation, α_p. Now, the range of estimates for the price stickiness parameter goes from 0.73 to 0.90, while for the wage one is 0.64 to 0.83. These imply substantial differences in the slopes of the corresponding Phillips curves. As in the case of the small scale model, the estimated values of the policy parameters are quite robust to changes in priors in the other policy parameters. This indicates that the cross derivatives of the likelihood functions among policy parameters are quite small. We conjecture that this is the consequence of a likelihood function that is quite flat over the values of the policy parameters, indicating difficulties in identifying the structural parameters of the policy rule. We leave a full quantitative analysis of the likelihood functions for further research.

4 Intuition from Analytical Results

The previous section established two key results regarding inferences of the Calvo parameters: *(i)* direct effects are strong for the aggregate model but quite weak for the state-level model; and *(ii)* indirect effects of policy rule parameters are quantitatively relevant for the aggregate model in some cases, but irrelevant for the state-level model overall. Why is this so? Because the empirical model is solved numerically, explicit analytical expressions between the structural parameters, priors and the data are not available. Here we consider a model, that because of its simplicity, allows us to explore analytically two key results that resemble the patterns observed

Table 7: Estimated Parameters Across Non-Calvo Priors, Medium-Scale Model, 1966 to 2017

	Baseline SW Priors			α_p Centered at 5		
	Mode	5%	95%	Mode	5%	95%
λ_p	**0.88**	0.85	0.91	**0.73**	0.70	0.77
λ_w	**0.83**	0.77	0.87	**0.71**	0.64	0.76
α_i	**0.75**	0.59	0.89	**0.75**	0.58	0.89
α_p	**1.61**	1.45	1.78	**5.02**	4.87	5.20
α_x	**0.23**	0.18	0.27	**0.29**	0.24	0.35
α_y	**0.12**	0.09	0.14	**0.22**	0.15	0.28

	α_y Centered at 1			α_x Centered at 1		
	Mode	5%	95%	Mode	5%	95%
λ_p	**0.89**	0.87	0.91	**0.90**	0.88	0.93
λ_w	**0.80**	0.74	0.85	**0.86**	0.81	0.90
α_i	**0.77**	0.59	0.89	**0.78**	0.59	0.89
α_p	**1.40**	1.22	1.60	**1.62**	1.39	1.78
α_x	**0.26**	0.20	0.31	**0.80**	0.72	0.88
α_y	**1.00**	0.93	1.08	**0.24**	0.19	0.30

in the larger estimated model.

We first take the simplest version of an New Keynesian model we can think of to illustrate how inferences of the slope of the Phillips curve are intertwined with monetary policy rule parameters when relying on aggregate data. We then show how relying on a state-level model written in terms of deviations from the aggregate can disentangle inferences of κ from monetary policy. The aggregate model is given by

$$x_t = -(i_t - \mathbb{E}_t \pi_{t+1}) + \mathbb{E}_t x_{t+1} + \varepsilon_t^x \tag{4}$$

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t + \varepsilon_t^\pi \tag{5}$$

$$i_t = \phi \pi_t, \tag{6}$$

where x_t is the output gap, i_t is the nominal interest rate and π_t is inflation. The shocks ε_t^x and ε_t^π are *iid* with mean 0 and standard deviation σ_x and σ_π respectively. The slope of the Phillips curve is κ and the monetary policy response to inflation is ϕ. Thus, the structural model has four structural parameters, $\{\phi, \kappa, \sigma_x, \sigma_\pi\}$.

One could consider different cases such as with an intertemporal elasticity of substitution different from 1, or with a monetary policy rule that responds to the output gap and has a monetary policy shock. But we deliberately abstract from these cases as it is hard to get clean analytical expressions to highlight the role of priors. We consider a case in which the structural parameters are not identified in the aggregate case – consistent with the findings of Komunjer and Ng (2011) and Iskrev (2010) in larger models – but are just identified at the state-level. In the discussion that follows, moving to a state-level model sharpens identification of the Phillips curve slope, consistent with our findings above. This is not to say, however, that other situations may not be possible in practice.

Substituting (6) in (4) gives a 2×2 system in terms of $y_t = (x_t, \pi_t)'$, which in the form of a general linear rational expectations model is

$$A_0 y_t = A_1 y_{t-1} + B_0 \mathbb{E}_t y_{t+1} + D_0 \varepsilon_t. \tag{7}$$

In our case, $A_1 = 0$ and $D_0 = I$, so by undetermined coefficients the solution to (7) is given by

$$y_t = A_0^{-1} \varepsilon_t. \tag{8}$$

In this case

$$\begin{pmatrix} x_t \\ \pi_t \end{pmatrix} = \frac{1}{1+\kappa\phi} \begin{pmatrix} 1 & -\phi \\ \kappa & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_t^x \\ \varepsilon_t^\pi \end{pmatrix}. \tag{9}$$

Equation (8) implies that $\mathbb{E}_t x_{t+1} = \mathbb{E}_t \pi_{t+1} = 0$ for all t. Hence, the structural equations in (7) collapse to the SVAR below

$$A_0 y_t = \varepsilon_t.$$

Let the variance covariance matrix of the structural shocks, ε_t, be the *diagonal* matrix, Σ. Assuming y_t is observed, one can estimate the variance covariance matrix of the reduced form shocks, $A_0^{-1}\varepsilon_t$, given by

$$\Omega = A_0^{-1} \Sigma (A_0^{-1})'. \tag{10}$$

Equation (10) maps the structural parameters to the reduced form variance covariance matrix. From estimation of the reduced form VAR, one obtains an estimate of the variance covariance matrix, that is, an estimate of the variance of x_t, the variance of π_t and the covariance between them. We denote these estimates by V_x, V_π, and $C_{x,\pi}$ respectively. Equation (10) implies the system below

$$V_x = \frac{\sigma_x^2}{(1+\kappa\phi)^2} + \frac{\phi^2\sigma_\pi^2}{(1+\kappa\phi)^2} \tag{11}$$

$$C_{x,\pi} = \frac{\kappa\sigma_x^2}{(1+\kappa\phi)^2} - \frac{\phi\sigma_\pi^2}{(1+\kappa\phi)^2} \tag{12}$$

$$V_\pi = \frac{\kappa^2\sigma_x^2}{(1+\kappa\phi)^2} + \frac{\sigma_\pi^2}{(1+\kappa\phi)^2}. \tag{13}$$

Notice that in the aggregate model inferences of the slope of the Phillips curve, κ, will be intertwined with inferences of the monetary policy rule parameter, ϕ. Notice also that the equations above map three reduced form moments into the four structural parameters of interest $\kappa, \phi, \sigma_x, \sigma_\pi$. Equation (10) has infinitely many solutions: the structural parameters are not identified. Just like in standard VAR analysis, one requires some restriction (or some prior) on the structural parameters to achieve identification (or to select one solution).

The analysis to this point has relied on an estimated reduced form VAR. In practice, of course, Bayesian estimation of structural models relies on the log posterior, $\mathcal{P}$, which is the sum of the log-likelihood, $\mathcal{L}$, and the prior, p. So we now discuss direct and indirect effects on κ in these terms. Collecting the observed data in $Y = \{x_t, \pi_t\}_{t=1}^{T}$, it is easy to show that the log-likelihood is

$$\mathcal{L}(Y|\kappa, \phi, \sigma_x, \sigma_\pi) = -T\ln(2\pi) - T\left[\ln\left(\frac{\sigma_x \sigma_\pi}{1+\kappa\phi}\right)\right] - \frac{1}{2}\sum_{t=1}^{T}\left[\left(\frac{\pi_t - \kappa x_t}{\sigma_\pi}\right)^2 + \left(\frac{\phi\pi_t + x_t}{\sigma_x}\right)^2\right],$$

and the log posterior, in turn, is given by

$$\mathcal{P}(\kappa, \phi, \sigma_x, \sigma_\pi|Y) = \mathcal{L}(Y|\kappa, \phi, \sigma_x, \sigma_\pi) + p(\kappa, \phi, \sigma_x, \sigma_\pi), \tag{14}$$

where $p(\kappa, \phi, \sigma_x, \sigma_\pi)$ stands for the priors on the structural parameters, which are taken to be independent.

The partial derivatives of $\mathcal{L}$ with respect to the four parameters are

$$\mathcal{L}_\kappa = T\left[\frac{\phi}{(1+\kappa\phi)}\right] + \sum_{t=1}^{T}\left(\frac{\pi_t - \kappa x_t}{\sigma_\pi}\right)\frac{x_t}{\sigma_\pi} \tag{15}$$

$$\mathcal{L}_\phi = T\left[\frac{\kappa}{(1+\kappa\phi)}\right] - \sum_{t=1}^{T}\left(\frac{\phi\pi_t + x_t}{\sigma_x}\right)\frac{\pi_t}{\sigma_x} \tag{16}$$

$$\mathcal{L}_{\sigma_\pi} = T\left[\frac{1}{\sigma_\pi}\right] - \sum_{t=1}^{T}\left(\frac{\pi_t - \kappa x_t}{\sigma_\pi}\right)^2\frac{1}{\sigma_\pi} \tag{17}$$

$$\mathcal{L}_{\sigma_x} = T\left[\frac{1}{\sigma_x}\right] - \sum_{t=1}^{T}\left(\frac{\phi\pi_t + x_t}{\sigma_x}\right)^2\frac{1}{\sigma_x}. \tag{18}$$

The maximum of the log-likelihood function is attained when those four partial derivatives are set equal to zero. In the appendix we show that, given any sample $Y = \{x_t, \pi_t\}_{t=1}^{T}$, if the vector $(\widetilde{\kappa}, \widetilde{\phi}, \widetilde{\sigma}_x, \widetilde{\sigma}_\pi)$ satisfies the first three equations, then it satisfies the fourth. This is another way of stating that the system is not identified. All these functions are continuous and differentiable, so this means that there exists a differentiable function

$$S(\kappa, \phi, \sigma_x, \sigma_\pi) = 0,$$

such that the solutions of the system above are given by the solution to this equation.

Once we chose priors, the maximum of the likelihood function is attained when those four partial derivatives are set equal to the negative of the partial derivative of the prior with respect to the corresponding parameter, that is

$$\begin{aligned}
\mathcal{L}_{\kappa} &= -p_{\kappa} \\
\mathcal{L}_{\phi} &= -p_{\phi} \\
\mathcal{L}_{\sigma_\pi} &= -p_{\sigma_\pi} \\
\mathcal{L}_{\sigma_x} &= -p_{\sigma_x}.
\end{aligned}$$

It follows that as long as the prior is not a trivial function of the parameters, so the partial derivatives of the prior are non-trivial functions of its arguments, the system of necessary conditions for a maximum of the posterior will identify, in general, a unique local solution. Thus, the choice of a prior can be thought of as way of selecting a solution (or a subset of them) among the infinitely many ones that satisfy $S(\kappa, \phi, \sigma_x, \sigma_\pi) = 0$.

To see how priors on the monetary policy rule parameter affect inferences of the slope of the Phillips curve, κ, it is useful to consider the case of uniform priors on $\kappa, \sigma_x, \sigma_\pi$ with a dogmatic prior on ϕ, equivalent to calibrating ϕ. Assuming we are evaluating the posterior at the interior of the support of the priors, uniform priors imply the derivatives of the priors are zero, so the system of partial derivatives of the posterior is the same as the partial derivatives of the likelihood, so the solutions to the partial derivatives is the intersection of the solutions of $S(\kappa, \phi, \sigma_x, \sigma_\pi) = 0$ and the support of the priors.

When ϕ is calibrated, the system is identified. The three equations that solve for $(\kappa, \sigma_x, \sigma_\pi)$ are the partial derivatives with respect to κ, σ_π and σ_x. The first order conditions are (15), (17) and (18). Equations (15) and (17) determine κ and σ_π, while (18) determines σ_x. One can show that the solution for κ satisfies

$$\kappa = \frac{C_{x,\pi} + \phi V_\pi}{V_x + C_{x,\pi}\phi}. \tag{19}$$

which reveals the determinants of the indirect effect, i.e. how inferences about κ are determined by the moments of the data, V_x, V_π and $C_{x,\pi}$ as well as the prior on ϕ. Indeed, in this case changing the mean of the dogmatic prior on ϕ affects posterior inferences on κ as given by

$$\frac{\partial \kappa}{\partial \phi} = \frac{V_\pi V_x}{(V_x + C_{x,\pi}\phi)^2}\left[1 - \frac{C_{x,\pi}^2}{V_\pi V_x}\right].$$

Moving to a state-level model, however, disentangles inferences of κ from ϕ. To see this, assume following McLeay and Tenreyro (2020), that there are n regions and aggregate inflation and the aggregate output gap are just weighted averages of inflation and the output gap in each of the n regions, that is

$$\pi_t = \sum_{s=1}^{n} \alpha(s)\pi_t(s), \tag{20}$$

and

$$x_t = \sum_{s=1}^{n} \alpha(s)x_t(s), \tag{21}$$

with $\sum_{s=1}^{n} \alpha(s) = 1$. In state s the regional output gap and regional inflation follow the equations below

$$\begin{aligned} x_t(s) &= -(i_t - \mathbb{E}_t\pi_{t+1}(s)) + \mathbb{E}_t x_{t+1}(s) + \varepsilon_t^x(s) & (22)\\ \pi_t(s) &= \beta\mathbb{E}_t\pi_{t+1}(s) + \kappa x_t(s) + \varepsilon_t^\pi(s). & (23) \end{aligned}$$

Define deviations from the aggregate as $\hat{\pi}_t(s) = \pi_t(s) - \sum_{s=1}^{n} \alpha(s)\pi_t(s)$ and $\hat{x}_t(s) = x_t(s) - \sum_{s=1}^{n} \alpha(s)x_t(s)$ and subtract aggregate equations from (23) and (22) to obtain

$$\begin{aligned} \hat{x}_t(s) &= \mathbb{E}_t\hat{\pi}_{t+1}(s) + \mathbb{E}_t\hat{x}_{t+1}(s) + \hat{\varepsilon}_t^x(s) & (24)\\ \hat{\pi}_t(s) &= \beta\mathbb{E}_t\hat{\pi}_{t+1}(s) + \kappa\hat{x}_t(s) + \hat{\varepsilon}_t^\pi(s). & (25) \end{aligned}$$

By undetermined coefficients the solution is given by A_0^{-1} which in this case reads

$$\begin{pmatrix} \hat{x}_t(s) \\ \hat{\pi}_t(s) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \kappa & 1 \end{pmatrix} \begin{pmatrix} \hat{\varepsilon}_t^x(s) \\ \hat{\varepsilon}_t^\pi(s) \end{pmatrix}. \tag{26}$$

For the state-level model, the system of equations implied by (10) is

$$\begin{aligned} V_x(s) &= \hat{\sigma}_x^2 & (27)\\ C_{x,\pi}(s) &= \kappa\hat{\sigma}_x^2 & (28)\\ V_\pi(s) &= \kappa^2\hat{\sigma}_x^2 + \hat{\sigma}_\pi^2. & (29) \end{aligned}$$

Now at the state-level, there are three equations for the three structural parameters of interest, $\kappa, \hat{\sigma}_x, \hat{\sigma}_\pi$, so the parameters, including κ, can be pinned down uniquely of state-level moments. As monetary policy is common across states, relying on deviations of state level data from the aggregate makes inferences of κ independent of monetary policy. As such, there is no indirect effect of priors on policy parameters on posterior inferences of κ.

The log-likelihood for state-level data of *state s* is given by

$$\mathcal{L}(\hat{Y}(s)|\kappa, \hat{\sigma}_x, \hat{\sigma}_\pi) = -T(s)\ln(2\pi) - T(s)\left[\ln(\hat{\sigma}_x\hat{\sigma}_\pi)\right] - \frac{1}{2}\sum_{t=1}^{T(s)}\left[\left(\frac{\hat{\pi}_t(s) - \kappa\hat{x}_t(s)}{\hat{\sigma}_\pi}\right)^2 + \left(\frac{\hat{x}_t(s)}{\hat{\sigma}_x}\right)^2\right].$$

In considering the whole economy, the log-likelihood of the state-level model is simply the sum over all states, that is $\sum_s \mathcal{L}(s)$. So moving from the aggregate model to the state level model changes the role of priors with respect to inferences of κ in two non-mutually exclusive ways. First, inferences on κ are disentangled from those of the monetary policy rule; econometrically, it reduces the number of structural parameters to estimate which has the potential to sharpen identification. In this simple example it does. Second, increasing the numbers of observations reduces the impact that priors exert on the posterior. These observations are consistent with our empirical findings. In Section 2 we found that inferences of nominal rigidities with aggregate models are quite sensitive to changes in the prior mean of the Calvo parameters as well as to changes in the prior mean of the policy rule parameters, but inferences of nominal rigidities were found to be quite robust to such changes in the state-level model.

5 Discussion and Conclusions

Estimating the slope of the Phillips curve in New Keynesian models has proven to be difficult and has spurred the use of Bayesian methods. We have explored how the choice of priors over the estimated parameters can influence the estimated slope of the Phillips curve. Our main innovation is to use state-level variation to estimate the Calvo parameters, and compare the estimation performance to the standard strategy of using aggregate data.

In using aggregate data, directly changing the prior on the Calvo price and wage adjustment parameters that govern the slope of the Phillips curve leads to very different posterior estimates of the Calvo parameters, in line with previous work. This implies very different inferences about the stickiness of prices and wages. Furthermore, we show that the priors on the other parameters of the model indirectly influence the final posterior estimate of the price and wage Calvo parameters in some cases.

In contrast, the estimates obtained using state-level data are notably robust to priors over the Calvo parameters. Furthermore, the indirect effect of priors specified on other parameters is also not important. This reflects the usefulness of state-level data for identification in two key ways: by first removing the role of endogenous monetary policy, and second, by the use of more data by exploiting the full panel of U.S. states.

Table 8: Posterior Distributions, Medium-Scale Model, 1966 to 2017

	Estimated			Calibrated		
	Mode	5%	95%	Mode	5%	95%
λ_p	0.88	0.85	0.91	-	-	-
λ_w	0.83	0.77	0.87	-	-	-
α_i	0.75	0.59	0.89	0.76	0.57	0.89
α_p	1.61	1.45	1.78	1.96	1.81	2.12
α_x	0.23	0.18	0.27	0.13	0.10	0.19
α_y	0.12	0.09	0.14	0.05	0.04	0.08

Our estimates of the Calvo parameters, obtained using state-level data, are not only substantially more robust to priors that the ones obtained using aggregate data: they also imply less rigid prices and even less rigid wages, in line with micro-level studies of price rigidities.

It is crucial to accurately estimate the price and wage stickiness parameters. Different values of these parameters can lead to different inferences about the drivers of business cycles in New Keynesian models, directly affecting the model's implications for optimal policy.

We use our results above to illustrate this point. To do so, we compare the unconditional variance decompositions in two versions of the SW model. In the first, we estimate the Calvo wage and price parameters using aggregate data. In the second, we calibrate both Calvo parameters to the values obtained in the state-level estimation and estimate the remaining parameters also using aggregate data (the parameter values are given in Table 8).

The variance decompositions in Table 9 show that, for instance, the wage markup shock becomes much more important in driving fluctuations in real and nominal variables, while the discount factor preference shock becomes less important. A full analysis of the implications for policy that our estimates using state-level data have is beyond the scope of this paper. We see the results of Table 9, however, as an encouraging indicator of the relevance of such an analysis.

Table 9: Unconditional Variance Decomposition, %

Variable \ Shock	Prod	Pref	Gov	Inve	Policy	Price MU	Wage MU
Estimated Calvo Price and Wage Parameters							
Consumption	7	68	2	4	11	3	5
Investment	2	8	0	88	1	1	0
Output	10	35	9	36	5	3	1
Hours	4	24	15	15	3	3	35
Inflation	1	1	0	0	0	29	68
Wages	1	5	0	2	1	12	79
Policy Rate	4	34	0	8	9	7	37
Calibrated Calvo Price and Wage Parameters							
Consumption	12	19	0	3	4	2	60
Investment	1	2	0	93	0	1	4
Output	15	14	7	25	3	2	34
Hours	4	0	0	1	0	1	93
Inflation	1	9	0	9	2	4	75
Wages	10	9	0	4	2	15	61
Policy Rate	1	18	0	19	3	2	57
Difference, Estimated Calvos to Calibrated Calvos							
Consumption	5	−48	−1	−1	−7	−1	55
Investment	−1	−7	0	5	−1	−0	4
Output	4	−22	−2	−11	−2	−1	33
Hours	−0	−24	−15	−14	−3	−2	58
Inflation	−0	8	0	9	2	−25	7
Wages	8	4	0	2	2	2	−18
Policy Rate	−3	−16	0	10	−7	−5	20

References

Canova, Fabio and Luca Sala, "Back to square one: Identification issues in DSGE models," *Journal of Monetary Economics*, 2009, *56* (4), 431–449.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 2005, *113* (1), 1–45.

Del Negro, Marco and Frank Schorfheide, "Forming priors for DSGE models (and how it affects the assessment of nominal rigidities)," *Journal of Monetary Economics*, 2008, *55* (7), 1191–1208.

Fitzgerald, Terry, Callum Jones, Mariano Kulish, and Juan Pablo Nicolini, "Is There a Stable Relationship between Unemployment and Future Inflation? Evidence from U.S. Cities," 2020.

Guerrieri, Luca and Matteo Iacoviello, "Occbin: A Toolkit to Solve Models with Occasionally Binding Constraints Easily," *Journal of Monetary Economics*, March 2015, *70*, 22–38.

Hamilton, James D., *Time Series Analysis*, Princeton University Press, 1994.

Ireland, Peter, "Technology Shocks in the New Keynesian Model," *Review of Economics and Statistics*, November 2004, *86* (4), 923–936.

Iskrev, Nikolay, "Local identification in DSGE models," *Journal of Monetary Economics*, 2010, *57* (2), 189–202.

Jones, Callum, "Unanticipated Shocks and Forward Guidance at the Zero Lower Bound," 2017.

_ , Virgiliu Midrigan, and Thomas Philippon, "Household Leverage and the Recession," Working Paper 2018.

Komunjer, Ivana and Serena Ng, "Dynamic identification of dynamic stochastic general equilibrium models," *Econometrica*, 2011, *79* (6), 1995–2032.

Kulish, Mariano, James Morley, and Tim Robinson, "Estimating DSGE Models with Zero Interest Rate Policy," *Journal of Monetary Economics*, 2017, *88*, 35–49.

McLeay, Michael and Silvana Tenreyro, "Optimal Inflation and the Identification of the Phillips Curve," *NBER Macroeconomics Annual 2019*, 2020, *34*, 199–255.

Mian, Atif and Amir Sufi, "House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis," *American Economic Review*, 2011, *101* (5), 2132–2156.

Müller, U., "Measuring Prior Sensitivity and Prior Informativeness in Large Bayesian Models," *Journal of Monetary Economics*, 2012, *59*, 581–597.

Schorfheide, Frank, "DSGE model-based estimation of the New Keynesian Phillips curve," *FRB Richmond Economic Quarterly*, 2008, *94* (4), 397–433.

Smets, Frank and Rafael Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 2007, *97* (3), 586–606.

Appendix

A Results starting in 1966

We start by describing results in the small-scale model estimated on aggregate data, starting in 1966 (in our baseline, the estimation starts instead in 1977). Table A.1 gives the estimates from 1966 to 2017 across priors on the Calvo parameters, quantifying the direct effect over this period. Table A.2 gives the estimates over priors on the other model parameters, quantifying the indirect effect.

B Results up to 2004

We next explore the robustness of our results to changing the final period of the estimation to 2004, the same as in Smets and Wouters (2007). Table B.3 gives the results on the direct effect oc changing priors over this sample when the small-scale model is estimated using aggregate data. Table B.4 shows the indirect effect using aggregate data. Table B.5 gives the direct effect when state-level data is used on a 1977 to 2004 sample, and Table B.6 gives estimates on the indirect effect using state-level data.

C Changing Mean of Priors

In these robustness exercises, we change the mean of the priors on the Taylor rule parameters without changing their standard deviations. Additional results for the SW model are shown in Table C.9.

D Full Description of the Model

The model description follows Jones, Midrigan and Philippon (2018). We describe the model with the full operative credit channel. But we note that absent this credit channel and the tradeable production structure, the model would reduce to the familiar 3-equation New Keynesian model.

Table A.1: Estimated Parameters and Slopes, Aggregate Data, 1966 to 2017

	Calvo Estimates			Phillips Curve Slopes*		
	Mode	5%	95%	Mode	5%	95%
			Uniform Distribution Priors, $U(0,1)$			
Prices (λ_p)	0.92	0.89	0.94	0.008	0.004	0.013
Wages (λ_w)	0.91	0.87	0.95	0.006	0.003	0.021
			Beta Distribution Priors, $Beta(0.5, 0.1)$			
Prices (λ_p)	0.90	0.88	0.92	0.012	0.008	0.018
Wages (λ_w)	0.86	0.83	0.90	0.022	0.012	0.034
			Beta Distribution Priors, $Beta(0.5, 0.05)$			
Prices (λ_p)	0.85	0.82	0.87	0.029	0.021	0.041
Wages (λ_w)	0.80	0.76	0.83	0.051	0.037	0.074

$\star$: Price Phillips curve slope is $(1 - \beta\lambda_p)(1 - \lambda_p)/\lambda_p$
$\star$: Wage Phillips curve slope is $(1 - \beta\lambda_w)(1 - \lambda_w)/\lambda_w$

Table A.2: Estimated Parameters Across Non-Calvo Priors, Aggregate Data, 1966 to 2017

	SW Priors on All			Uniform on TR Only			Uniform on All Else		
	Mode	5%	95%	Mode	5%	95%	Mode	5%	95%
				Parameter Estimates					
λ_p	0.90	0.88	0.92	0.90	0.87	0.92	0.89	0.87	0.92
λ_w	0.86	0.83	0.90	0.88	0.84	0.90	0.82	0.76	0.87
α_i	0.82	0.79	0.86	0.85	0.79	0.90	0.82	0.79	0.86
α_p	1.56	1.42	1.74	1.97	1.63	2.85	1.62	1.44	1.77
α_x	0.32	0.29	0.37	0.47	0.40	0.61	0.32	0.28	0.37
α_y	0.14	0.11	0.18	0.14	0.09	0.23	0.14	0.11	0.19
				Phillips Curve Slopes*					
Prices	0.012	0.008	0.018	0.012	0.008	0.019	0.013	0.008	0.020
Wages	0.022	0.012	0.034	0.018	0.011	0.032	0.040	0.020	0.080

$\star$: Price Phillips curve slope is $(1 - \beta\lambda_p)(1 - \lambda_p)/\lambda_p$
$\star$: Wage Phillips curve slope is $(1 - \beta\lambda_w)(1 - \lambda_w)/\lambda_w$

Table B.3: Estimated Parameters and Slopes, Aggregate Data, 1977 to 2004

	Calvo Estimates			Phillips Curve Slopes*		
	Mode	5%	95%	Mode	5%	95%
	Uniform Distribution Priors, $U(0,1)$					
Prices (λ_p)	0.98	0.94	1.00	0.000	0.000	0.004
Wages (λ_w)	0.95	0.90	0.96	0.003	0.002	0.012
	Beta Distribution Priors, $Beta(0.5, 0.1)$					
Prices (λ_p)	0.91	0.87	0.94	0.008	0.005	0.020
Wages (λ_w)	0.88	0.85	0.92	0.013	0.008	0.029
	Beta Distribution Priors, $Beta(0.5, 0.05)$					
Prices (λ_p)	0.80	0.76	0.83	0.053	0.035	0.080
Wages (λ_w)	0.79	0.76	0.83	0.053	0.038	0.078

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

Table B.4: Estimated Parameters Across Non-Calvo Priors, Aggregate Data, 1977 to 2004

	SW Priors on All			Uniform on TR Only			Uniform on All Else		
	Mode	5%	95%	Mode	5%	95%	Mode	5%	95%
	Parameter Estimates								
λ_p	0.91	0.87	0.94	0.91	0.87	0.94	0.90	0.87	0.94
λ_w	0.88	0.85	0.92	0.88	0.85	0.92	0.72	0.59	0.80
α_i	0.91	0.85	0.93	0.91	0.85	0.93	0.94	0.90	0.95
α_p	3.44	2.58	5.36	3.44	2.58	5.36	6.50	4.22	7.34
α_x	0.45	0.37	0.59	0.45	0.37	0.59	0.44	0.36	0.60
α_y	0.53	0.28	0.82	0.53	0.28	0.82	0.73	0.58	1.30
	Phillips Curve Slopes*								
Prices	0.009	0.004	0.020	0.009	0.004	0.020	0.012	0.004	0.020
Wages	0.017	0.007	0.027	0.017	0.007	0.027	0.110	0.051	0.287

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

Table B.5: Estimated Parameters and Slopes, State-Level Data, 1977 to 2004

	Calvo Estimates			Phillips Curve Slopes*		
	Mode	5%	95%	Mode	5%	95%
	Uniform Distribution Priors, $U(0,1)$					
Prices (λ_p)	0.52	0.50	0.55	0.44	0.37	0.50
Wages (λ_w)	0.30	0.28	0.32	1.65	1.41	1.81
	Beta Distribution Priors, $Beta(0.5, 0.1)$					
Prices (λ_p)	0.54	0.50	0.56	0.40	0.36	0.49
Wages (λ_w)	0.31	0.29	0.33	1.54	1.35	1.75
	Beta Distribution Priors, $Beta(0.5, 0.05)$					
Prices (λ_p)	0.53	0.51	0.56	0.40	0.35	0.48
Wages (λ_w)	0.32	0.30	0.34	1.43	1.25	1.65

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

Table B.6: Estimated Parameters Across Non-Calvo Priors, State Data, 1977 to 2004

	SW Priors on All			Uniform on All Else		
	Mode	5%	95%	Mode	5%	95%
	Calvo Estimates					
λ_p	0.53	0.50	0.55	0.55	0.53	0.57
λ_w	0.31	0.29	0.33	0.39	0.37	0.41
	Phillips Curve Slopes*					
Prices	0.43	0.37	0.50	0.37	0.32	0.42
Wages	1.53	1.36	1.76	0.93	0.85	1.05

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

Table B.7: Estimated Parameters, Aggregate Data, Medium-Scale Model, 1966 to 2004

	Mode	5%	95%	Mode	5%	95%	Mode	5%	95%
	Direct Effect, Across Priors on Calvos								
	Uniform			*Beta*(0.5,0.1)			*Beta*(0.5,0.05)		
λ_p	0.81	0.73	0.85	0.70	0.63	0.77	0.61	0.56	0.67
λ_w	0.98	0.81	1.00	0.73	0.65	0.80	0.61	0.56	0.68
α_i	0.77	0.56	0.89	0.78	0.56	0.89	0.77	0.59	0.90
α_p	1.54	1.39	1.72	1.72	1.57	1.88	1.77	1.63	1.93
α_x	0.28	0.24	0.32	0.25	0.21	0.30	0.24	0.19	0.28
α_y	0.06	0.05	0.09	0.07	0.05	0.10	0.07	0.05	0.10
	Direct Effect, Phillips Curve Slopes*								
Prices	0.046	0.027	0.101	0.130	0.070	0.219	0.251	0.164	0.348
Wages	0.001	0.000	0.046	0.101	0.051	0.190	0.251	0.152	0.348
	Indirect Effect, Across Priors on Other Parameters								
	SW Priors on All			Uniform on TR Only			Uniform on All Else		
λ_p	0.70	0.63	0.77	0.68	0.61	0.75	0.67	0.60	0.74
λ_w	0.73	0.65	0.80	0.70	0.62	0.78	0.72	0.63	0.79
α_i	0.78	0.56	0.89	0.19	0.05	0.95	0.37	0.05	0.95
α_p	1.72	1.57	1.88	2.14	1.87	2.55	2.15	1.83	2.53
α_x	0.25	0.21	0.30	0.35	0.28	0.41	0.34	0.28	0.42
α_y	0.07	0.05	0.10	0.09	0.07	0.14	0.10	0.07	0.14
	Indirect Effect, Phillips Curve Slopes*								
Prices	0.130	0.070	0.219	0.152	0.085	0.251	0.164	0.093	0.269
Wages	0.101	0.051	0.190	0.130	0.063	0.235	0.110	0.057	0.219

*: Price Phillips curve slope is $(1-\beta\lambda_p)(1-\lambda_p)/\lambda_p$
*: Wage Phillips curve slope is $(1-\beta\lambda_w)(1-\lambda_w)/\lambda_w$

Table C.8: Estimated Parameters Across Non-Calvo Priors, Aggregate Data, 1966 to 2017

	Baseline SW Priors			α_p Centered at 5		
	Mode	5%	95%	Mode	5%	95%
λ_p	0.90	0.88	0.92	0.89	0.87	0.92
λ_w	0.86	0.83	0.90	0.88	0.84	0.91
α_i	0.82	0.79	0.86	0.95	0.93	0.96
α_p	1.56	1.42	1.74	4.96	4.78	5.12
α_x	0.32	0.29	0.37	0.40	0.36	0.45
α_y	0.14	0.11	0.18	0.19	0.14	0.28

	α_y Centered at 1			α_x Centered at 1		
	Mode	5%	95%	Mode	5%	95%
λ_p	0.90	0.87	0.91	0.89	0.87	0.91
λ_w	0.88	0.86	0.92	0.89	0.85	0.91
α_i	0.97	0.96	0.97	0.74	0.65	0.81
α_p	1.62	1.37	1.77	1.65	1.49	1.86
α_x	0.36	0.32	0.41	0.94	0.85	1.01
α_y	0.96	0.87	1.05	0.10	0.06	0.16

Table C.9: Estimated Parameters Across Non-Calvo Priors, Medium-Scale Model, 1966 to 2017

	α_p Centered at 5, α_x at 1			α_r at .5, α_p at 5, α_x/α_y at 1			α_r at .5, α_p at 5, α_x/α_y at 2		
	Mode	5%	95%	Mode	5%	95%	Mode	5%	95%
λ_p	0.82	0.79	0.86	0.86	0.83	0.88	0.89	0.87	0.91
λ_w	0.84	0.80	0.88	0.82	0.76	0.86	0.86	0.81	0.89
α_i	0.79	0.60	0.90	0.52	0.36	0.65	0.51	0.36	0.65
α_p	4.92	4.74	5.15	4.94	4.75	5.16	5.01	4.79	5.20
α_x	0.89	0.81	0.96	0.91	0.83	0.98	1.93	1.84	2.01
α_y	0.25	0.19	0.31	0.99	0.91	1.07	2.00	1.91	2.08

D.1 Full Model with Credit Channel

Household problem The economy consists of a continuum of ex ante identical islands of measure 1 that belong to a trading bloc in a monetary union. Consumers on each island derive utility from the consumption of a final good, leisure, and housing. Let s index an individual island and $p_t(s)$ denote the price of the final consumption good. Individual households on each island belong to labor unions that sell differentiated varieties of labor. We assume perfect risk-sharing across households belonging to different labor unions on a given island. Labor is immobile across islands and the housing stock on each island is in fixed supply. The problem of a household that belongs to labor union ι is to

$$\max \mathbb{E}_0 \sum_{t=0}^{\infty} \left(\prod_{j=0}^{t-1} \beta_j(s) \right) \left[\int_0^1 v_{it}(s) \log\left(c_{it}(s)\right) \mathrm{d}i + \eta_t^h(s) \log\left(h_t(s)\right) - \frac{\eta_t^n(s)}{1+\nu} n_t(\iota, s)^{1+\nu} \right] \tag{30}$$

where $h_t(s)$ is the amount of housing the household owns, $n_t(\iota, s)$ is the amount of labor it supplies, and $c_{it}(s)$ is the consumption of an individual member i. The term $v_{it}(s) \geq 1$ represents a taste shifter, an i.i.d random variable drawn from a Pareto distribution:

$$\Pr(v_{it}(s) \leq v) = F(v) = 1 - v^{-\alpha}. \tag{31}$$

Here, $\alpha > 1$ determines the amount of uncertainty about v. A lower α implies more uncertainty. The terms $\eta_t^h(s)$ and $\eta_t^n(s)$ affect the preference for housing and the disutility from work, while $\beta_t(s)$ is the household's one-period-ahead discount factor. We assume that each of these preference shifters have an island-specific component and an aggregate component, all of which follow AR(1) processes with independent Gaussian innovations. The household's budget constraint is:

$$p_t(s)x_t(s) + e_t(s)(h_{t+1}(s) - h_t(s)) = w_t(\iota, s)n_t(\iota, s) + q_t l_t(s) - b_t(s) + (1 + \gamma q_t)a_t(s) + T_t(\iota, s), \tag{32}$$

where $x_t(s)$ are transfers made to individual members in the goods market, $e_t(s)$ is the price of housing, $w_t(\iota, s)$ is the wage rate, and $T_t(\iota, s)$ collects the profits households earn from their ownership of intermediate goods firms, transfers from the government aimed at correcting the steady state markup distortion, and the transfers stemming from the risk-sharing arrangement.[8] We let $a_t(s)$ denote the amount of coupon payments the household is entitled to receive in period t, $b_t(s)$ the amount it must repay, and q_t the economy-wide price of the securities

[8] We assume that households on island s exclusively own firms on that particular island.

described below. Thus, $q_t a_t(s)$ represents the household's total asset holdings (savings), while $q_t b_t(s)$ represents its outstanding debt. We describe a household's holdings of the security by recording the amount of coupon payments b_t that the household has to make period t. Letting $l_t(s)$ denote the amount of securities the household sells in period t, the date $t+1$ coupon payments are

$$b_{t+1}(s) = \sum_{i=0}^{\infty} \gamma^i l_{t-i}(s) = l_t(s) + \gamma b_t(s). \tag{33}$$

The household also faces a liquidity constraint limiting the consumption of an individual member to be below the amount of real balances the member holds:

$$p_t(s) c_{it}(s) \leq p_t(s) x_t(s). \tag{34}$$

The household also faces a borrowing constraint

$$q_t l_t(s) \leq m_t(s) e_t(s) h_{t+1}(s). \tag{35}$$

The law of motion for a household's assets is given by

$$q_t a_{t+1}(s) = p_t(s) \left(x_t(s) - \int_0^1 c_{it}(s) \mathrm{d}i \right). \tag{36}$$

There are no barriers to capital flows, so all islands trade securities at a common price q_t. The credit limit $m_t(s)$ evolves as the product of an island-specific and aggregate component, both of which are AR(1) processes with Gaussian disturbances.

At this point, we note that as $\alpha \to \infty$, $v_{it}(s) \to 1$. In this case, there is no idiosyncratic uncertainty. There is no meaningful role for the liquidity constraints and, since housing is separable in the utility function and exogenously fixed, there is no role for credit, and the economy collapses to the standard 3-equation New Keynesian model (see Jones, Midrigan and Philippon, 2018, for details and a discussion of this point).

Final goods producers Final goods producers on island s produce $y_t(s)$ units of the final good using $y_t^N(s)$ units of non-tradable goods produced locally and $y_t^M(s,j)$ units of tradable goods produced on island j and imported to island s:

$$y_t(s) = \left(\omega^{\frac{1}{\sigma}} y_t^N(s)^{\frac{\sigma-1}{\sigma}} + (1-\omega)^{\frac{1}{\sigma}} \left(\int_0^1 y_t^M(s,j)^{\frac{\kappa-1}{\kappa}} \mathrm{d}j \right)^{\frac{\kappa}{\kappa-1}\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{37}$$

where ω determines the share of non-traded goods, σ is the elasticity of substitution between traded and non-traded goods and κ is the elasticity of substitution between varieties of the traded goods produced on different islands. Letting $p_t^N(s)$ and $p_t^M(s)$ denote the prices of these goods on island s, the final goods price on an island is

$$p_t(s) = \left(\omega p_t^N(s)^{1-\sigma} + (1-\omega)\left(\int_0^1 p_t^M(j)^{1-\kappa}\mathrm{d}j\right)^{\frac{1-\sigma}{1-\kappa}}\right)^{\frac{1}{1-\sigma}}. \tag{38}$$

The demand for non-tradable intermediate goods produced on an island is

$$y_t^N(s) = \omega\left(\frac{p_t^N(s)}{p_t(s)}\right)^{-\sigma} y_t(s), \tag{39}$$

while demand for an island's tradable exports $y_t^X(s)$ is an aggregate of what all other islands purchase:

$$y_t^X(s) = (1-\omega)p_t^M(s)^{-\kappa}\left(\int_0^1 p_t^M(j)^{1-\kappa}\mathrm{d}j\right)^{\frac{\kappa-\sigma}{1-\kappa}}\left(\int_0^1 p_t(j)^{\sigma} y_t(j)\mathrm{d}j\right). \tag{40}$$

Intermediate goods producers Traded and non-traded goods on each island are themselves CES composites of varieties of differentiated intermediate inputs with an elasticity of substitution ϑ. The demand for an individual variety k for non-tradeable goods (for example) are

$$y_t^N(s,k) = \left(p_t^N(s,k)/p_t^N(s)\right)^{-\vartheta} y_t^N(s).$$

Individual producers of intermediate goods are subject to Calvo price adjustment frictions. Let λ_p denote the probability that a firm does not reset its price in a given period. A firm that resets its price maximizes the present discounted flow of profits weighted by the probability that the price it chooses at t will still be in effect at any particular date. As was the case earlier, the production function is linear in labor, but it is now subject to sector-specific productivity disturbances $z_t^N(s)$ and $z_t^X(s)$, so that

$$y_t^j(s,k) = z_t^j(s) n_t^j(s,k), \text{ for } j \in \{N, X\}$$

so that the unit cost of production is simply $w_t(s)/z_t^j(s)$ in both sectors.

For example, a traded intermediate goods firm that resets its price solves

$$\max_{p_t^{X*}(s)} \sum_{j=0}^{\infty}\left(\lambda_p^j \prod_{i=0}^{j-1}\beta_{t+i}(s)\right)\mu_{t+j}(s)\left(p_t^{X*}(s) - \tau_p\frac{w_{t+j}(s)}{z_{t+j}^X(s)}\right)\left(\frac{p_t^{X*}(s)}{p_{t+j}^X(s)}\right)^{-\vartheta} y_{t+j}^X(s), \tag{41}$$

where $\mu_{t+j}(s)$ is the shadow value of wealth of the representative household on island s – that is, the multiplier on the flow budget constraint (32) – and $\tau_p = \frac{\vartheta-1}{\vartheta}$ is a tax the government levies to eliminate the steady state markup distortion. This tax is rebated lump sum to households on island s. The composite price of traded exports or non-traded goods is then a weighted average of the prices of individual differentiated intermediates. For example, the price of export goods is

$$p_t^X(s) = \left[(1-\lambda_p)p_t^{X*}(s)^{1-\vartheta} + \lambda_p p_{t-1}^X(s)^{1-\vartheta}\right]^{\frac{1}{1-\vartheta}}. \tag{42}$$

Wage setting We assume that individual households are organized in unions that supply differentiated varieties of labor. The total amount of labor services available in production is

$$n_t(s) = \left(\int_0^1 n_t(\iota,s)^{\frac{\psi-1}{\psi}} \mathrm{d}\iota\right)^{\frac{\psi}{\psi-1}}, \tag{43}$$

where ψ is the elasticity of substitution between labor varieties. Demand for an individual union's labor given its wage $w_t(\iota,s)$ is therefore $n_t(\iota,s) = (w_t(\iota,s)/w_t(s))^{-\psi} n_t(s)$. The problem of a union that resets its wage is to choose a new wage $w_t^*(s)$ to

$$\max_{w_t^*(s)} \sum_{j=0}^{\infty} \left(\lambda_w^j \prod_{i=0}^{j-1} \beta_{t+i}(s)\right) \times \tag{44}$$
$$\left[\tau_w \mu_{t+j}(s) w_t^*(s) \left(\frac{w_t^*(s)}{w_{t+j}(s)}\right)^{-\psi} n_{t+j}(s) - \frac{\eta_{t+j}^n(s)}{1+\nu}\left(\left(\frac{w_t^*(s)}{w_{t+j}(s)}\right)^{-\psi} n_{t+j}(s)\right)^{1+\nu}\right],$$

where λ_w is the probability that a given union leaves its wage unchanged and $\tau_w = \frac{\psi-1}{\psi}$ is a labor income subsidy aimed at correcting the steady state markup distortion. The composite wage at which labor services are sold to producers is

$$w_t(s) = \left[(1-\lambda_w)w_t^*(s)^{1-\psi} + \lambda_w w_{t-1}(s)^{1-\psi}\right]^{\frac{1}{1-\psi}}. \tag{45}$$

D.2 Monetary Policy

Let $y_t = \int_0^1 p_t(s) y_t(s)/p_t \,\mathrm{d}s$ be total real output in this economy, where $p_t = \int_0^1 p_t(s)\mathrm{d}s$ is the aggregate price index. Let $\pi_t = p_t/p_{t-1}$ denote the rate of inflation and

$$1 + i_t = \mathbb{E}_t R_{t+1} \tag{46}$$

be the expected nominal return on the long-term security, which we refer to as the nominal interest rate. Aggregation over the pricing choices of individual producers implies, up to a first-order approximation,

$$\log(\pi_t/\bar{\pi}) = \bar{\beta}\mathbb{E}_t \log(\pi_{t+1}/\bar{\pi}) + \frac{(1-\lambda_p)(1-\lambda_p\bar{\beta})}{\lambda_p}\left(\log(w_t) - \log(z_t)\right) + \theta_t,$$

where we add an AR(1) disturbance θ_t to individual firms' desired markups, $\bar{\beta}$ is the steady state discount factor, and $\bar{\pi}$ is the steady-state level of inflation.

We assume that monetary policy is characterized by a Taylor rule when the ZLB does not bind:

$$1 + i_t = (1 + i_{t-1})^{\alpha_i}\left[(1+\bar{\imath})\,\pi_t^{\alpha_\pi}\left(\frac{y_t}{y_t^*}\right)^{\alpha_y}\right]^{1-\alpha_i}\left(\frac{y_t/y_t^*}{y_{t-1}/y_{t-1}^*}\right)^{\alpha_x}\exp(\varepsilon_t^i),$$

where ε_t^i is a monetary policy shock; α_i determines the persistence; and α_π; α_y; and α_x determine the extent to which monetary policy responds to inflation, deviations of output from its flexible price level y_t^*, and the growth rate of the output gap, respectively. We assume that $\bar{\imath}$ is set to a level that ensures a steady state level of inflation of $\bar{\pi}$. When the ZLB binds, then

$$i_t = 0.$$

The interest rate may be at zero either because aggregate shocks cause the ZLB to bind, or because the Fed commits to keeping i_t at 0 for a longer time period than implied by the constraint. We thus implicitly assume that the Fed can manipulate expectations of how the path of interest rates evolves, as in Eggertsson and Woodford (2003) and Werning (2015). In our estimation we use survey data from the New York Federal Reserve to discipline the expected duration of the zero interest rate regime during the 2009 to 2015 period.

Since we assume that an individual island is of measure zero, monetary policy does not react to island-specific disturbances. The monetary union is closed so aggregate savings must equal aggregate debt:

$$\int_0^1 a_{t+1}(s)\mathrm{d}s = \int_0^1 b_{t+1}(s)\mathrm{d}s. \tag{47}$$

D.3 Likelihood of the State Component

We use Bayesian likelihood methods to estimate the parameters of the island economy and the shocks. We use a panel dataset across states for the state-level estimation, and aggregate data and the ZLB for the aggregate-level estimation. We formulate the state-space of the model so as to separate our estimation into a collection of regional components to make it computationally feasible.

We discuss separately the likelihood function of the state/regional component and then the likelihood function of the aggregate component.

We use Bayesian methods. We first log-linearize the model. The log-linearized model has the state space representation

$$x_t = \mathbf{J} + \mathbf{Q} x_{t-1} + \mathbf{G}\varepsilon_t \tag{48}$$

$$z_t = \mathbf{H}_t x_t. \tag{49}$$

The state vector is x_t. The error is distributed $\varepsilon_t \sim N(0, \Omega)$, where Ω is the covariance matrix of ε_t. We assume no observation error of the data z_t.

Denote by ϑ the vector of parameters to be estimated. Denote by $\mathcal{Z} = \{z_\tau\}_{\tau=1}^{T}$ the sequence of $N_z \times 1$ vectors of observable variables, combined over states. By Bayes law, the posterior $\mathcal{P}(\vartheta \mid \mathcal{Z})$ satisfies

$$\mathcal{P}(\vartheta \mid \mathcal{Z}) \propto L(\mathcal{Z} \mid \vartheta) \times \mathcal{P}(\vartheta).$$

With Gaussian errors ε_t, the likelihood function $L(\mathcal{Z} \mid \vartheta)$ is computed using the sequence of structural matrices and the Kalman filter, described below:

$$\log L(\mathcal{Z} \mid \vartheta) = -\left(\frac{N_z T}{2}\right) \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \log \det \mathbf{S}_t - \frac{1}{2} \sum_{t=1}^{T} \tilde{y}_t^\top \left(\mathbf{S}_t\right)^{-1} \tilde{y}_t,$$

where $\tilde{y}_t$ is the vector of forecast errors and $\mathbf{S}_t$ is its associated covariance matrix.

D.3.1 Kalman Filter

The Kalman filter recursion is given by the following equations. The state of the system is $(\hat{x}_t, \mathbf{P}_{t-1})$. In the predict step, the structural matrices **J**, **Q** and **G** are used to compute a forecast of the state $\hat{x}_{t|t-1}$ and the forecast covariance matrix $\mathbf{P}_{t|t-1}$ as

$$\begin{aligned} \hat{x}_{t|t-1} &= \mathbf{J} + \mathbf{Q}\hat{x}_t \\ \mathbf{P}_{t|t-1} &= \mathbf{Q}\mathbf{P}_{t-1}\mathbf{Q}^\top + \mathbf{G}\Omega\mathbf{G}^\top. \end{aligned} \tag{50}$$

We update these forecasts with imperfect observations of the state vector. This update step involves computing forecast errors $\tilde{y}_t$ and their associated covariance matrix $\mathbf{S}_t$ as

$$\begin{aligned} \tilde{y}_t &= z_t - \mathbf{H}_t \hat{x}_{t|t-1} \\ \mathbf{S}_t &= \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top. \end{aligned}$$

The Kalman gain matrix is given by

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^\top \mathbf{S}_t^{-1}.$$

With $\tilde{y}_t$, $\mathbf{S}_t$ and $\mathbf{K}_t$ in hand, the optimal filtered update of the state x_t is

$$\hat{x}_t = \hat{x}_{t|t-1} + \mathbf{K}_t\tilde{y}_t,$$

and for its associated covariance matrix,

$$\mathbf{P}_t = (I - \mathbf{K}_t\mathbf{H}_t)\,\mathbf{P}_{t|t-1}.$$

The Kalman filter is initialized with x_0 and $\mathbf{P}_0$ determined from their unconditional moments and is computed until the final time period T of data. We can show that the stationary $\mathbf{P}_0$ has the expression

$$\operatorname{vec}(\mathbf{P}_0) = (\mathbf{I} - \mathbf{Q}\otimes\mathbf{Q})^{-1}\operatorname{vec}(\mathbf{G}\Omega\mathbf{G}^\top) \tag{51}$$

D.3.2 Kalman Smoother

With the estimates of the parameters on a sample up to time period T, the Kalman smoother gives an estimate of $x_{t|T}$, or an estimate of the state vector at each point in time given all available information. With $\hat{x}_{t|t-1}$, $\mathbf{P}_{t|t-1}$, $\mathbf{K}_t$, and $\mathbf{S}_t$ in hand from the Kalman filter, the vector $x_{t|T}$ is computed by

$$x_{t|T} = \hat{x}_{t|t-1} + \mathbf{P}_{t|t-1}r_{t|T},$$

where the vector $r_{T+1|T} = 0$ and is updated with the recursion

$$r_{t|T} = \mathbf{H}_t^\top\mathbf{S}_t^{-1}\left(z_t - \mathbf{H}_t\hat{x}_{t|t-1}\right) + (I - \mathbf{K}_t\mathbf{H}_t)^\top\,\mathbf{P}_{t|t-1}^\top r_{t+1|T}.$$

Finally, to get an estimate of the shocks to each state variable under this model's shock structure, denoted by e_t, we can compute

$$e_t = \mathbf{G}\varepsilon_t = \mathbf{G}r_{t|T}.$$

D.3.3 Block Structure

The regional component of the model has a block structure separated by state. For example, consider two states so that the log-linearized state-space representation is

$$\begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} = \begin{bmatrix} \mathbf{J}^1 \\ \mathbf{J}^2 \end{bmatrix} + \begin{bmatrix} \mathbf{Q}^1 & 0 \\ 0 & \mathbf{Q}^2 \end{bmatrix}\begin{bmatrix} x_{t-1}^1 \\ x_{t-1}^2 \end{bmatrix} + \begin{bmatrix} \mathbf{G}^1 & 0 \\ 0 & \mathbf{G}^2 \end{bmatrix}\begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix}$$

Under this block structure, the forecast covariance matrix $P_{t|t-1}$ also has a block structure. This is clear from the expressions (50) and (51).

The block structure is also helpful for computational reasons. The log-likelihood becomes a weighted sum of state-by-state log-likelihood functions. To show this: because $\mathbf{P}_{t|t-1}$ has a block structure, so does $\mathbf{S}_t$. And because $\mathbf{S}_t$ has a block structure

$$\log \det \mathbf{S}_t = \log \prod_j \det \mathbf{S}_t^j = \sum_j \log \det \mathbf{S}_t^j.$$

Also, because $\mathbf{S}_t$ has a block structure, so does its inverse, so that the last term in the log-likehood can also be separated by state. The log-likelihood is then

$$\log L(\mathcal{Z} \mid \vartheta) = \sum_s \log L^s(\mathcal{Z}^s \mid \vartheta).$$

D.4 Likelihood of the Aggregate Component

D.4.1 Solution with Zero Lower Bound

The state-space of the model is

$$\begin{aligned} x_t &= \mathbf{J}_t + \mathbf{Q}_t x_{t-1} + \mathbf{G}_t \varepsilon_t \\ z_t &= \mathbf{H}_t x_t. \end{aligned}$$

The reduced-form matices $\mathbf{J}_t$ and $\mathbf{G}_t$ are time-varying because of the occasionally-binding ZLB (see Kulish et al., 2017), and the observation equation is time-varying because the nominal interest rate becomes unobserved when it is at its bound. A sequence of ZLB durations maps to a sequence of reduced-form matrices.

Denote by ϑ the vector of parameters to be estimated and by $\mathbf{T}$ the vector of ZLB durations that are observed each period. Denote by $\mathcal{Z} = \{z_\tau\}_{\tau=1}^{T}$ the sequence of vectors of observable variables. With Gaussian errors, the likelihood function $L(\mathcal{Z}, \mathbf{T} \mid \vartheta)$ for the aggregate component is computed using the sequence of structural matrices associated with the sequence of ZLB durations, and the Kalman filter:

$$\log L(\mathcal{Z}, \mathbf{T} \mid \vartheta) = -\left(\frac{N_z T}{2}\right) \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \log \det \mathbf{H}_t \mathbf{S}_t \mathbf{H}_t^\top - \frac{1}{2} \sum_{t=1}^{T} \tilde{y}_t^\top \left(\mathbf{H}_t \mathbf{S}_t \mathbf{H}_t^\top\right)^{-1} \tilde{y}_t.$$

D.4.2 Kalman filter

The state of the system is $(\hat{x}_t, \mathbf{P}_{t-1})$. In the predict step, the structural matrices $\mathbf{J}_t$, $\mathbf{Q}_t$, and $\mathbf{G}_t$ are used to compute a forecast of the state $\hat{x}_{t|t-1}$ and the forecast covariance matrix $\mathbf{P}_{t|t-1}$ as

$$\hat{x}_{t|t-1} = \mathbf{J}_t + \mathbf{Q}_t \hat{x}_t$$
$$\mathbf{P}_{t|t-1} = \mathbf{Q}_t \mathbf{P}_{t-1} \mathbf{Q}_{t|t-1}^\top + \mathbf{G}_t \Omega \mathbf{G}_t^\top.$$

This formulation differs from the time-invariant Kalman filter used at the state level, because in the forecast stage, the matrices $\mathbf{J}_t$, $\mathbf{Q}_t$ and $\mathbf{G}_t$ can vary over time. We update these forecasts with imperfect observations of the state vector. This update step involves computing forecast errors $\tilde{y}_t$ and its associated covariance matrix $\mathbf{S}_t$ as

$$\tilde{y}_t = z_t - \mathbf{H}_t \hat{x}_{t|t-1}$$
$$\mathbf{S}_t = \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top.$$

The Kalman gain matrix is given by

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\top \mathbf{S}_t^{-1}.$$

With $\tilde{y}_t$, $\mathbf{S}_t$, and $\mathbf{K}_t$ in hand, the optimal filtered update of the state x_t is

$$\hat{x}_t = \hat{x}_{t|t-1} + \mathbf{K}_t \tilde{y}_t,$$

and for its associated covariance matrix:

$$\mathbf{P}_t = (I - \mathbf{K}_t \mathbf{H}_t)\, \mathbf{P}_{t|t-1}.$$

The Kalman filter is initialized with x_0 and $\mathbf{P}_0$ determined from their unconditional moments and is computed until the final time period T of data.

D.4.3 Kalman Smoother

With the estimates of the parameters and durations in hand at time period T, the Kalman smoother gives an estimate of $x_{t|T}$, or an estimate of the state vector at each point in time given all available information (Hamilton, 1994). With $\hat{x}_{t|t-1}$, $\mathbf{P}_{t|t-1}$, $\mathbf{K}_t$ and $\mathbf{S}_t$ in hand from the Kalman filter, the vector $x_{t|T}$ is computed by

$$x_{t|T} = \hat{x}_{t|t-1} + \mathbf{P}_{t|t-1} r_{t|T},$$

where the vector $r_{T+1|T} = 0$ and is updated with the recursion:

$$r_{t|T} = \mathbf{H}_t^\top \mathbf{S}_t^{-1} \left(z_t - \mathbf{H}_t \hat{x}_{t|t-1} \right) + (I - \mathbf{K}_t \mathbf{H}_t)^\top \mathbf{P}_{t|t-1}^\top r_{t+1|T}.$$

Finally, to get an estimate of the shocks to each state variable under this model's shock structure, denoted by e_t, we compute:

$$e_t = \mathbf{G}_t \varepsilon_t = \mathbf{G}_t r_{t|T}.$$

D.5 Posterior Sampler

This section describes the sampler used to obtain the posterior distribution of interest. We compute the likelihood function at the state level and the aggregate level, together with the prior. The posterior of our full model $\mathcal{P}(\vartheta \mid \mathbf{T}, \mathcal{Z})$ satisfies

$$\mathcal{P}(\vartheta \mid \mathbf{T}, \mathcal{Z}) \propto L(\mathcal{Z}, \mathbf{T} \mid \vartheta) \times \mathcal{P}(\vartheta).$$

We use a Markov Chain Monte Carlo procedure to sample from the posterior. It has a single block, corresponding to the parameters ϑ.[9] The sampler at step j is initialized with the last accepted draw of the structural parameters ϑ_{j-1}.

First, start by selecting which parameters to propose new values. For those parameters, draw a new proposal ϑ_j from a proposal density centered at ϑ_{j-1} and with thick tails to ensure sufficient coverage of the parameter space and an acceptance rate of roughly 20% to 25%. The proposal ϑ_j is accepted with probability $\frac{\mathcal{P}(\vartheta_j|\mathbf{T},\mathcal{Z})}{\mathcal{P}(\vartheta_{j-1}|\mathbf{T},\mathcal{Z})}$. If ϑ_j is accepted, then set $\vartheta_{j-1} = \vartheta_j$.

E Description of Data Used in Structural Estimation

E.1 State Level

We use the MSA-level inflation data, described above, and map the 27 MSA regions into 20 states with the mapping in Table E.10. For states which contain multiple MSA regions (for example, Cincinnati and Cleveland are both in Ohio), we select only the data of one of the MSA regions.

For the other state-level data series, we use state-level data on employment, output, and compensation. The observed state data are annual. To construct the data, we first take each

[9] It is worth noting that as in Kulish, Morley and Robinson (2017), in addition to the structural parameters, one can estimate the expected zero lower bound durations, in which case an additional block is needed in the posterior sampler.

Table E.10: MSA to State Mapping

State	MSA
AK	Anchorage
AZ	Phoenix
CA	Los Angeles
CO	Denver
FL	Miami
GA	Atlanta
HI	Honolulu
IL	Chicago
KS	Kansas City
MA	Tampa
MD	Baltimore
MI	Detroit
MO	St. Louis
NY	New York
OH	Cincinnati
OR	Portland
PA	Philadelphia
TX	Dallas
WA	Seattle
WI	Minneapolis

state's series relative to its initial value, compute the devation of each state's observation from the state mean, regress that series on time dummies, weighted by the state's relative population, and work with the residuals. We then take out a linear trend from the resulting series, for each subsample studied.

Main estimations Here, we provide more details on each series.

- Output: We use state-level data on Gross Domestic Product in current dollars. (BEA SAGDP2S). The data are available for download at the BEA website.

- Employment: We use state-level data on total employment from the BEA annual table SA4. In our empirical analysis, we scale this measure of employment by each state's population.

- Labor Compensation: We use state-level data on compensation of employees from the BEA annual table SA6N.

- Wages: To construct our wages series, we divide total labor compensation by the number of employed individuals, using the two series described above.

- Population: We use state-level data on population from the BEA annual table SA1-3.

E.2 Aggregate Level

At the aggregate level, we use the GDP deflator for inflation, employment, output, wages, the Fed Funds rates, and ZLB durations from NY Federal Reserve Survey Data. The codes for each raw data series are as follows:

- Gross Domestic Product: Implicit Price Deflator (GDPDEF).

- Gross Domestic Product: (GDP).

- Cumulated nonfarm business section compensation (PRS85006062) minus employment growth (PRS85006012) and deflated by the GDP deflator.

- Total employment net of construction, over the civilian noninstitutional population.

Fed Funds rate: the interest rate is the Federal Funds Rate, taken from the Federal Reserve Economic Database.

ZLB Durations: we follow the approach of Kulish, Morley and Robinson (2017) and use the ZLB durations extracted from the New York Federal Reserve Survey of Primary Dealers, conducted eight times a year from 2011Q1 onwards.[10] We take the mode of the distribution implied by these surveys. Before 2011, we use responses from the Blue Chip Financial Forecasts survey.

F Description of the Smets and Wouters (2007) Model

We list here the linearized equations of the Smets and Wouters (2007) model. We use the same notation for variables and parameters as in that paper. A full description of the model is available there and its accompanying Appendix.

[10] See the website here. For example, in the survey conducted on January 18 2011, one of the questions asked was: "Of the possible outcomes below, please indicate the percent chance you attach to the timing of the first federal funds target rate increase" (Question 2b). Responses were given in terms of a probability distribution across future quarters.

F.1 Sticky price economy

Factor prices:

$$mc_t = \alpha r_t + (1-\alpha) w_t - \epsilon_{a,t} \tag{52}$$

$$r_t = w_t + l_t - k_t^s \tag{53}$$

$$z_t = \tfrac{1-\psi}{\psi} r_t \tag{54}$$

Investment:

$$i_t = \tfrac{1}{1+\bar{\beta}\gamma} \left(i_{t-1} + \bar{\beta}\gamma \mathbb{E}_t i_{t+1} + \tfrac{1}{\gamma^2 \phi} q_t \right) + \epsilon_{i,t} \tag{55}$$

$$q_t = \tfrac{\sigma_c(1+\lambda/\gamma)}{1-\lambda/\gamma} \epsilon_{b,t} + \tfrac{1-\delta}{1-\delta+R^k} \mathbb{E}_t q_{t+1} + \tfrac{R^k}{1-\delta+R^k} \mathbb{E}_t r_{t+1} - r_t + \pi_{t+1} \tag{56}$$

Consumption:

$$c_t = \epsilon_{b,t} + \tfrac{\lambda/\gamma}{1+\lambda/\gamma} c_{t-1} + \tfrac{1}{1+\lambda/\gamma} \mathbb{E}_t c_{t+1} + \tfrac{(\sigma_c - 1) W^* L^* / C^*}{\sigma_c, (1+\lambda/\gamma)} \left(l_t - \mathbb{E}_t l_{t+1} \right) - \tfrac{1-\lambda/\gamma}{\sigma_c, (1+\lambda/\gamma)} \left(r_t - \mathbb{E}_t \pi_{t+1} \right) \tag{57}$$

Resource constraint:

$$y_t = c_t c_y + i_t i_y + \epsilon_{g,t} + z_t z_y \tag{58}$$

Production function:

$$y_t = \phi_p \left(\epsilon_{a,t} + \alpha k_t^s + (1-\alpha) \ l_t \right) \tag{59}$$

$$k_t^s = z_t + k_{t-1} \tag{60}$$

Monetary policy rule:

$$r_t = (1-\alpha_i) \alpha_p \pi_t + (1-\alpha_i) \alpha_y \left(y_t - y_t^f \right) + \alpha_x \left(y_t - y_t^f - \left(y_{t-1} - y_{t-1}^f \right) \right) + \alpha_i r_{t-1} + \epsilon_{r,t} \tag{61}$$

Evolution of capital:

$$k_t = (1 - i_k) \ k_{t-1} + i_k i_t + \epsilon_{i,t} \, \phi \, \gamma^2 \, i_k \tag{62}$$

Price and wage Philips curves:

$$\pi_t = \tfrac{1}{1+\bar{\beta}\gamma \, \iota_p} \left(\bar{\beta}\gamma \, \mathbb{E}_t \pi_{t+1} + \iota_p \, \pi_{t-1} + mc_t \, \frac{\frac{(1-\xi_p)\left(1-\bar{\beta}\gamma\,\xi_p\right)}{\xi_p}}{1+\left(\phi_p - 1\right)\epsilon_p} \right) + \epsilon_{p,t} \tag{63}$$

$$w_t = w_1 w_{t-1} + w_2 \mathbb{E}_t w_{t+1} + w_3 \pi_{t-1} - w_4 \pi_t + w_2 \mathbb{E}_t \pi_{t+1} + w_5 \left(\sigma_l \, l_t + \tfrac{1}{1-\lambda/\gamma} c_t - \tfrac{\lambda/\gamma}{1-\lambda/\gamma} c_{t-1} - w_t \right) + \epsilon_{w,t} \tag{64}$$

where $w_1 = \frac{1}{1+\bar{\beta}\gamma}$, $w_2 = \frac{\bar{\beta}\gamma}{1+\bar{\beta}\gamma}$, $w_3 = \frac{\iota_w}{1+\bar{\beta}\gamma}$, $w_4 = \frac{1+\bar{\beta}\gamma\,\iota_w}{1+\bar{\beta}\gamma}$, and $w_5 = \frac{(1-\xi_w)\left(1-\bar{\beta}\gamma\,\xi_w\right)}{\left(1+\bar{\beta}\gamma\right)\xi_w} \frac{1}{1+(\phi_w - 1)\epsilon_w}$.

F.2 Flexible price economy

The corresponding equations defining the flexible price economy are:

$$\epsilon_{a,t} = \alpha r_t^f + (1-\alpha) w_t^f \tag{65}$$

$$r_t^f = w_t^f + l_t^f - k_t^f \tag{66}$$

$$z_t^f = \tfrac{1-\psi}{\psi} r_t^f \tag{67}$$

$$k_t^f = z_t^f + kp_{t-1}^f \tag{68}$$

$$i_t^f = \tfrac{1}{1+\bar{\beta}\gamma} \left(i_{t-1}^f + \bar{\beta}\gamma \, \mathbb{E}_t i_{t+1}^f + \tfrac{1}{\gamma^2 \phi} \, q_t^f \right) + \varepsilon_{i,t} \tag{69}$$

$$q_t^f = \tfrac{1-\delta}{1-\delta+Rk} \, \mathbb{E}_t q_{t+1}^f + \tfrac{R^k}{1-\delta+R^k} \, \mathbb{E}_t r k_{t+1}^f - rr_t^f + \tfrac{\sigma_c,(1+\lambda/\gamma)}{1-\lambda/\gamma} \, \epsilon_{b,t} \tag{70}$$

$$c_t^f = \epsilon_{b,t} + \tfrac{\lambda/\gamma}{1+\lambda/\gamma} \, c_{t-1}^f + \tfrac{1}{1+\lambda/\gamma} \, \mathbb{E}_t c_{t+1}^f + \tfrac{(\sigma_c - 1) \, W^* L^* / C^*}{\sigma_c,(1+\lambda/\gamma)} \left(l_t^f - \mathbb{E}_t l_{t+1}^f \right) - \tfrac{1-\lambda/\gamma}{\sigma_c,(1+\lambda/\gamma)} rr_t^f \tag{71}$$

$$y_t^f = c_t^f c_y + i_t^f i_y + \epsilon_{g,t} + z_t^f z_y \tag{72}$$

$$y_t^f = \phi_p \left(\epsilon_{a,t} + \alpha k_t^f + (1-\alpha) \; l_t^f \right) \tag{73}$$

$$k_t^{p,f} = k_{t-1}^{p,f} \; (1 - i_k) + i_t^f \, i_k + \epsilon_{i,t} \, \gamma^2 \, \phi \, i_k \tag{74}$$

$$w_t^f = \sigma_l l_t^f + \tfrac{1}{1-\lambda/\gamma} c_t^f - \tfrac{\lambda/\gamma}{1-\lambda/\gamma} c_{t-1}^f \tag{75}$$

F.3 Shocks

$$\epsilon_{a,t} = \rho_a \, \epsilon_{a,t-1} + \sigma_a \, \eta_{a,t} \tag{76}$$

$$\epsilon_{b,t} = \rho_b \, \epsilon_{b,t-1} + \sigma_b \, \eta_{b,t} \tag{77}$$

$$\epsilon_{g,t} = \rho_g \, \epsilon_{g,t-1} + \sigma_g \, \eta_{g,t} + \eta_{a,t} \, \sigma_a \, \rho_{ga} \tag{78}$$

$$\epsilon_{i,t} = \rho_i \, \epsilon_{i,t-1} + \sigma_i \, \eta_{i,t} \tag{79}$$

$$\epsilon_{r,t} = \rho_r \, \epsilon_{r,t-1} + \sigma_r \, \eta_{r,t} \tag{80}$$

$$\epsilon_{p,t} = \rho_p \, \epsilon_{p,t-1} + \eta_{p,ma,t} - \mu_p \, \eta_{p,ma,t-1} \tag{81}$$

$$\eta_{p,ma,t} = \sigma_p \, \eta_{p,t} \tag{82}$$

$$\epsilon_{w,t} = \rho_w \, \epsilon_{w,t-1} + \eta_{w,ma,t} - \mu_w \, \eta_{w,ma,t-1} \tag{83}$$

$$\eta_{w,ma,t} = \sigma_w \, \eta_{w,t} \tag{84}$$

F.4 Measurement equations

$$dy_t = \bar{\gamma} + y_t - y_{t-1} \tag{85}$$

$$dc_t = \bar{\gamma} + c_t - c_{t-1} \tag{86}$$

$$di_t = \bar{\gamma} + i_t - i_{t-1} \tag{87}$$

$$dw_t = \bar{\gamma} + w_t - w_{t-1} \tag{88}$$

$$\pi_t^{obs} = \bar{\pi} + \pi_t \tag{89}$$

$$r_t^{obs} = \bar{r} + r_t \tag{90}$$

$$l_t^{obs} = \bar{l} + l_t \tag{91}$$

G Additional analytical results

Proof that the system of first order conditions to maximize the likelihood function has a continuuum of solutions.

The partial derivatives of the likelihood with respect to the four parameters are

$$\mathcal{L}_\kappa = T\left[\frac{\phi}{(1+\kappa\phi)}\right] + \frac{1}{\sigma_\pi^2}\sum_{t=1}^{T}(\pi_t - \kappa x_t)\, x_t$$

$$\mathcal{L}_\phi = T\left[\frac{\kappa}{(1+\kappa\phi)}\right] - \frac{1}{\sigma_x^2}\sum_{t=1}^{T}(\phi\pi_t + x_t)\, \pi_t$$

$$\mathcal{L}_{\sigma_\pi} = -T\left[\frac{1}{\sigma_\pi}\right] + \frac{1}{\sigma_\pi^3}\sum_{t=1}^{T}(\pi_t - \kappa x_t)^2$$

$$\mathcal{L}_{\sigma_x} = -T\left[\frac{1}{\sigma_x}\right] + \frac{1}{\sigma_x^3}\sum_{t=1}^{T}(\phi\pi_t + x_t)^2$$

So, if we set them to zero, we have

$$\begin{aligned}
T\left[\frac{\phi}{(1+\kappa\phi)}\right] &= -\sum_{t=1}^{T}\left(\frac{\pi_t - \kappa x_t}{\sigma_\pi}\right)\frac{x_t}{\sigma_\pi} \\
T\left[\frac{\kappa}{(1+\kappa\phi)}\right] &= \sum_{t=1}^{T}\left(\frac{\phi\pi_t + x_t}{\sigma_x}\right)\frac{\pi_t}{\sigma_x} \\
T\left[\frac{1}{\sigma_\pi}\right] &= \sum_{t=1}^{T}\left(\frac{\pi_t - \kappa x_t}{\sigma_\pi}\right)^2\frac{1}{\sigma_\pi} \\
T\left[\frac{1}{\sigma_x}\right] &= \sum_{t=1}^{T}\left(\frac{\phi\pi_t + x_t}{\sigma_x}\right)^2\frac{1}{\sigma_x}
\end{aligned}$$

which can be written

$$
\begin{aligned}
T\left[\frac{1}{(1+\kappa\phi)}\right] &= -\sum_{t=1}^{T}\left(\frac{\pi_t-\kappa x_t}{\sigma_\pi}\right)\frac{x_t}{\phi\sigma_\pi} \\
T\left[\frac{1}{(1+\kappa\phi)}\right] &= \sum_{t=1}^{T}\left(\frac{\phi\pi_t+x_t}{\sigma_x}\right)\frac{\pi_t}{\kappa\sigma_x} \\
T &= \sum_{t=1}^{T}\left(\frac{\pi_t-\kappa x_t}{\sigma_\pi}\right)^2 \\
T &= \sum_{t=1}^{T}\left(\frac{\phi\pi_t+x_t}{\sigma_x}\right)^2
\end{aligned}
$$

or

$$
\begin{aligned}
T\sigma_\pi^2 &= -\sum_{t=1}^{T}(\pi_t-\kappa x_t)\frac{x_t}{\phi}(1+\kappa\phi) \\
T\sigma_x^2 &= \sum_{t=1}^{T}(\phi\pi_t+x_t)\frac{\pi_t}{\kappa}(1+\kappa\phi) \\
T\sigma_\pi^2 &= \sum_{t=1}^{T}(\pi_t-\kappa x_t)^2 \\
T\sigma_x^2 &= \sum_{t=1}^{T}(\phi\pi_t+x_t)^2
\end{aligned}
$$

The first two imply

$$
\begin{aligned}
T\sigma_\pi^2 &= -\sum_{t=1}^{T}(\pi_t-\kappa x_t)\frac{x_t}{\phi}(1+\kappa\phi) = -\frac{1}{\phi}\sum_{t=1}^{T}\left(\pi_t x_t-\kappa x_t^2\right)(1+\kappa\phi) \\
T\sigma_x^2 &= \sum_{t=1}^{T}(\phi\pi_t+x_t)\frac{\pi_t}{\kappa}(1+\kappa\phi) = \frac{1}{\kappa}\sum_{t=1}^{T}\left(\phi\pi_t^2+x_t\pi_t\right)(1+\kappa\phi)
\end{aligned}
$$

or, reversing the sig in the first equation

$$
\begin{aligned}
T\sigma_\pi^2 &= -\sum_{t=1}^{T}(\pi_t-\kappa x_t)\frac{x_t}{\phi}(1+\kappa\phi) = \frac{1}{\phi}\sum_{t=1}^{T}\left(\kappa x_t^2-\pi_t x_t\right)(1+\kappa\phi) \\
T\sigma_x^2 &= \sum_{t=1}^{T}(\phi\pi_t+x_t)\frac{\pi_t}{\kappa}(1+\kappa\phi) = \frac{1}{\kappa}\sum_{t=1}^{T}\left(\phi\pi_t^2+x_t\pi_t\right)(1+\kappa\phi)
\end{aligned}
$$

Then, reversing the difference in the third equation, we can write the system as

$$\begin{aligned}
T\sigma_\pi^2 &= \frac{1}{\phi}\sum_{t=1}^{T}\left(\kappa x_t^2 - \pi_t x_t\right)(1+\kappa\phi) \\
T\sigma_x^2 &= \frac{1}{\kappa}\sum_{t=1}^{T}\left(\phi \pi_t^2 + x_t \pi_t\right)(1+\kappa\phi) \\
T\sigma_\pi^2 &= \sum_{t=1}^{T}(\kappa x_t - \pi_t)^2 \\
T\sigma_x^2 &= \sum_{t=1}^{T}(\phi \pi_t + x_t)^2
\end{aligned}$$

We can write them as

$$\begin{aligned}
\sigma_\pi^2 &= \left(\kappa\frac{1}{T}\sum_{t=1}^{T} x_t^2 - \frac{1}{T}\sum_{t=1}^{T}\pi_t x_t\right)\frac{(1+\kappa\phi)}{\phi} \\
\sigma_x^2 &= \left(\phi\frac{1}{T}\sum_{t=1}^{T} \pi_t^2 + \frac{1}{T}\sum_{t=1}^{T} x_t \pi_t\right)\frac{(1+\kappa\phi)}{\kappa} \\
\sigma_\pi^2 &= \kappa^2\frac{1}{T}\sum_{t=1}^{T} x_t^2 + \frac{1}{T}\sum_{t=1}^{T}\pi_t^2 - 2\kappa\frac{1}{T}\sum_{t=1}^{T} x_t\pi_t \\
\sigma_x^2 &= \phi^2\frac{1}{T}\sum_{t=1}^{T} \pi_t^2 + \frac{1}{T}\sum_{t=1}^{T} x_t^2 + 2\phi\frac{1}{T}\sum_{t=1}^{T}\pi_t x_t
\end{aligned}$$

Let

$$\frac{1}{T}\sum_{t=1}^{T} x_t^2 = V_x,\ \frac{1}{T}\sum_{t=1}^{T}\pi_t^2 = V_\pi \text{ and } \frac{1}{T}\sum_{t=1}^{T}\pi_t x_t = C_{x,\pi}$$

so we have

$$\begin{aligned}
\sigma_\pi^2 &= (\kappa V_x - C_{x,\pi})\frac{(1+\kappa\phi)}{\phi} \\
\sigma_x^2 &= (\phi V_\pi + C_{x,\pi})\frac{(1+\kappa\phi)}{\kappa} \\
\sigma_\pi^2 &= \kappa^2 V_x + V_\pi - 2\kappa C_{x,\pi} \\
\sigma_x^2 &= \phi^2 V_\pi + V_x + 2\phi C_{x,\pi}
\end{aligned}$$

Now, use the first and the third to eliminate σ_π^2 and obtain

$$\begin{aligned}
\kappa^2 V_x + V_\pi - 2\kappa C_{x,\pi} &= (\kappa V_x - C_{x,\pi})\frac{(1+\kappa\phi)}{\phi} \\
\sigma_x^2 &= (\phi V_\pi + C_{x,\pi})\frac{(1+\kappa\phi)}{\kappa} \\
\sigma_x^2 &= \phi^2 V_\pi + V_x + 2\phi C_{x,\pi}
\end{aligned}$$

Then, we now manipulate the first equation, replace on the second and obtain the third.

To see this, write the first equation as

$$\kappa^2\phi V_x + \phi V_\pi - 2\phi\kappa C_{x,\pi} = \kappa V_x - C_{x,\pi} + \kappa^2\phi V_x - \kappa\phi C_{x,\pi}$$

or

$$\phi V_\pi + C_{x,\pi} = \kappa V_x + \phi\kappa C_{x,\pi}$$

or

$$\phi V_\pi \frac{1}{\kappa} + C_{x,\pi}\frac{1}{\kappa} - \phi C_{x,\pi} = V_x \tag{92}$$

Now, write the second as

$$\sigma_x^2 = \phi V_\pi \frac{1}{\kappa} + \phi^2 V_\pi + C_{x,\pi}\frac{1}{\kappa} + \phi C_{x,\pi}$$

or

$$\sigma_x^2 = \phi^2 V_\pi + 2\phi C_{x,\pi} + \phi V_\pi \frac{1}{\kappa} + C_{x,\pi}\frac{1}{\kappa} - \phi C_{x,\pi}$$

Using (92) , we obtain

$$\sigma_x^2 = \phi^2 V_\pi + 2\phi C_{x,\pi} + V_x$$

which is the fourth condition. QED