Monetary policy implications of state-dependent prices and wages

James Costain
(Banco de España)

Anton Nakov
(European Central Bank and CEPR)

Borja Petit Zarzalejos
(CEMFI)

PRELIMINARY AND INCOMPLETE
January 2016

Abstract

This paper studies the dynamic general equilibrium effects of monetary policy shocks in a "control cost" model of state-dependent retail price adjustment and state-dependent wage adjustment. Both suppliers of retail goods and suppliers of labor are monopolistic competitors subject to idiosyncratic productivity shocks and nominal rigidities. Nominal rigidity arises because precise choice is costly: decision-makers tolerate errors both in the timing of price and wage adjustments, and in the new level at which the price or wage is set, because achieving perfect precision in these decisions would be excessively costly.

The model is calibrated to data on the size and frequency of price and wage adjustments. We find that sticky wages by themselves account for much of the nonneutrality that occurs in the model where both sticky wages and sticky prices are present. Hence, a model in which both prices and wages are sticky implies substantially larger real effects of monetary shocks than does a model with sticky prices only.

Keywords: Sticky prices, sticky wages, state-dependent adjustment, logit equilibrium, near rationality, control costs

JEL Codes: E31, D81, C73

1 Introduction[1]

The nominal rigidity of prices and/or wages is a prominent assumption in contemporary monetary macroeconomics. For reasons of analytical tractability, many models are based on Calvo's (1983) framework, in which the probability of adjustment is constant. But several influential papers have claimed that if nominal stickiness is derived from rational decision-making, instead of being imposed in an *ad hoc* way, then monetary policy ceases to have large real effects (see for example the menu cost models of Caplin and Spulber, 1987, and Golosov and Lucas, 2007). This finding has recently inspired a large literature that has investigated how the conclusions of Calvo-style models and menu cost models hold up in a variety of state-dependent pricing frameworks that are closely calibrated to microdata on retail price adjustments (*e.g.* Klenow and Kryvtsov, 2008; Gagnon, 2009; Matejka, 2010; Midrigan, 2011; Álvarez, González-Rozada, Neumeyer, and Beraja, 2011; Eichenbaum, Jaimovich, and Rebelo, 2011; Kehoe and Midrigan, 2012; Dotsey, King, and Wolman, 2013; Álvarez, Lippi, and Paciello, 2014; Costain and Nakov, 2011, 2015).

Much of this new literature concludes, to quote Kehoe and Midrigan, that "prices are sticky after all". In other words, while stripped-down menu cost models like Golosov and Lucas (2007) imply that monetary policy is nearly neutral, related models that fit retail microdata better imply that price stickiness does in fact matter at the aggregate level, leading to nontrivial real effects of monetary policy. This result is encouraging, since it improves the link between microdata and modern macroeconomics. However, it may be premature to draw strong quantitative conclusions from this literature, which revolves around models where sticky retail prices are the only form of friction. In contrast, contemporary DSGE models that aim to match business cycle dynamics well rely on multiple rigidities— not just Calvo stickiness in prices and wages, but many other frictions, such as consumption habits, investment adjustment costs, labor matching frictions, and so forth. The computationally intensive nature of recent state-dependent pricing models has restricted attention to economies with price stickiness only, ignoring any other frictions. Likewise, computational difficulties have limited most previous papers on state dependence to considering only one-time, zero-probability money supply shocks, instead of solving a fully stochastic equilibrium, which severely limits the types of policy responses that can be analyzed. This suggests that further progress in understanding the quantitative relevance of mechanisms related to nominal rigidity requires us to study models in which multiple frictions interact.

Therefore, this paper studies the implications of adding another layer of state-dependent adjustment, allowing for wage stickiness as well as price stickiness.[2] A natural point of departure for our exercise is Erceg, Henderson, and Levin's (2000) study of monopolistic retail price setters and monopolistic wage

[1]Thanks to Jordi Galí, Julián Messina, and Ernesto Villanueva for helpful comments. We are grateful to Virgiliu Midrigan for making his data available to us, and to the James M. Kilts Center at the Univ. of Chicago GSB, which is the original source of those data. Views expressed here are those of the authors and do not necessarily coincide with those of the Bank of Spain, the Eurosystem, the ECB, or the CEPR.

[2]We know of only one previous study of state-dependent prices and wages in a DSGE model, Takahashi (2015). But that paper abstracts from idiosyncratic shocks, so it cannot be closely calibrated to microdata. Takahashi's paper also differs from ours in that it analyzes a stochastic menu cost model (following Dotset *et al.*, 2009) rather than a control cost model.

setters under the Calvo framework. Following Erceg *et al.*, we set up the wage setters' problem so that it closely parallels the price setting problem, but we allow for state dependence in both decisions. More precisely, we compare a framework in which both price and wage setters are constrained by the Calvo friction to a framework in which price and wage setters are both constrained by a state-dependent friction, and in addition we compare these with scenarios in which price setting and/or wage setting is almost perfectly flexible. Note therefore that the purpose of our paper is to study the implications of different models of price stickiness and wage stickiness while abstracting from any other frictions that might affect the labor market (or other markets). While the interaction of nominal rigidities with labor market matching has been a major area of research recently, our goal in this paper is to quantify the effects of state-dependent prices and wages by themselves, leaving their interaction with matching frictions for future work.

Our model of state-dependent adjustment is an extension of the "control cost" model of price stickiness proposed by Costain and Nakov (2015), henceforth CN15. Control costs are a modeling device from game theory designed to capture the fact that decision-making is costly, and that therefore some decisions are mistakes.[3] A decision is treated as a random variable defined over a set of feasible alternatives, and the decision-maker is assumed to face a cost function that increases with the precision of that random variable. Placing probability one on the optimal alternative is a very precise decision, and therefore the decision-maker may instead economize on the costs of choice by tolerating some randomness (some errors) in the alternative chosen. CN15 models nominal rigidity by applying this framework *both* to the prices firms choose, and to firms' control of the *timing* of their adjustments. In equilibrium, in their model, managers of retail firms economize on time devoted to decision-making by tolerating some small errors in the prices they set, and some small errors in the timing of their price adjustments.

There are a number of reasons why it seems interesting to extend the CN15 framework to other frictions, beyond price stickiness. First, it describes adjustment costs in a sparsely parameterized way; the benchmark scenario in CN15 simultaneously fits many "puzzling" features of retail price setting on the basis of just two free parameters in the decision cost function. Second, these costs have an appealing interpretation: the costs of price adjustment are interpreted as time devoted by management to decision-making. These may be plausibly larger than the menu-type fixed costs associated with the physical act of changing the price, and may be compared, at least roughly, to case studies on time use by management. Third, the model is no more difficult to solve numerically than comparable menu cost models, but it is far simpler to solve than "rational inattention" models in the tradition of Sims (2003). Fourth, the mathematical structure of the model— resetting a control variable at intermittent points of time— seems applicable to many other decisions besides price adjustment, potentially allowing us to describe many margins of a general equilibrium model in a mutually consistent and mutually comparable way. Finally, since the calibration strategy in the recent state-dependent pricing literature involves matching many moments of the distribution of individual price adjustments, it stretches credulity to abstract from errors. When matching (for example) the standard deviation of observed price adjustments, inferences about the

[3] See Stahl (1990), Mattsson and Weibull (2002), or van Damme (1991), Ch. 4.

standard deviation of the underlying shocks will differ greatly depending on whether or not we insist that every single price adjustment represents a precisely optimal action.

1.1 Further discussion / Related literature

2 Model

This discrete-time model embeds near-rational price adjustment and wage adjustment in an otherwise standard New Keynesian general equilibrium framework based on Golosov and Lucas (2007). Prices are set by a continuum of monopolistically competitive firms that sell differentiated retail goods. Wages are set by a continuum of monopolistically competitive labor unions that sell differentiated types of labor. In addition, there is a representative household, and a monetary authority that sets an exogenous growth process for the nominal money supply.

2.1 Household

The household's period utility function is $u(C_t) - X(H_t) + v\left(M_t/P_t\right)$, where C_t is consumption, H_t is total time devoted to working *or* decision-making, and M_t/P_t is real money balances. The functions u and v are assumed increasing and concave. We assume the increasing, convex disutility function $X(H) = \chi\frac{H^{1+\zeta}}{1+\zeta}$. We will focus initially on the linear case $\zeta = 0$, implying $X(H) = \chi H$, which is easier to solve, but we will see that the nonlinear specification $\zeta > 0$ is necessary to match wage adjustment data. Utility is discounted by factor β per period. Consumption is a CES aggregate of differentiated products C_{jt}, with elasticity of substitution ϵ:

$$C_t = \left\{\int_0^1 C_{jt}^{\frac{\epsilon-1}{\epsilon}} dj\right\}^{\frac{\epsilon}{\epsilon-1}}. \tag{1}$$

The household's nominal period budget constraint is

$$\int_0^1 P_{jt}C_{jt}dj + M_t + R_t^{-1}B_t = \int_0^1 W_{it}N_{it}di + M_{t-1} + B_{t-1} + T_t^M + T_t^D, \tag{2}$$

Here $\int_0^1 P_{jt}C_{jt}dj$ is total nominal consumption, and $\int_0^1 W_{it}H_{it}di$ is total labor compensation received from monopolistically supplying a continuum of differentiated types of labor H_{it}. B_t represents nominal bond holdings, with interest rate $R_t - 1$; T_t^M is a lump sum transfer from the central bank, and T_t^D is a dividend payment from the firms.

Households choose $\{C_{jt}, B_t, M_t\}_{t=0}^{\infty}$ to maximize expected discounted utility, subject to the budget constraint (2).[4] They also set nominal wages intermittently, as we will discuss in section 2.3, and they

[4]We are abusing notation here for the sake of brevity. The time subscript on the household's decision variables should not be interpreted as indicating deterministic dependence on time; instead, it indicates dependence on the stochastic aggregate state of the economy.

supply labor to fulfill the demand that arises given the nominal wages they have set. Optimal consumption across the differentiated goods implies

$$C_{jt} = (P_{jt}/P_t)^{-\epsilon} C_t, \tag{3}$$

so nominal spending can be written as $P_t C_t = \int_0^1 P_{jt} C_{jt} dj$ under the price index

$$P_t \equiv \left\{ \int_0^1 P_{jt}{}^{1-\epsilon} dj \right\}^{\frac{1}{1-\epsilon}}. \tag{4}$$

The household's first-order conditions for total consumption and for money use can be written as follows:

$$R_t^{-1} = \beta E_t \left(\frac{P_t u'(C_{t+1})}{P_{t+1} u'(C_t)} \right), \tag{5}$$

$$1 - \frac{v'(M_t/P_t)}{u'(C_t)} = \beta E_t \left(\frac{P_t u'(C_{t+1})}{P_{t+1} u'(C_t)} \right). \tag{6}$$

2.2 Monopolistic firms

Each firm j produces output Y_{jt} under a constant returns technology $Y_{jt} = A_{jt} N_{jt}$. Here labor N_{jt} is the only input, and A_{jt} is an idiosyncratic productivity process that follows a time-invariant Markov process on a bounded set, $A_{jt} \in \Gamma^A \subseteq [\underline{A}, \overline{A}]$. Productivity innovations are *iid* across firms. Thus, A_{jt} is correlated with $A_{j,t-1}$, but it is uncorrelated with other firms' shocks. Firm i is a monopolistic competitor that sets a price P_{jt}, facing the demand curve $Y_{jt} = C_t P_t^{\epsilon} P_{jt}^{-\epsilon}$. We assume each firm must fulfill all demand at its chosen price. Note that since firms are infinitesimal, each firm j assumes that its own price P_{jt} has no effect on the aggregate price level P_t. It hires labor at wage rate W_t, generating profits

$$U_{jt} = P_{jt} Y_{jt} - W_t N_{jt} = \left(P_{jt} - \frac{W_t}{A_{jt}} \right) C_t P_t^{\epsilon} P_{jt}^{-\epsilon} \tag{7}$$

per period. Firms are owned by the household, so they discount nominal income between times t and $t+1$ at the rate $\beta \frac{P_t u'(C_{t+1})}{P_{t+1} u'(C_t)}$, consistent with the household's marginal rate of substitution.

It will help to distinguish value functions at several different points in time. First, let $V_t(P, A)$ be the value of a firm that begins period t with nominal price P and productivity A, prior to any time t decisions, and prior to time t output. We assume that decision-making takes time, and therefore if the firm decides to make a price adjustment during period t, this new price becomes effective at time $t+1$.[5]

[5] A one-period lag would be unrealistic if the time period were very long. But on the other hand, an infinitesimal time lag in decision-making is logically essential in a continuous-time model (decisions must be "non-anticipative"). When we calibrate the model, we will impose a very short time period, close enough to continuous-time that a one-period lag is not restrictive.

Then, let $O_t(P, A)$ be the value of the option to adjust prices at time t, net of current profits. That is,[6]

$$V_t(P, A) \; = \; \left(P - \frac{W_t}{A}\right) C_t P_t^{\epsilon} P^{-\epsilon} + O_t(P, A) \tag{8}$$

The option value $O_t(P, A)$ summarizes the value of the firm's two possible time-t decisions: whether to adjust its price, and if so, which new price P' to set (for period $t + 1$). The firm may make errors in either of these choices. We discuss these two decisions in turn, beginning with the latter.

2.2.1 Choosing a new price

Our model develops the idea that nominal rigidities may derive primarily from the costs of decision-making. One possible approach would be to assume that upon paying a fixed cost, a firm can make an optimal choice. But this would seem to be an extreme assumption, and a sort of corner solution. We find it more appealing and realistic to assume that firms can devote more or less time and resources to decision-making, in order to choose with more or less precision. In equilibrium in our framework firms will typically prefer to make choices with an interior degree of precision. Therefore their chosen action will not always be the optimal one; instead, firms will sometimes (indeed, usually) make errors.

Consistent with this general description, we adopt the "control cost" approach from game theory (see van Damme, 1991, Chapter 4). A key feature of this approach is that we model the price decision indirectly: we write the decision problem "as if" firms choose a probability distribution over prices, instead of choosing a price directly and deterministically.[7] The decision problem incorporates a cost function that increases with precision: concentrating greater probability on a small range of prices increases costs. Thus, the control cost approach both takes account of the fact that choice is costly, and links this observation to the fact that decisions frequently involve error, while allowing the degree of errors to be controlled by the efforts of the decision-maker.

There are many possible measures of precision. We choose a measure based on relative entropy, also known as Kullback-Leibler divergence, which is a measure of distance between one probability distribution and another. For two distributions $\pi_1(x)$ and $\pi_2(x)$, for some random variable x with support on set $\mathcal{X}$, the Kullback-Leibler divergence $\mathcal{D}(\pi_1||\pi_2)$ of π_1 relative to π_2 is defined by[8]

$$\mathcal{D}(\pi_1||\pi_2) = \int_{\mathcal{X}} \pi_1(x) \ln\left(\frac{\pi_1(x)}{\pi_2(x)}\right) dx. \tag{9}$$

[6] Again, we abuse notation for the sake of brevity: time subscripts on the value functions represent dependence on the aggregate state. Thus, if the aggregate state of the economy is Ω_t, we define $V_t(P, A) \equiv V(P, A, \Omega_t)$ and $O_t(P, A) \equiv O(P, A, \Omega_t)$, Time-subscripted variables in equation (8) represent aggregate quantities: $P_t \equiv P(\Omega_t)$ is the aggregate price level, $W_t \equiv W(\Omega_t)$ is the aggregate wage, and $C_t \equiv C(\Omega_t)$ is aggregate consumption demand.

[7] Luce (1959) and Machina (1985) are early advocates of analyzing decisions in terms of a probability distribution over alternatives; this approach is also adopted by Sims (2003). See Chapter 2 of Anderson *et al.* (1992) for discussion.

[8] While we write (9) with an integral, we can be agnostic at this point about whether $\mathcal{X}$ is a discrete or continuous set. If it is a continuous set, then π_1 and π_2 should be interpreted as density functions. If it is a discrete set, then π_1 and π_2 should be interpreted as vectors of probabilities, and the integral in (9) should be interpreted as a sum.

Following Stahl (1990) and Mattsson and Weibull (2002), we will assume that the decision cost is proportional to the Kullback-Leibler divergence of the chosen distribution, relative to an exogenous benchmark distribution. Thus, if no decision costs are paid, the action x is distributed according to the benchmark distribution. But by putting more effort into the decision process, the decision-maker can bias the distribution of the action towards the most desirable alternatives.

We assume that decision costs are denominated in units of time, since we regard managers' time as the main input to decision-making. Consistent with typical US retail data, we will assume that when the firm sets a new nominal price $\tilde{P}$, this remains constant in nominal terms until a new adjustment occurs. We benchmark the cost of the decision process against an exogenous benchmark distribution $\eta_t(\widetilde{P})$. The time subscript on η allows the benchmark price distribution to change over time, which allows the model to have a nominal trend; we will later detrend the model by restating it in real terms.

Assumption 1. The time cost of choosing a nominal distribution $\pi(\widetilde{P})$ is $\kappa_\pi \mathcal{D}(\pi || \eta_t)$, where $\kappa_\pi > 0$ is a constant, and $\eta_t(\widetilde{P})$ is an exogenous nominal price distribution.

Here κ_π represents the marginal cost of entropy reduction, in units of labor time. The cost function described in Assumption 1 is nonnegative and convex.[9] The upper bound on the cost function is associated with a distribution that places all probability on a single price $\widetilde{P}$ (concretely, costs are maximized when all probability is placed on one price that minimizes the benchmark probability $\eta(\widetilde{P})$). The lower bound on this cost function is zero, associated with choosing the distribution $\pi(\widetilde{P})$ equal to the benchmark distribution $\eta_t(\widetilde{P})$.

Now consider the pricing decision under this cost function. Suppose the firm, at time t, has already decided to update its price, but has not yet chosen which new nominal price to set. Since the new price only becomes effective at time $t+1$, the relevant value of each possible chosen price $\widetilde{P}$ is

$$V_t^e(\widetilde{P}, A) \;\equiv\; E_t\left[\beta \frac{P_t u'(C_{t+1})}{P_{t+1} u'(C_t)} V_{t+1}(\widetilde{P}, A') \middle| A\right], \tag{10}$$

where $E_t\left[\bullet | A\right]$ represents an expectation over the time $t+1$ variables $\Omega' \equiv \Omega_{t+1}$ and $A' \equiv A_{t+1}$ conditional on the time t aggregate state Ω_t and firm j's time t productivity $A_{jt} = A$.

We will write the nominal value of its price decision problem as $\tilde{V}_t(A)$, where $A_{jt} = A$ is the firm's current productivity. This value can be calculated in terms of the expected discounted value V_t^e through the following Bellman equation:

$$\tilde{V}_t(A) = \max_{\pi(\widetilde{P})} \int \pi(\widetilde{P}) V_t^e(\widetilde{P}, A) d\widetilde{P} - \kappa_\pi W_t \int \pi(\widetilde{P}) \ln\left(\frac{\pi(\widetilde{P})}{\eta_t(\widetilde{P})}\right) d\widetilde{P} \quad \text{s.t.} \quad \int \pi(\widetilde{P}) d\widetilde{P} = 1 \tag{11}$$

Thus, the firm chooses a price distribution that maximizes its value, net of computational costs (which

[9] Cover and Thomas (2006), Theorem 2.7.2.

we convert to nominal terms by multiplying by the wage). The first-order condition for $\pi(\widetilde{P})$ is

$$V_t^e(\widetilde{P}, A) - \kappa_\pi W_t \left(1 + \ln\left(\frac{\pi(\widetilde{P})}{\eta_t(\widetilde{P})}\right)\right) - \mu = 0,$$

where μ is the multiplier on the constraint. Some rearrangement yields the following weighted multinomial logit formula:

$$\pi_t(\widetilde{P}|A) \equiv \frac{\eta_t(\widetilde{P}) \exp\left(\frac{V_t^e(\widetilde{P},A)}{\kappa_\pi W_t}\right)}{\int_{\Gamma^P} \eta_t(P') \exp\left(\frac{V_t^e(P',A)}{\kappa_\pi W_t}\right) dP'} \tag{12}$$

The parameter κ_π in the logit function can be interpreted as the degree of noise in the decision process; in the limit as $\kappa_\pi \to 0$, (12) converges to the policy function under full rationality, so that the optimal price is chosen with probability one.

By calculating the logarithm of π_t and plugging it into the objective, we can also obtain an analytical formula for the value function:

$$\tilde{V}_t(A) = \kappa_\pi W_t \ln\left(\int \eta_t(\widetilde{P}) \exp\left(\frac{V_t^e(\widetilde{P}, A)}{\kappa_\pi W_t}\right) d\widetilde{P}\right). \tag{13}$$

This solution gives the value of adjusting the current price, net of decision costs.

Some interpretive comments may be helpful at this point. First, while we write the decision problem "as if" the firm chooses a probability distribution over prices, this should not be taken literally— actually choosing a distribution would be a complex, costly diversion from the true task of choosing a price *per se*. Rather, we describe the decision as a choice of a mixed strategy because this is a way to incorporate errors into the model. And we describe the decision as an optimization problem because this disciplines the errors; it amounts to assuming that the firm devotes sufficient effort to avoiding especially costly errors. Aspects of the model that we do take seriously include (a) making decisions is costly in terms of time and other resources; (b) therefore decision-makers do not always take the action that would otherwise be optimal; (c) *ceteris paribus*, more valuable actions are more probable than less valuable ones; (d) in a retail pricing context, these considerations apply both to the timing of price adjustment, and to the actual price chosen. We will argue, when we come to the quantitative results, that this framework, without any additional type of friction, provides a very successful model of nominal rigidity, in spite of the fact that we restrict the implementation to fairly strong functional form assumptions.

Second, the problem is written conditional on the true expected discounted values $V_t^e(\widetilde{P}, A)$ of the possible nominal prices $\widetilde{P}$, instead of conditioning on a prior, as in a "rational inattention" model. This reflects the fact that we are *not* assuming imperfect information. But this is not equivalent to saying that the firm "knows" the true values $V_t^e(\widetilde{P}, A)$. Instead, our assumption is that the firm has sufficient information to calculate $V_t^e(\widetilde{P}, A)$. Nonetheless, drawing correct conclusions from that information, and acting accordingly, may be costly.[10]

2.2.2 Choosing the timing of price adjustment

We next analyze, in an analogous manner, the decision whether or not to adjust at time t. Note that in the previous subsection we defined costs in terms of the Kullback-Leibler divergence of the price distribution, relative to some benchmark distribution. Now we will make an analogous assumption to model the timing decision. But in this case, at any point in time there are only two options: adjust now, or not. Since the probability of these two actions must sum to one, effectively the relevant benchmark is just a single number, which we can interpret as an exogenous hazard rate.

Now, suppose the time period is sufficiently short so that we can approximately ignore multiple adjustments within a single period. If the firm chooses not to adjust its current price P, then its nominal price next period will be unchanged: $\widetilde{P}' = P$; the expected value of this unchanged price, from the point of view of period t, is $V_t^e(P, A)$. If instead the firm adjusts its price at time t, then its new expected value is $\tilde{V}_t(A)$, as given by (11) and (13). Now suppose it adjusts its price with probability λ. We measure the cost of this adjustment probability in terms of Kullback-Leibler divergence, relative to some arbitrary Poisson process with arrival rate $\bar{\lambda}$:

> **Assumption 2.** The time cost incurred in period t by choosing to adjust in period t with probability $\lambda \in [0, 1]$ is $\kappa_\lambda \mathcal{D}((\lambda, 1-\lambda)||(\bar{\lambda}, 1-\bar{\lambda}))$, where $\kappa_\lambda > 0$ and $\bar{\lambda} \in [0, 1]$ are exogenous parameters.

Here κ_λ is the marginal cost of entropy reduction in the timing decision, which might or might not equal the corresponding parameter κ_π from the pricing decision.

Given this cost function, which we rewrite using the definition (9) of Kullback-Leibler divergence, the optimal adjustment probability is given by solving

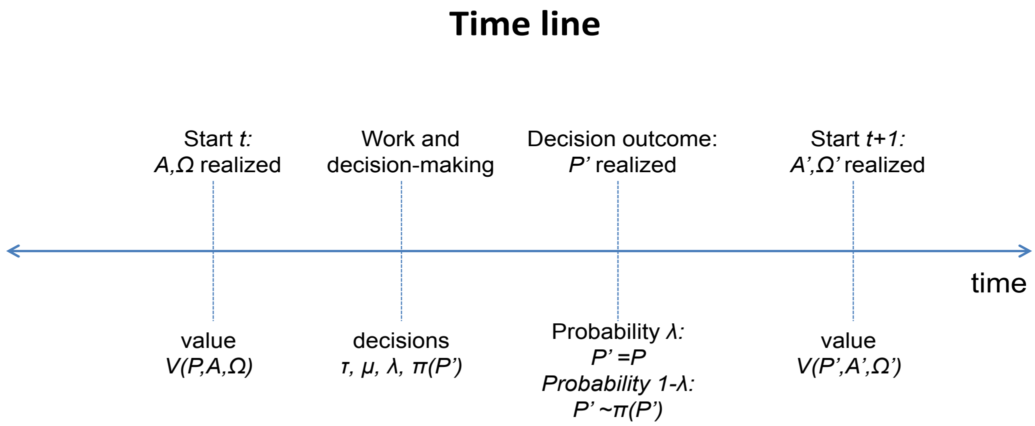
$$O_t(P, A) \;=\; \max_{\lambda}\,(1-\lambda)V_t^e(P, A) + \lambda\tilde{V}_t(A) - \kappa_\lambda W_t\left[\lambda \ln\left(\frac{\lambda}{\bar{\lambda}}\right) + (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right)\right]. \quad (14)$$

Here, the value function O represents the value of the option to adjust, or not, net of decision costs. The first order condition from (14) is

$$V_t^e(P, A) - \tilde{V}_t(A) = \kappa_\lambda W_t\left[\ln\lambda + 1 - \ln\bar{\lambda} - \ln(1-\lambda) - 1 + \ln(1-\bar{\lambda})\right]. \quad (15)$$

[10] Since economists are accustomed to models of perfect rationality, they often equate observing a given information set with knowing all quantities that can be calculated from that information set. But when rationality is less than perfect, we cannot equate these two assumptions. Here, we assume firms can observe all relevant shocks and state variables, but we do not equate this with actually knowing $V_t^e(\widetilde{P}, A)$ or knowing the optimal action, and therefore we do not equate it with implementing the optimal action with probability one.

Figure 1: Timing of firms' decisions.



Rearranging, we can solve (15) to obtain[11]

$$\lambda_t(P,A) = \frac{\bar{\lambda}\exp\left(\frac{\tilde{V}_t(A)}{\kappa_\lambda W_t}\right)}{\bar{\lambda}\exp\left(\frac{\tilde{V}_t(A)}{\kappa_\lambda W_t}\right) + (1-\bar{\lambda})\exp\left(\frac{V_t^e(P,A)}{\kappa_\lambda W_t}\right)} \tag{16}$$

$$= \frac{\bar{\lambda}}{\bar{\lambda} + (1-\bar{\lambda})\exp\left(\frac{-D_t(P,A)}{\kappa_\lambda W_t}\right)} \tag{17}$$

$$\equiv \Lambda\left(\frac{D_t(P,A)}{\kappa_\lambda W_t}\right), \tag{18}$$

where $D_t(P,A)$ is the expected gain from adjustment:

$$D_t(P,A) \equiv \tilde{V}_t(A) - V_t^e(P,A). \tag{19}$$

The weighted binary logit hazard (16) was also derived by Woodford (2008) from a model with a Shannon constraint.[12] The free parameter $\bar{\lambda}$ measures the rate of decision making; concretely, the probability of adjustment in one discrete time period is $\bar{\lambda}$ when the firm is indifferent between adjusting and not adjusting, that is, in the state (P,A) such that $D_t(P,A) = 0$.[13]

[11]Note also that (16) has a well-defined continuous-time limit. If $\bar{\lambda}$ is a continuous-time constant hazard against which we benchmark the costs of a time-varying hazard λ_t, then the continuous-time analogue of (16) is $\lambda_t(P,A) = \bar{\lambda}\exp\left(\frac{V_t^e(P,A)-\tilde{V}_t(A)}{\kappa_\lambda W_t}\right)$.

[12]Woodford's (2009) paper only states a first-order condition like (15); his (2008) manuscript points out that the first-order condition implies a logit hazard of the form (16).

[13]A version of this model nests Calvo price adjustment as a special case. If we set $\kappa_\pi = 0$ and $\kappa_\lambda = \infty$, then the firm always sets the optimal price, conditional on adjustment, and adjustment occurs with a constant probability $\bar{\lambda}$.

2.2.3 Deriving the Bellman equation

A two-step decision

While we have separated the two steps of the decision process in order to describe each one in detail, we can concatenate them into in a single Bellman equation. There is more than one way to do this: first, we write the Bellman equation in a way that shows the two steps separately. If the firm starts period t with nominal price P, then its value $V_t(P,A) \equiv V_t(P,A,\Omega_t)$ at the beginning of t satisfies:

$$\begin{aligned} V_t(P,A) \;=\; \max_{\lambda,\pi(\widetilde{P})} & \left(P - \frac{W_t}{A}\right) C_t P_t^{\epsilon} P^{-\epsilon} + (1-\lambda)V_t^e(P,A) + \lambda \int \pi(\widetilde{P}) V_t^e(\widetilde{P},A) d\widetilde{P} \qquad (20) \\ & -\; \lambda\kappa_\pi W_t \int \pi(\widetilde{P}) \ln\left(\frac{\pi(\widetilde{P})}{\eta_t(\widetilde{P})}\right) d\widetilde{P} \;-\; \kappa_\lambda W_t \left[\lambda \ln\left(\frac{\lambda}{\bar{\lambda}}\right) + (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right)\right] \\ & \text{s.t.} \quad \int \pi(\widetilde{P}) d\widetilde{P} = 1. \end{aligned}$$

This equation decomposes the firm's labor demand into three parts. Time devoted to producing output is $N_t(P,A) \equiv \frac{W_t}{A} C_t \left(\frac{P_t}{P}\right)^{\epsilon}$. Labor time devoted to monitoring whether or not it is a good moment to make a price adjustment is $\mu_t(P,A) \equiv \kappa_\lambda \left[\lambda \ln\left(\frac{\lambda}{\bar{\lambda}}\right) + (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right)\right]$, and (expected) labor time devoted to choosing which new price to set is $\tau_t(P,A) \equiv \lambda\kappa_\pi \int \pi(\widetilde{P}) \ln\left(\frac{\pi(\widetilde{P})}{\eta_t(\widetilde{P})}\right) d\widetilde{P}$.

A one-step decision

Equivalently, the Bellman equation can be written to show the decision as a single step. Consider the following set of alternatives:

$$\Gamma_t^{\dagger} \equiv \{\text{Don't adjust, Adjust to } \tilde{P},\, \tilde{P} \in \Gamma_t^P\}.$$

This set contains the discrete alternative of not adjusting, together with all the possible adjustments in Γ_t^P, which may in principle be a continuum of alternatives. Without loss of generality, suppose the firm chooses the first alternative with probability $1-\lambda$, and draws from the remaining alternatives with density $\pi^{\dagger}(\tilde{P}) = \lambda\pi(\tilde{P})$, where $1 - \lambda + \int \pi^{\dagger}(\tilde{P}) d\tilde{P} = 1$.

Now suppose $\kappa_\pi = \kappa_\lambda = \kappa$.[14] Note that the cost terms from the previous Bellman equation can be rewritten as follows:

$$\lambda\kappa_\pi W_t \int \pi(\widetilde{P}) \ln\left(\frac{\pi(\widetilde{P})}{\eta_t(\widetilde{P})}\right) d\widetilde{P} \;+\; \kappa_\lambda W_t \left[\lambda \ln\left(\frac{\lambda}{\bar{\lambda}}\right) + (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right)\right]$$

[14]The following argument demonstrates that when $\kappa_\pi = \kappa_\lambda$, the two relative entropy cost terms can be combined into a single relative entropy cost term, which then makes it easy to argue that each backwards induction step has a unique solution. When instead $\kappa_\pi \neq \kappa_\lambda$, the single cost term does not take the form of a relative entropy. Nonetheless, it can still be proved that each backwards induction step has a unique solution; see Costain (2016), Proposition 8.

$$= \lambda\kappa W_t \left\{ (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right) \;+\; \lambda\left[\int \pi(\widetilde{P})\ln\left(\frac{\pi(\widetilde{P})}{\eta_t(\widetilde{P})}\right)d\widetilde{P} \;+\; \ln\left(\frac{\lambda}{\bar{\lambda}}\right)\right]\right\}$$

$$= \lambda\kappa W_t \left\{ (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right) \;+\; \lambda\left[\int \pi(\widetilde{P})\ln\left(\frac{\lambda\pi(\widetilde{P})}{\bar{\lambda}\eta_t(\widetilde{P})}\right)d\widetilde{P}\right]\right\}$$

$$= \lambda\kappa W_t \left\{ (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right) \;+\; \left[\int \pi^\dagger(\widetilde{P})\ln\left(\frac{\pi^\dagger(\widetilde{P})}{\eta_t^\dagger(\widetilde{P})}\right)d\widetilde{P}\right]\right\} \tag{21}$$

The expression in (21) is a relative entropy measure over the set of alternatives $\Gamma_t^\dagger$: it is the Kullback-Leibler divergence of the probabilities $(1-\lambda, \pi^\dagger(\tilde{P}))$ relative to the default distribution $(1-\bar{\lambda}, \eta^\dagger(\tilde{P}))$, where $\eta^\dagger(\tilde{P}) \equiv \lambda\eta(\tilde{P})$.

Therefore, the following Bellman equation is equivalent to (20):

$$\begin{aligned} V_t(P,A) \;=\; & \max_{1-\lambda,\pi^\dagger(\widetilde{P})} \left(P - \frac{W_t}{A}\right) C_t P_t^\epsilon P^{-\epsilon} + (1-\lambda)V_t^e(P,A) + \int \pi^\dagger(\widetilde{P})V_t^e(\widetilde{P},A)d\widetilde{P} \\ & - \lambda\kappa W_t \left\{ (1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right) \;+\; \left[\int \pi^\dagger(\widetilde{P})\ln\left(\frac{\pi^\dagger(\widetilde{P})}{\eta_t^\dagger(\widetilde{P})}\right)d\widetilde{P}\right]\right\} \\ & \text{s.t.}\;\; 1-\lambda + \int \pi^\dagger(\widetilde{P})d\widetilde{P} = 1. \end{aligned} \tag{22}$$

Thus, the firm's choice at any time t can be regarded as a single decision across the set of alternatives $\Gamma_t^\dagger$, subject to a relative entropy cost. An advantage of writing the problem this way is that the payoffs from the choice are a linear function of the probabilities $(1-\lambda, \pi^\dagger(\tilde{P}))$, and the costs are strictly convex, because relative entropy is a strictly convex function of probabilities. Therefore the Bellman equation maximizes a stictly concave function (expected payoffs minus costs) over a convex set (the simplex on which the probabilities integrate to one). Therefore we know that a given backwards induction step has a unique solution, given by the unique solution to the first-order conditions.

2.3 Labor market

BEFORE FINAL CALIBRATION WE MIGHT ADD A RETIREMENT/REBIRTH SHOCK.

Each worker i is the monopolistic supplier of a specific type of labor H_{it}, sold at wage W_{it}. Over time, worker i receives shocks Z_{it} that affect the value of her labor time H_{it}. The process Z_{it} is a time-invariant Markov process on a bounded set, $Z_{it} \in \Gamma^Z \subset [\underline{Z}, \overline{Z}]$. We will define $N_{it} = Z_{it}H_{it}$ as her "effective labor". By this definition, we can say that the price of effective labor is $\frac{W_{it}}{Z_{it}}$. These idiosyncratic shocks can be regarded as worker-specific changes in the value of labor, including various forms of human capital accumulation.

The labor demand of each firm j is defined in terms of effective labor. Its total labor input N_{jt} is a

CES aggregate across varieties of effective labor i, with elasticity of substitution ϵ_n:

$$N_{jt} = \left\{ \int_0^1 N_{ijt}^{\frac{\epsilon_n - 1}{\epsilon_n}} di \right\}^{\frac{\epsilon_n}{\epsilon_n - 1}}. \tag{23}$$

It is straightforward to show that under this demand structure, the firm's optimal hiring satisfies

$$H_{ijt} \equiv \frac{N_{ijt}}{Z_{it}} = Z_{it}^{\epsilon_n - 1} \left(\frac{W_{it}}{W_t} \right)^{-\epsilon_n} N_{jt}, \tag{24}$$

when we define the wage index

$$W_t \equiv \left\{ \int_0^1 \left(\frac{W_{it}}{Z_{it}} \right)^{1-\epsilon_n} di \right\}^{\frac{1}{1-\epsilon_n}}. \tag{25}$$

Firm j's nominal wage bill can then be written as

$$\int_0^1 W_{it} H_{ijt} di = W_t N_{jt}. \tag{26}$$

Given the firm's labor demand equation (24), total demand for worker i's labor time is given by

$$H_{it} = Z_{it}^{\epsilon_n - 1} \left(\frac{W_{it}}{W_t} \right)^{-\epsilon_n} N_t, \tag{27}$$

where N_t is total labor demand across all firms j. Note that under the notation of Sec. 2.2, N_{jt} refers to firm j's labor demand *for production*. But firm j also uses labor for decision-making. Our implicit assumption thus far has been that the firm uses the same types of labor for decision-making as it does for production; this is why the same wage W_t appears in the firm's flow of profits from production and in its decision costs. Thus, to account explicitly for all uses of time, let us define $K_{jt}^{\pi} \equiv \kappa_\pi \mathcal{D}(\pi_{jt}||\eta)$, $\tau_{jt} \equiv \lambda_{jt}\kappa_\pi \mathcal{D}(\pi_{jt}||\eta)$, and $\mu_{jt} \equiv \kappa_\lambda \mathcal{D}((\lambda_{jt}, 1-\lambda_{jt})||(\bar{\lambda}, 1-\bar{\lambda}))$, where π_{jt} and λ_{jt} refer to the probability vector and the hazard rate chosen by firm j at time t. Then total labor demand can be defined as

$$N_t = \int_0^1 (N_{jt} + \mu_{jt} + \tau_{jt}) dj. \tag{28}$$

The worker adjusts her nominal wage intermittently to maximize the value of labor income net of labor distutility. This decision is subject to control costs, both on the timing decision, and on the choice of which wage to set. We assume workers act in the interest of the households of which they form part, and that their consumption is fully insured by the household. Therefore they discount future income at the same rate $\beta\frac{P_t u'(C_{t+1})}{P_{t+1}u'(C_t)}$ that applies to the household and firm. Now let $L_t(W, Z)$ be the nominal value of a worker with wage W and productivity Z at the beginning of period t, before supplying labor, and before making any decisions. As in the case of price decisions, we assume that a wage adjustment in period t becomes effective in period $t+1$. Therefore the value of setting the wage to some arbitrary new value $\widetilde{W}$ is

$$L_t^e(\widetilde{W}, Z) \;\equiv\; E_t \left[\beta \frac{P_t u'(C_{t+1})}{P_{t+1} u'(C_t)} L_{t+1}(\widetilde{W}, Z') \middle| Z \right].$$

We make two assumptions about the cost of workers' decision-making that are analogous to those we made in the case of the firm.

Assumption 3. The time cost of choosing a nominal wage distribution $\pi^W(\widetilde{W})$ is $\kappa_w \mathcal{D}(\pi^W || \eta_t^W)$, where $\kappa_w > 0$ is a constant, and $\eta_t^W(\widetilde{W})$ is an exogenous benchmark distribution of nominal wages.

Assumption 4. The time cost incurred in period t by choosing to adjust the wage in period t with probability $\rho_t \in [0,1]$ is $\kappa_\rho \mathcal{D}((\rho_t, 1-\rho_t)||(\bar{\rho}, 1-\bar{\rho}))$, where $\kappa_\rho > 0$ and $\bar{\rho} \in [0,1]$ are exogenous parameters.

Now, let τ be the (expected) amount of time dedicated in period t to setting a new wage; let μ be the amount of time dedicated in period t to monitoring whether it is a good moment to reset the wage; and let $H_t(W, Z) = Z^{\epsilon_n - 1} N_t \left(\frac{W_t}{W}\right)^{\epsilon_n}$ be labor demand conditional on the current wage W and current productivity Z. We can then write the wage setting problem of a worker with convex disutility of labor in a form analogous to the firm's price setting problem (20).

$$L_t(W, Z) \;=\; \max_{\tau, \mu, \rho, \pi^W(\widetilde{W})} W H_t(W, Z) - \frac{P_t}{u'(C_t)} X(H_t(W,Z) + \tau + \mu) + (1-\rho) L_t^e(Z, W) + \rho \int \pi^W(\widetilde{W}) L_t^e(\widetilde{W}, Z) d\widetilde{W} \tag{29}$$

$$\text{s.t.} \quad \int \pi^W(\widetilde{W}) d\widetilde{W} = 1,$$

$$\rho \kappa_w \int \pi^W(\widetilde{W}) \ln \left(\frac{\pi^W(\widetilde{W})}{\eta_t^W(\widetilde{W})} \right) d\widetilde{W} \;=\; \tau, \tag{30}$$

$$\kappa_\rho \left[\rho \ln \left(\frac{\rho}{\bar{\rho}} \right) + (1-\rho) \ln \left(\frac{1-\rho}{1-\bar{\rho}} \right) \right] \;=\; \mu.$$

Here the labor disutility X is multiplied by $P_t/u'(C_t)$ in order to express the whole Bellman equation in nominal units.

In the case of nominal price adjustment, we broke apart the firm's decision into two separate pieces, representing the decision of whether or not to adjust prices, and the decision of what price to set conditional on adjustment. The reason we could break up the decision this way is that we assumed that the firm could hire any quantity of labor at the (aggregate) wage rate W_t: in effect, we defined the firms' labor costs as a linear function of its labor demand. However, from the point of view of a worker, assuming a linear cost function for time use would be highly restrictive. While we will compute an example with a linear labor disutility function $X(h) = \chi h$, we will see that we need a more general, nonlinear specification $X(h) = \chi \frac{h^{1+\zeta}}{1+\zeta}$ in order to match wage adjustment data. But therefore we cannot simply condition on a constant marginal cost of labor: the quantity of labor supplied to the firm affects the marginal cost of time used for each type of decision-making. This is why the two decisions are linked in the wage setting decision (29).

Nonetheless, the policy functions for the wage distribution and the timing of wage adjustment resemble the policy functions implied by the firm's problem. In equation (29), we see that the wage at the beginning of period t is still W, since the firm has not yet made any decision to change its price. In analogy to the firm's problem, we have assumed that the worker is obliged to supply the quantity of labor $H_t(W, Z)$ demanded by its customers, given its current wage. Thus total time use by the worker in period t is $H_t(W, Z) + \tau_t(W, Z) + \mu_t(W, Z)$, and labor disutility is calculated accordingly. Solving the Bellman equation, the density of possible new wages $\widetilde{W}$ is:

$$\pi_t^W(\widetilde{W}|W, Z) \equiv \frac{\eta_t^W(\widetilde{W}) \exp\left(\frac{L_t^e(\widetilde{W}, Z)}{\kappa_w x_t(W, Z)}\right)}{\int \eta_t^W(W') \exp\left(\frac{L_t^e(W', Z)}{\kappa_w x_t(W, Z)}\right) dW'}. \tag{31}$$

Here

$$x_t(W, Z) \equiv \frac{P_t}{u'(C_t)} X'(H_t(W, Z) + \tau_t(W, Z) + \mu_t(W, Z)) \tag{32}$$

denotes the marginal disutility of time in period t, expressed in nominal units so that it is comparable to the value function L^e. Note that x_t is a function of the current wage W. Therefore, although the worker's logit formula (31) is analogous to the firm's logit (12), the worker's probabilities require a higher-dimensional calculation, since they vary across W, Z, and $\widetilde{W}$.

Likewise, if the worker's beginning-of-period wage and productivity are W and Z, her optimal adjustment probability must satisfy:

$$\rho_t(W, Z) = \frac{\bar{\rho} \exp\left(\frac{\tilde{L}_t(W, Z)}{\kappa_\rho x_t(W, Z)}\right)}{\bar{\rho} \exp\left(\frac{\tilde{L}_t(W, Z)}{\kappa_\rho x_t(W, Z)}\right) + (1 - \bar{\rho}) \exp\left(\frac{L_t^e(W, Z)}{\kappa_\rho x_t(W, Z)}\right)} \tag{33}$$

$$= \frac{\bar{\rho}}{\bar{\rho} + (1 - \bar{\rho}) \exp\left(\frac{-D_t^W(W, Z)}{\kappa_\rho x_t(W, Z)}\right)} \tag{34}$$

$$\equiv \rho\left(\frac{D_t^W(W, Z)}{\kappa_\rho x_t(W, Z)}\right), \tag{35}$$

where

$$D_t^W(W, Z) \equiv \tilde{L}_t(W, Z) - L_t^e(W, Z). \tag{36}$$

2.4 Detrending

Before we describe the dynamics of the distributions of firms and workers, it is helpful to remove the nominal trend from the description of firms' and workers' behavior. If we choose the default distributions for nominal prices and wages, $\eta_t^P(\widetilde{P})$ and $\eta_t^W(\widetilde{W})$, so that they can be interpreted as unchanging distributions $\eta^p(\widetilde{p})$ and $\eta^w(\widetilde{w})$ of *real* prices and wages, then the maximization problems of the firms and workers are homogeneous of degree one in nominal prices. Then the Bellman equations can be stated equivalently in real terms, rather than nominal terms as we did above.

Let Ω_t be a nominal aggregate state variable for this economy at time t. This implies that there exist functions P and W that define the nominal price level and the nominal wage level as a function of Ω_t:

$$P_t = P(\Omega_t), \tag{37}$$

$$W_t = W(\Omega_t). \tag{38}$$

We will define real variables by dividing by the aggregate price level, and we will treat all idiosyncratic real variables in logs. In particular, we define the following idiosyncratic quantities:

$$p_{jt} \equiv \ln P_{jt} - \ln P(\Omega_t), \tag{39}$$

$$\widetilde{p}_{jt} \equiv \ln \widetilde{P}_{jt} - \ln P(\Omega_t), \tag{40}$$

$$a_{jt} \equiv \ln A_{jt}, \tag{41}$$

$$w_{it} \equiv \ln W_{it} - \ln P(\Omega_t), \tag{42}$$

$$\widetilde{w}_{it} \equiv \ln \widetilde{W}_{it} - \ln P(\Omega_t), \tag{43}$$

$$z_{it} \equiv \ln Z_{it}, \tag{44}$$

$$\xi_{it} \equiv x(W_{it}, Z_{it}, \Omega_t)/P(\Omega_t). \tag{45}$$

Since we wish to define the default distributions of *real* prices and wages as time invariant, we have two restrictions on the default distributions of nominal variables. Given $\widetilde{P} \equiv P(\Omega_t)e^{\widetilde{p}}$, we must have $\eta_t^P(\widetilde{P}) = \widetilde{P}^{-1}\eta^p(\widetilde{p})$. Likewise, given $W \equiv P(\Omega_t)e^{\widetilde{w}}$, we must have $\eta_t^W(\widetilde{W}) = \widetilde{W}^{-1}\eta^w(\widetilde{w})$.[15]

Now let Ξ_t be the real variable constructed by replacing all nominal state variables that are included in Ω_t by their log real counterparts, and by likewise replacing any distributions of nominal idiosyncratic state variables that are included in Ω_t by the corresponding distributions of log real state variables. It is reasonable to conjecture that Ξ_t is a valid real aggregate state variable for this economy at time t. In particular, this would mean that there exist functions m, w, x, and i that define the real money supply, the real aggregate wage, and the inflation rate in terms of Ξ:

$$m_t \equiv M_t/P(\Omega_t) = m(\Xi_t), \tag{46}$$

$$w_t \equiv W(\Omega_t)/P(\Omega_t) = w(\Xi_t), \tag{47}$$

$$i_t \equiv \ln P(\Omega_t) - \ln P(\Omega_{t-1}) = i(\Xi_t, \Xi_{t-1}). \tag{48}$$

[15] To see this, when we say that there is an unchanging distribution of $\widetilde{p}$, we mean that $cdf_t^P(\widetilde{P}) = cdf^p(\widetilde{p})$, evaluated at the point $\widetilde{P} = P_t e^{\widetilde{p}}$. Using the chain rule, this implies $\frac{\partial cdf_t^P}{\partial P}(\widetilde{P})P_t e^{\widetilde{p}} = \frac{\partial cdf^p}{\partial p}(\widetilde{p})$. Then since $\eta_t^P(\widetilde{P}) \equiv \frac{\partial cdf_t^P}{\partial P}(\widetilde{P})$ and $\eta^p(\widetilde{p}) \equiv \frac{\partial cdf^p}{\partial p}(\widetilde{p})$ we obtain $\eta_t^P(\widetilde{P}) = \widetilde{P}^{-1}\eta^p(\widetilde{p})$.

Note that the variable i_t represents the consumer price inflation rate between periods $t-1$ and t. Likewise, it must be possible to write aggregate consumption and labor as functions of the real state, so that

$$c(\Xi_t) \;=\; C_t \;\equiv\; C(\Omega_t), \tag{49}$$
$$n(\Xi_t) \;=\; N_t \;\equiv\; N(\Omega_t), \tag{50}$$

and firm-specific labor demand can be written as

$$h(w, z, \Xi_t) \;\equiv\; H(P(\Omega_t)e^w, e^z, \Omega_t) \;=\; e^{z(\epsilon_n - 1)} n(\Xi_t) w(\Xi_t)^{\epsilon_n} e^{-\epsilon_n w}. \tag{51}$$

We will now argue that if such a real state variable exists, then the Bellman equations of the firms and workers can be rewritten in real terms. (Then, in the next subsections, we will describe the distributional dynamics, in real terms, and finally identify a candidate for the real state variable Ξ.) Writing the Bellman equation in real terms implies defining the firm's real value as a function of real state variables. In particular, a real representation of Bellman equation (20) requires real value functions v, v^e, and $\tilde{v}$ that satisfy the identities

$$v(p, a, \Xi) \;\equiv\; \frac{V(P(\Omega)e^p, e^a, \Omega)}{P(\Omega)}, \tag{52}$$
$$v^e(p, a, \Xi) \;\equiv\; \frac{V^e(P(\Omega)e^p, e^a, \Omega)}{P(\Omega)} \;=\; \beta E\left\{ \frac{u'(c(\Xi_{t+1}))}{u'(c(\Xi_t))} v(p - i_{t+1}, a', \Xi_{t+1}) \middle| a, \Xi_t \right\},. \tag{53}$$

We see in (53) that in the absence of any nominal price adjustment, a log real price p at time t becomes $p - i_{t+1}$ at time $t + 1$.[16]

Assuming that these real value functions exist, (20) can be rewritten as

$$v(p, a, \Xi_t) \;=\; \max_{\lambda, \pi^p(\tilde{p})} \left(e^p - \frac{w(\Xi_t)}{e^a} \right) c(\Xi_t) e^{-\epsilon p} + (1-\lambda) v^e(p, a, \Xi_t) + \lambda \int \pi^p(\tilde{p}) v^e(\tilde{p}, a, \Xi_t) d\tilde{p} \tag{57}$$
$$-\; \lambda \kappa_\pi w(\Xi_t) \int \pi^p(\tilde{p}) \ln\left(\frac{\pi^p(\tilde{p})}{\eta^p(\tilde{p})} \right) d\tilde{p} \;-\; \kappa_\lambda w(\Xi_t) \left[\lambda \ln\left(\frac{\lambda}{\bar{\lambda}}\right) + (1-\lambda) \ln\left(\frac{1-\lambda}{1-\bar{\lambda}} \right) \right]$$
$$\text{s.t.} \quad \int \pi^p(\tilde{p}) d\tilde{p} = 1.$$

[16]To prove (53), note that

$$P(\Omega)v^e(p, a, \Xi_t) \equiv V^e(P(\Omega)e^p, e^a, \Omega) \;=\; E\left\{ \beta \frac{P(\Omega_t) u'(C(\Omega_{t+1}))}{P(\Omega_{t+1}) u'(c(\Omega_t))} V(P(\Omega_t)e^p, A', \Omega_{t+1}) \middle| A, \Omega_t \right\} \tag{54}$$
$$= E\left\{ \beta \frac{P(\Omega_t) u'(C(\Omega_{t+1}))}{P(\Omega_{t+1}) u'(C(\Omega_t))} V(P(\Omega_{t+1}) e^{p - i_{t+1}}, A', \Omega_{t+1}) \middle| A, \Omega_t \right\} \tag{55}$$
$$= P(\Omega_t) E\left\{ \beta \frac{u'(c(\Xi_{t+1}))}{u'(c(\Xi_t))} v(e^{p - i_{t+1}}, a', \Xi_{t+1}) \middle| a, \Xi_t \right\}. \tag{56}$$

(((COULD COMMENT OUT THE FOLLOWING INTERMEDIATE STAGES OF BELLMAN)))

Note that (8) can be decomposed as

$$v(p,a,\Xi_t) \;=\; \left(e^p - w(\Xi_t)e^{-a}\right)c(\Xi_t)e^{-\epsilon p} \;+\; o(p,a,\Xi_t), \tag{58}$$

where

$$o(p,a,\Xi) \;\equiv\; \frac{O(P(\Omega)e^p,e^a,\Omega)}{P(\Omega)} \;=\; \max_{\lambda}(1-\lambda)v^e(p,a,\Xi)+\lambda\tilde{v}(a,\Xi)-\kappa_\lambda w(\Xi)\left\{\lambda\ln\left(\frac{\lambda}{\bar{\lambda}}\right)+(1-\lambda)\ln\left(\frac{1-\lambda}{1-\bar{\lambda}}\right)\right\}, \tag{59}$$

and

$$\tilde{v}(a,\Xi) \;\equiv\; \frac{\tilde{V}(e^a,\Omega)}{P(\Omega)} \;=\; \max_{\pi(\tilde{p})}\int_{\gamma^p}\pi(\tilde{p})v^e(\tilde{p},a,\Xi)d\tilde{p}-\kappa_\pi w(\Xi)\int_{\gamma^p}\pi(\tilde{p})\ln\left(\frac{\pi(\tilde{p})}{\eta(\tilde{p})}d\tilde{p}\right) \quad \text{s.t.} \int_{\gamma^p}\pi(\tilde{p})d\tilde{p}=1, \tag{60}$$

The solutions of these equations are

$$\tilde{v}(a,\Xi) \;=\; \kappa_\pi w(\Xi)\ln\left(\int_{\gamma^p}\eta(\tilde{p})\exp\left(\frac{v^e(\tilde{p},a,\Xi)}{\kappa_\pi w(\Xi)}\right)d\tilde{p}\right), \tag{61}$$

$$o(p,a,\Xi) \;=\; v^e(p,a,\Xi)+\kappa_\lambda w(\Xi)\ln\left(1-\bar{\lambda}+\bar{\lambda}\exp\left(\frac{d(p,a,\Xi)}{\kappa_\lambda w(\Xi)}\right)\right), \tag{62}$$

where

$$d(p,a,\Xi) \;\equiv\; \tilde{v}(a,\Xi)-v^e(p,a,\Xi). \tag{63}$$

Obviously, the worker's Bellman equation (29) can be detrended in analogy with that of the firm. To do so, we postulate real value functions l and l^e that satisfy the identities

$$l(w,z,\Xi) \;\equiv\; \frac{L(P(\Omega)e^w,e^z,\Omega)}{P(\Omega)}, \tag{64}$$

$$l^e(w,z,\Xi) \;\equiv\; \frac{L^e(P(\Omega)e^w,e^z,\Omega)}{P(\Omega)}, \;=\; \beta E\left\{\frac{u'(c(\Xi_{t+1}))}{u'(c(\Xi_t))}l(w-i_{t+1},z',\Xi_{t+1})|z,\Xi_t\right\}. \tag{65}$$

The worker's Bellman equation can now be rewritten in real terms as follows:

$$l(w,z,\Xi_t) \;=\max_{\tau,\mu,\rho,\pi^w(\tilde{w})} e^w h_t(w,z,\Xi_t)-\frac{X(h(w,z,\Xi_t)+\tau+\mu)}{u'(c(\Xi_t))}+(1-\rho)l_t^e(w,z,\Xi_t)+\rho\int\pi^w(\tilde{w})l^e(\tilde{w},z,\Xi_t)d\tilde{w} \tag{66}$$

$$\text{s.t.} \quad \int\pi^w(\tilde{w})d\tilde{w}=1,$$

$$\rho\kappa_w\int\pi^w(\tilde{w})\ln\left(\frac{\pi^w(\tilde{w})}{\eta^w(\tilde{w})}\right)d\tilde{w} \;=\; \tau,$$

$$\kappa_\rho\left[\rho\ln\left(\frac{\rho}{\bar{\rho}}\right)+(1-\rho)\ln\left(\frac{1-\rho}{1-\bar{\rho}}\right)\right] \;=\; \mu.$$

Analyzing (66), it is straightforward to show that the chosen distribution of wages takes the form[17]

$$\pi_t^w(\widetilde{w}|w,z) \;\equiv\; \frac{\eta^w(\widetilde{w})\exp\left(\frac{l_t^e(\widetilde{w},w)}{\kappa_w \xi_t(w,z)}\right)}{\int \eta^w(w')\exp\left(\frac{l_t^e(w',z)}{\kappa_w \xi_t(w,z)}\right)dw'}\,, \tag{67}$$

where

$$\xi_t(w,z) \;\equiv\; \frac{X'(h_t(w,z)+\tau_t(w,z)+\mu_t(w,z))}{u'(C_t)} \tag{68}$$

is the worker's marginal disutility of time spent working, expressed in units of consumption goods. Similarly, using the first-order condition for ρ, it can be shown that the adjustment hazard takes the following form:

$$\rho_t(w,z) \;=\; \frac{\bar{\rho}\exp\left(\frac{\tilde{l}_t(w,z)}{\kappa_\rho \xi_t(w,z)}\right)}{\bar{\rho}\exp\left(\frac{\tilde{l}_t(w,z)}{\kappa_\rho \xi_t(w,z)}\right)+(1-\bar{\rho})\exp\left(\frac{l_t^e(w,z)}{\kappa_\rho \xi_t(w,z)}\right)}. \tag{69}$$

Thus, the decision noise in both the timing choice and the wage-setting choice is proportional to the worker's marginal disutility of labor.

Indeed, to characterize the worker's decision in a given state (w,z,Ξ), it suffices to find the unique value of $\xi_t(w,z)$ that solves (68).[18] The time devoted to decision-making is

$$\begin{aligned}\mu_t(w,z)+\tau_t(w,z) \;&= \kappa_\rho \rho_t(w,z)\log\left(\frac{\rho_t(w,z)}{\bar{\rho}}\right)+(1-\rho_t(w,z))\log\left(\frac{1-\rho_t(w,z)}{1-\bar{\rho}}\right)\\ &\quad + \kappa_w \rho_t(w,z)\int \pi^w(\widetilde{w}|w,z)\ln\left(\frac{\pi^w(\widetilde{w}|w,z)}{\eta^w(\tilde{w})}\right)d\tilde{w}.\end{aligned} \tag{70}$$

This can be calculated using formulas (67) and (69), and it is strictly decreasing as a function of ξ.[19] Since marginal disutility increases strictly with total time use (and since $h_t(w,z)$ is predetermined, so it does not depend on ξ), the right-hand side of (68) can be regarded as a strictly decreasing function of ξ. Therefore (68) can be solved by bisection to give a unique solution $\xi_t(w,z) \geq 0$ in any given state (w,z,Ξ_t).

[17] As we showed earlier for the worker's problem, it is possible to rewrite this problem in terms of a single entropy cost term (a convex function) and a linear objective function. Since labor disutility is also convex, a unique well-defined solution exists for the maximization problem involved in a single backwards induction step.

[18] This discussion refers to the calculations required to solve a backwards induction step of the Bellman equation (66). Therefore we regard the future value $\tilde{l}_t^e(w,z)$ as a known function.

[19] Again, the easiest way to prove this is to recognize that the timing choice and the wage-setting choice can be written as a single decision problem. Then $\mu+\tau$ simply represents the relative entropy of the decision, multiplied by the scalar κ. Relative entropy is a strictly decreasing function of the decision noise ξ (for a proof, see Lemma 1e, Costain 2016). Total decision time is also a decreasing function of ξ in the general case when $\kappa_\rho \neq \kappa_w$; see Prop. 8, Costain 2016.

Again, the decision can be decomposed into various parts:

$$l(w,z,\Xi_t) \;= \max_{\tau,\mu} e^w h_t(w,z,\Xi_t) - \frac{X(h(w,z,\Xi_t)+\tau+\mu)}{u'(c(\Xi_t))} + o^w(\mu,\tau,w,z,\Xi_t), \tag{71}$$

$$o^w(\mu,\tau,w,z,\Xi_t) \;\equiv\; \max_{\rho}\,(1-\rho)l^e(w,z,\Xi_t) + \rho\tilde{l}(\rho,\tau,w,z,\Xi_t) \quad \text{s.t. } \kappa_\rho\left\{\rho\ln\left(\frac{\rho}{\bar{\rho}}\right) + (1-\rho)\ln\left(\frac{1-\rho}{1-\bar{\rho}}\right)\right\} = \mu, \tag{72}$$

$$\tilde{l}(\rho,\tau,w,z,\Xi_t) \;\equiv\; \max_{\pi^w(\tilde{w})} \int_{\gamma^w} \pi^w(\tilde{w})l^e(\tilde{w},z,\Xi_t)d\tilde{w} \quad \text{s.t. } \rho\kappa_w \int_{\gamma^w} \pi^w(\tilde{w})\ln\left(\frac{\pi^w(\tilde{w})}{\eta^w(\tilde{w})}d\tilde{w}\right) = \tau\;, \;\; \int_{\gamma^w}\pi^w(\tilde{w})d\tilde{w} = 1. \tag{73}$$

The solutions of (73)-(72) are

$$\tilde{l}(w,z,\Xi) \;=\; \kappa_w\xi(w,z,\Xi)\ln\left(\int_{\gamma^w}\eta^w(\tilde{w})\exp\left(\frac{l^e(\tilde{w},z,\Xi)}{\kappa_w\xi(w,z,\Xi)}\right)d\tilde{w}\right), \tag{74}$$

$$o^w(w,z,\Xi) \;=\; l^e(w,z,\Xi) + \kappa_\rho\xi(w,z,\Xi)\ln\left(1-\bar{\rho}+\bar{\rho}\exp\left(\frac{d^w(w,z,\Xi)}{\kappa_\rho\xi(w,z,\Xi)}\right)\right), \tag{75}$$

where

$$d^w(w,z,\Xi) \;\equiv\; \tilde{l}(w,z,\Xi) - l^e(w,z,\Xi). \tag{76}$$

2.5 Distributional dynamics

The distribution of firms' prices and productivities, and likewise that of workers' wages and productivities, evolves over time as firms and workers respond to idiosyncratic and aggregate shocks. We first state the equations governing the dynamics of the distribution across firms.

We continue to use the notation P_{jt} to refer to the nominal price at which firm j produces in period t, prior to adjustment. This may of course differ from its price $\widetilde{P}_{jt}$ at the end of t, when price adjustments are realized. Therefore we will distinguish the beginning-of-period distribution of prices and log productivities, $\Phi_t(P_{jt},a_{jt})$, from the distribution of prices and log productivities at the end of t, $\widetilde{\Phi}_t(\widetilde{P}_{jt},a_{jt})$. But instead of tracking nominal prices P_{jt}, it is simpler to focus on log real prices p_{jt}. Therefore, in analogy to the nominal distributions, we define $\Psi_t(p_{jt},a_{jt})$ as the real distribution at the beginning of t, when production takes place, and $\widetilde{\Psi}_t(\widetilde{p}_{jt},a_{jt})$ as the real distribution at the end of t. Finally, we also use lower-case letters to represent the joint densities associated with these distributions, which we write as $\phi_t(P_{jt},a_{jt})$, $\widetilde{\phi}_t(\widetilde{P}_{jt},a_{jt})$, $\psi_t(p_{jt},a_{jt})$, and $\widetilde{\psi}_t(\widetilde{p}_{jt},a_{jt})$, respectively.[20]

Two stochastic processes drive the dynamics of the distribution. First, there is the Markov process

[20] Our notation in this section assumes that all densities are well-defined on a continuous support, but we do not actually impose this assumption on the model. With slightly more sophisticated notation we could allow explicitly for distributions with mass points, or with discrete support.

for firm-specific log productivity, which we can write in terms of the following *c.d.f.*:

$$S(a'|a) \;=\; prob(a_{j,t} \leq a'|a_{j,t-1} = a), \tag{77}$$

or in terms of the corresponding density function:

$$s(a'|a) \;=\; \frac{\partial}{\partial a'} S(a'|a). \tag{78}$$

Thus, suppose that the density of nominal prices and log productivities at the end of period $t-1$ is $\tilde{\phi}_{t-1}(\tilde{P}, a)$. This density is then affected by productivity shocks; the density at the beginning of t will therefore be

$$\phi_t(\tilde{P}, a') \;=\; \int s(a'|a)\widetilde{\phi}_{t-1}(\tilde{P}, a)da. \tag{79}$$

But this equation conditions on a given nominal price $\tilde{P}$. Holding fixed a firm's nominal price, its real log price is changed by inflation, from $\widetilde{p}_{i,t-1}$ to $p_{i,t} \equiv \widetilde{p}_{i,t-1} - i_t$, where $i_t = \log(P_t/P_{t-1})$. Therefore the density of real log prices and log productivities at the beginning of t is given by

$$\psi_t\left(\widetilde{p} - \log\frac{P_t}{P_{t-1}}, a'\right) \;=\; \int s(a'|a)\widetilde{\psi}_{t-1}(\widetilde{p}, a)da, \tag{80}$$

and hence the cumulative distribution at the beginning of t, in real terms, is

$$\Psi_t(p, a') \;=\; \int^p \int^{a'} \left(\int s(b|a)\widetilde{\psi}_{t-1}\left(q + i_t, a\right) da\right) db\, dq. \tag{81}$$

The second stochastic process that determines the dynamics is the process of real price updates, which we have defined in terms of a conditional density of logit form in (12). A firm with real log price p and log productivity a at the beginning of period t adjusts its price with probability $\lambda\left(\frac{d_t(p,a)}{\kappa_\lambda w_t}\right)$, where

$$d_t(p, a) \;\equiv\; \widetilde{v}_t(a) - v^e_t(p, a). \tag{82}$$

Upon adjustment, its new real log price is distributed according to $\pi_t(\tilde{p}|a)$. Therefore, if the density of firms at the beginning of t is $\psi_t(p, a)$, the density at the end of t is given by

$$\widetilde{\psi}_t(\widetilde{p}, a) \;=\; \left(1 - \lambda\left(\frac{d_t(\widetilde{p}, a)}{\kappa_\lambda w_t}\right)\right)\psi_t(\widetilde{p}, a) + \int \lambda\left(\frac{d_t(p, a)}{\kappa_\lambda w_t}\right)\pi_t(\widetilde{p}|a)\psi_t(p, a)dp. \tag{83}$$

The cumulative distribution at the end of t is simply given by integrating up this density:

$$\widetilde{\Psi}_t(p, a) \;=\; \int^{\widetilde{p}} \int^{a} \widetilde{\psi}_t(q, b)db\, dq. \tag{84}$$

The dynamics of wages and worker productivities is analogous; it suffices to go directly to the real log dynamics, without developing notation for the nominal dynamics. Let $\Psi_t^w(w_{it}, z_{it})$ be the distribution of real log prices and log worker productivities at the beginning of the period, when production takes place, and let $\widetilde{\Psi}_t^w(\widetilde{w}_{it}, z_{it})$ as the corresponding distribution at the end of the period. We write the densities associated with these distributions as $\psi_t^w(w_{it}, z_{it})$ and $\widetilde{\psi}_t^w(\widetilde{w}_{it}, z_{it})$, respectively.

We assume worker productivity is driven by the Markov process S^z:

$$S^w(z'|z) \;=\; prob(z_{i,t+1} \leq z'|z_{i,t} = z), \tag{85}$$

with the following density function:

$$s^z(z'|z) \;=\; \frac{\partial}{\partial z'} S(z'|z). \tag{86}$$

Meanwhile, holding fixed a worker's nominal wage, her real log wage is changed by inflation, from $\widetilde{w}_{i,t-1}$ at the end of $t-1$, to $w_{i,t} \equiv \widetilde{w}_{i,t-1} - i_t$. Therefore the density of real log wages and log worker productivities at the beginning of t is given by

$$\psi_t^w\left(\widetilde{w} - i_t, z'\right) \;=\; \int s^z(z'|z)\tilde{\psi}_{t-1}^w(\widetilde{w}, z)dz. \tag{87}$$

Hence the corresponding cumulative distribution at the beginning of t is

$$\Psi_t^w(w, z') \;=\; \int^w \int^{z'} \left(\int s^z(b|z)\tilde{\psi}_{t-1}^w\left(q + i_t, z\right)dz \right) db\, dq. \tag{88}$$

Next, a worker with real log wage w and log productivity z at the beginning of period t adjusts her wage with probability $\rho\left(\frac{d_t^w(w,z)}{\kappa_\rho \xi_t(w,z)}\right)$, where

$$d_t^w(w, z) \;\equiv\; \widetilde{l}_t(w, z) - l_t^e(w, z). \tag{89}$$

Upon adjustment, her new real log wage is distributed according to $\pi_t^w(\tilde{w}|w, z)$. Therefore, if the density of workers at the beginning of t is $\psi_t^w(w, z)$, the density at the end of t is given by

$$\widetilde{\psi}_t^w(\widetilde{w}, z) \;=\; \left(1 - \rho\left(\frac{d_t^w(\widetilde{w}, z)}{\kappa_\rho \xi_t(\widetilde{w}, z)}\right)\right)\psi_t^w(\widetilde{w}, z) + \int \rho\left(\frac{d_t^w(w, z)}{\kappa_\rho \xi_t(w, z)}\right)\pi_t^w(\widetilde{w}|w, z)\psi_t^w(w, z)dw. \tag{90}$$

The cumulative distribution at the end of t is simply given by integrating up this density:

$$\widetilde{\Psi}_t^w(\widetilde{w}, z) \;=\; \int^{\tilde{w}} \int^z \psi_t(q, b)db\, dq. \tag{91}$$

2.6 Monetary policy and aggregate consistency

The nominal money supply is affected by an AR(1) shock process g,[21]

$$g_t = \phi_g g_{t-1} + \epsilon^g_t, \tag{92}$$

where $0 \leq \phi_g < 1$ and $\epsilon^g_t \sim i.i.d. N(0, \sigma^2_g)$. Here g_t represents the time t rate of money growth:

$$M_t / M_{t-1} \equiv \mu_t = \mu^* \exp(g_t). \tag{93}$$

Seigniorage revenues are paid to the household as a lump sum transfer T^M_t, and the government budget is balanced each period, so that $M_t = M_{t-1} + T^M_t$.

Bond market clearing is simply $B_t = 0$. When supply equals demand for each good j, total labor supply and demand satisfy

$$N_t - K^\lambda_t - K^\pi_t = \int_0^1 \frac{C_{jt}}{A_{jt}} dj = C_t \int \int \psi_t(p, a) \exp(-\epsilon p - a) da\, dp \equiv \Delta_t C_t, \tag{94}$$

where K^λ_t is total time devoted to deciding whether to adjust prices, and K^π_t is total time devoted to choosing which price to set by firms that adjust. Equation (94) also defines a measure of price dispersion, $\Delta_t \equiv P^\epsilon_t \int_0^1 P^{-\epsilon}_{jt} A^{-1}_{jt} dj$, weighted to allow for heterogeneous productivity. As in Yun (2005), an increase in Δ_t decreases the goods produced per unit of labor, effectively acting like a negative aggregate productivity shock.

In nominal terms, the price level and wage level are given as follows

$$\int \int P^{1-\epsilon} \phi_t(P, A) dA\, dP \;=\; P(\Omega_t)^{1-\epsilon}. \tag{95}$$

$$\int \int \left(\frac{W}{Z}\right)^{1-\epsilon_N} \phi^W_t(W, Z) dZ\, dW \;=\; W(\Omega_t)^{1-\epsilon_N}. \tag{96}$$

Given (95), the real price level is one by definition:

$$\int \int \exp((1-\epsilon)p) \psi_t(p, a) da\, dp \;=\; 1. \tag{97}$$

The real wage level satisfies

$$\int \int \exp((1-\epsilon_N)(w - z)) \psi^W_t(w, z) dz\, dw \;=\; w(\Xi_t)^{1-\epsilon_N}. \tag{98}$$

[21] In related work (Costain and Nakov 2011B) we have studied state-dependent pricing when the monetary authority follows a Taylor rule. Our conclusions about the degree of state-dependence, microeconomic stylized facts, and the real effects of monetary policy were not greatly affected by the type of monetary policy rule considered. Therefore we focus here on the simple, transparent case of a money growth rule.

3 Results

3.1 Parameters

Consumption utility and labor disutility are parameterized as $u(C) = \frac{1}{1-\gamma}(C^{1-\gamma} - 1)$ and $X(h) = \frac{\chi}{1+\eta} h^{1+\eta}$.

The productivity processes and workers are assumed to follow discretized approximations of the following AR(1) processes:

$$a_{jt} = \rho_a a_{jt-1} + \epsilon_t^a, \tag{99}$$

$$z_{it} = \rho_z z_{it-1} + \epsilon_t^z, \tag{100}$$

where ϵ_t^a and ϵ_t^z are *i.i.d.* normal shocks with mean zero. Thus the variances of a_{jt} and z_{it} are $\sigma_a^2 = \frac{\sigma_{\epsilon a}^2}{1-\rho_a^2}$ and $\sigma_z^2 = \frac{\sigma_{\epsilon z}^2}{1-\rho_z^2}$, where $\sigma_{\epsilon a}^2$ and $\sigma_{\epsilon z}^2$ are the variances of the innovations ϵ_t^a and ϵ_t^z, respectively.

3.2 Frictionless steady state

As a basis for comparison, and as an initial guess for calculating the general equilibrium of our costly adjustment model, we will also calculate a version of our model without any frictions.

When prices are frictionless, firms choose P_{jt} to optimize static profits $P_{jt}^{1-\epsilon} C_t P_t^{\epsilon} - \frac{W_t P_{jt}^{-\epsilon}}{A_{jt}} C_t P_t^{\epsilon}$ each period, implying the first-order condition $P_{jt} = \left(\frac{\epsilon}{\epsilon-1}\right) \frac{W_t}{A_{jt}}$. In real terms, this becomes $p_{jt} = \log(P_{jt}/P_t) = \log\left(\frac{\epsilon(W_t/P_t)}{\epsilon-1}\right) - a_{jt}$. Hence in an aggregate steady state with no frictions (indicated by subscript nf),

$$p_{jt} \sim N\left(\log\left(\frac{\epsilon w_{nf}}{\epsilon - 1}\right), \sigma_a^2\right).$$

The aggregate price identity can be written as $E\exp((1-\epsilon)p) = 1$. Hence in an aggregate steady state we must have

$$\begin{aligned} 1 &= \exp\left((1-\epsilon)\log\left(\frac{\epsilon w_{nf}}{\epsilon-1}\right)\right)\exp\left(\frac{(1-\epsilon)^2\sigma_a^2}{2}\right) \\ &= \left(\frac{\epsilon w_{nf}}{\epsilon-1}\right)^{1-\epsilon}\exp\left(\frac{(1-\epsilon)^2\sigma_a^2}{2}\right). \end{aligned}$$

But w_{nf} is the only endogenous variable in this equation, so we can invert it to calculate w_{nf} explicitly:

$$w_{nf} = \frac{\epsilon - 1}{\epsilon} \exp\left(\frac{(\epsilon-1)\sigma_a^2}{2}\right). \tag{101}$$

In the limit where decisions are costless, labor market clearing is simply $N_t = \Delta_t C_t$, where $\Delta_t =$

$E\exp(-\epsilon p_{jt} - a_{jt})$. We then have

$$-\epsilon p_{jt} - a_{jt} \;=\; -\epsilon \log\left(\frac{\epsilon w_{nf}}{\epsilon - 1}\right) - (1-\epsilon)a_{jt} \;=\; -\epsilon\left(\frac{(\epsilon-1)\sigma_a^2}{2}\right) - (1-\epsilon)a_{jt},$$

where we have used $\log\left(\frac{\epsilon w_{nf}}{\epsilon-1}\right) = \frac{(\epsilon-1)\sigma_a^2}{2}$. Hence in the frictionless steady state,

$$\Delta_{nf} \;=\; E\exp(-\epsilon p_{jt} - a_{jt}) \;=\; \exp\left(-\epsilon\frac{(\epsilon-1)\sigma_a^2}{2}\right)\exp\left(\frac{(\epsilon-1)^2\sigma_a^2}{2}\right) \;=\; \exp\left(\frac{(1-\epsilon)\sigma_a^2}{2}\right) \tag{102}$$

This is an explicit formula for the dispersion variable Δ_{nf} in the steady state of the frictionless model. Hence we know labor demand as a linear multiple of consumption $N_{nf} = \Delta_{nf}C_{nf}$.

Likewise, when wages are frictionless, workers choose W_{it} to optimize their static payoffs $W_{it}H_t(W_{it}, Z_{it}) - \frac{P_t}{u'(C_t)}X(H_t(W_{it}, Z_{it}))$, where $H_t(W_{it}, Z_{it}) \equiv N_t W_t^{\epsilon_n} Z^{\epsilon_n - 1} W_{it}^{-\epsilon_n}$. (Since decisions are costless, only time spent working enters as an argument of the disutility function X.) The first-order condition is

$$\frac{W_{it}}{P_t} \;=\; \frac{\epsilon_n}{\epsilon_n - 1}C_t^\gamma \chi H_t(W_{it}, Z_{it})^\eta \;=\; \frac{\epsilon_n \chi}{\epsilon_n - 1}C_t^\gamma\left[N_t w_t^{\epsilon_n} Z_{it}^{\epsilon_n - 1}(W_{it}/P_t)^{-\epsilon_n}\right]^\eta$$

In steady state, this yields

$$w_{it} \;\equiv\; \log\left(\frac{W_{it}}{P_{nf}}\right) \;=\; \frac{1}{1+\epsilon_n\eta}\log\left(\frac{\epsilon_n\chi}{\epsilon_n - 1}C_{nf}^\gamma (N_{nf} w_{nf}^{\epsilon_n})^\eta\right) + \frac{\eta(\epsilon_n - 1)}{1+\epsilon_n\eta}z_{it}.$$

Hence the steady-state distribution is

$$w_{it} \;\sim\; N\left(\frac{1}{1+\epsilon_n\eta}\log\left(\frac{\epsilon_n\chi}{\epsilon_n - 1}C_{nf}^\gamma (N_{nf} w_{nf}^{\epsilon_n})^\eta\right), \left(\frac{\eta(\epsilon_n - 1)}{1+\epsilon_n\eta}\right)^2 \sigma_z^2\right).$$

The definition of the wage index can be written as

$$w_t^{1-\epsilon_n} \;=\; E\exp\left((1-\epsilon_n)(w_{it} - z_{it})\right),$$

where, in steady state,

$$(1-\epsilon_n)(w_{it} - z_{it}) \;=\; \frac{1-\epsilon_n}{1+\epsilon_n\eta}\log\left(\frac{\epsilon_n\chi}{\epsilon_n - 1}C_{nf}^\gamma (N_{nf} w_{nf}^{\epsilon_n})^\eta\right) + \frac{(\epsilon_n - 1)(1+\eta)}{1+\epsilon_n\eta}z_{it}.$$

Hence in steady state the wage index must satisfy

$$w_{nf}^{1-\epsilon_n} \;=\; \left(\frac{\epsilon_n\chi}{\epsilon_n - 1}C_{nf}^\gamma (N_{nf} w_{nf}^{\epsilon_n})^\eta\right)^{\frac{1-\epsilon_n}{1+\epsilon_n\eta}} \exp\left(\left(\frac{(1-\epsilon_n)(1+\eta)}{1+\epsilon_n\eta}\right)^2 \sigma_z^2\right).$$

Raising each side to the power $\frac{1}{1-\epsilon_n}$, and rearranging, we obtain

$$w_{nf}^{\frac{1}{1+\epsilon_n\eta}} = \left(\frac{\epsilon_n\chi}{\epsilon_n-1}C_{nf}^{\gamma}N_{nf}^{\eta}\right)^{\frac{1}{1+\epsilon_n\eta}} \exp\left((1-\epsilon_n)\left(\frac{1+\eta}{1+\epsilon_n\eta}\right)^2 \sigma_z^2\right),$$

or equivalently

$$w_{nf} = \left(\frac{\epsilon_n\chi}{\epsilon_n-1}C_{nf}^{\gamma+\eta}\Delta_{nf}^{\eta}\right)\exp\left((1+\eta)^2\left(\frac{1-\epsilon_n}{1+\epsilon_n\eta}\right)\sigma_z^2\right),$$

Since w_{nf} and Δ_{nf} are already known, we can solve this equation to find steady-state frictionless consumption:

$$C_{nf}^{\gamma+\eta} = \left(\frac{\epsilon_n-1}{\epsilon_n\chi}\Delta_{nf}^{-\eta}\right)\exp\left((1+\eta)^2\left(\frac{\epsilon_n-1}{1+\epsilon_n\eta}\right)\sigma_z^2\right)w_{nf}. \tag{103}$$

3.3 Steady state results

We compare:

- V1: Benchmark. Sticky prices and sticky wages.
- V2: Semi-flexible prices and sticky wages.
- V3: Flexible prices and sticky wages.
- V4: Sticky prices and semi-flexible wages.
- V5: Sticky prices and flexible wages.
- V6: Flexible prices and Flexible wages.

	V 1	V. 2	V. 3 *	V. 4	V. 5	V. 6
$\kappa_\pi = \kappa_\lambda$	1.7%	0.17%	0. 017%	1.7%	1.7%	0.017%
$\kappa_w = \kappa_\rho$	1.7%	1.7%	1.7%	0.17%	0.017%	0.017%

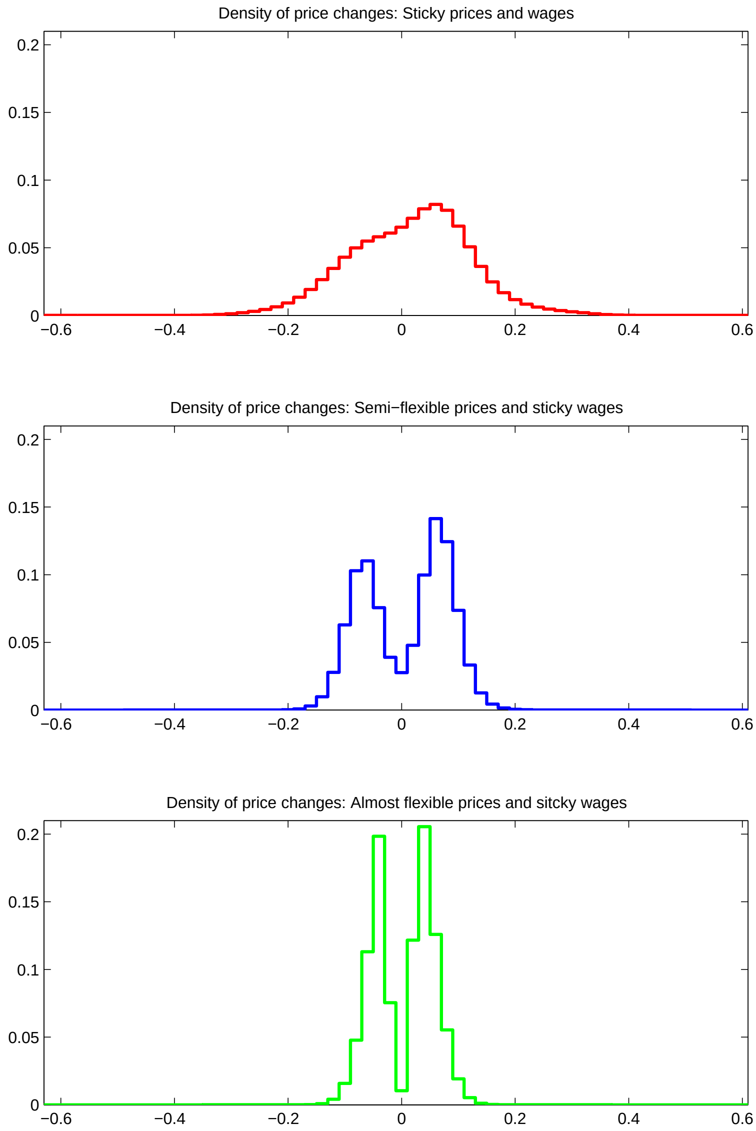
3.4 Results: price and wage adjustments.

Evaluating the LPW model with different values of κ_π, κ_λ, κ_w and κ_ρ

Decreasing κ_π and κ_λ								
	V1 (κ_0)		V2 ($\kappa_0/10$)		V3 ($\kappa_0/100$)		V6 ($\kappa_0/100$)	
Consumption	0.349572		0.350833		0.351400		0.352165	
Labor	0.352963		0.349296		0.348088		0.348845	
Wage	0.857613		0.863830		0.866628		0.866630	
	Prices	Wages	Prices	Wages	Prices	Wages	Prices	Wages
Frequency of adj.	10.1	6.02	22.5	6.03	54.4	6.04	54.4	6.95
Mean change	1.68	2.83	0.76	2.82	0.31	2.82	0.31	2.45
Std of prices/wages	5.26	3.07	5.63	3.08	6.06	3.08	6.06	0.96
Mean abs(Change)	8.57	6.14	6.80	6.16	4.76	6.16	4.76	2.29
Std of changes	10.6	8.53	7.50	8.53	5.30	8.52	5.26	1.79
Kurtosis of changes	3.20	10.0	1.82	9.77	1.94	9.66	1.94	3.33
Price/wage increses	58.5	72.1	55.5	72.0	54.0	71.9	54.0	92.0
Price/wage increses ($\leq$5%)	28.8	48.1	27.0	47.9	60.7	47.9	60.8	99.9
Price/wage increses ($\leq$2.5%)	14.2	24.7	8.90	24.6	19.9	24.6	20.0	61.6
Pricing cost as % Rev.	0.51	0.33	0.19	0.33	0.07	0.33	0.07	0.01
Timing cost as % Rev.	0.37	0.35	0.11	0.35	0.03	0.35	0.03	0.01
Loss. relative to Flex	1.78	0.57	0.56	0.61	0.13	0.59	0.13	0.41

Dercreasing κ_w and κ_ρ								
	V1 (κ_0)		V4 ($\kappa_0/10$)		V5 ($\kappa_0/100$)		V6 ($\kappa_0/100$)	
Consumption	0.349572		0.350059		0.348929		0.352165	
Labor	0.352963		0.353450		0.352321		0.348845	
Wage	0.857613		0.857618		0.857606		0.866630	
	Prices	Wages	Prices	Wages	Prices	Wages	Prices	Wages
Frequency of adj.	10.1	6.02	10.1	6.41	10.41	7.28	54.4	6.95
Mean change	1.68	2.83	1.68	2.66	1.68	2.34	0.31	2.45
Std of prices/wages	5.26	3.07	5.26	1.21	5.26	0.75	6.06	0.96
Mean abs(Change)	8.57	6.14	8.57	2.70	8.57	1.98	4.76	2.29
Std of changes	10.6	8.53	10.6	2.10	10.6	1.26	5.26	1.79
Kurtosis of changes	3.20	10.0	3.20	5.21	3.20	3.62	1.94	3.33
Price/wage increses	58.5	72.1	58.5	92.2	58.5	99.1	54.0	92.0
Price/wage increses ($\leq$5%)	28.8	48.1	28.8	91.7	28.8	100	60.8	99.9
Price/wage increses ($\leq$2.5%)	14.2	24.7	14.2	57.4	14.2	81.1	20.0	61.6
Pricing cost as % Rev.	0.51	0.33	0.51	0.07	0.51	0.01	0.07	0.01
Timing cost as % Rev.	0.37	0.35	0.37	0.05	0.37	0.01	0.03	0.01
Loss. relative to Flex	1.78	0.57	1.78	0.40	1.78	0.42	0.13	0.41

Figure 2: Distribution of nonzero price changes: varying price stickiness.



Notes:

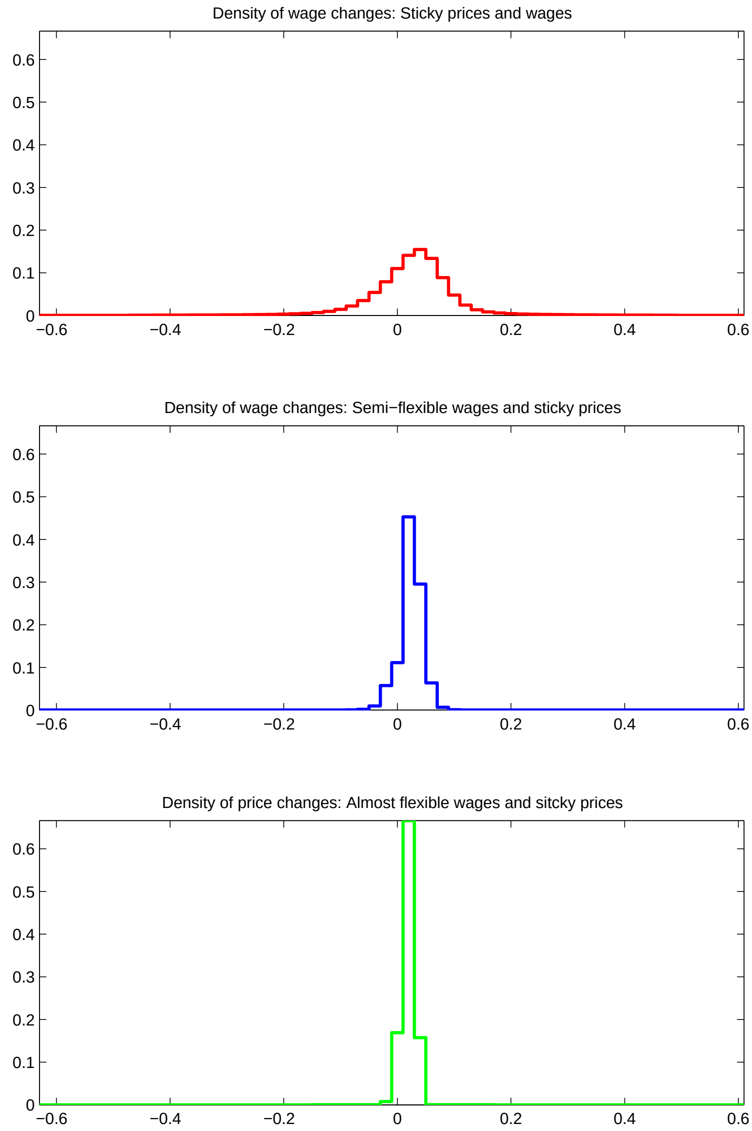
Effect of price stickiness on histogram of nonzero price adjustments.

Red: both prices and wages sticky (V1).

Blue: prices semi-flexible, wages sticky (V2).

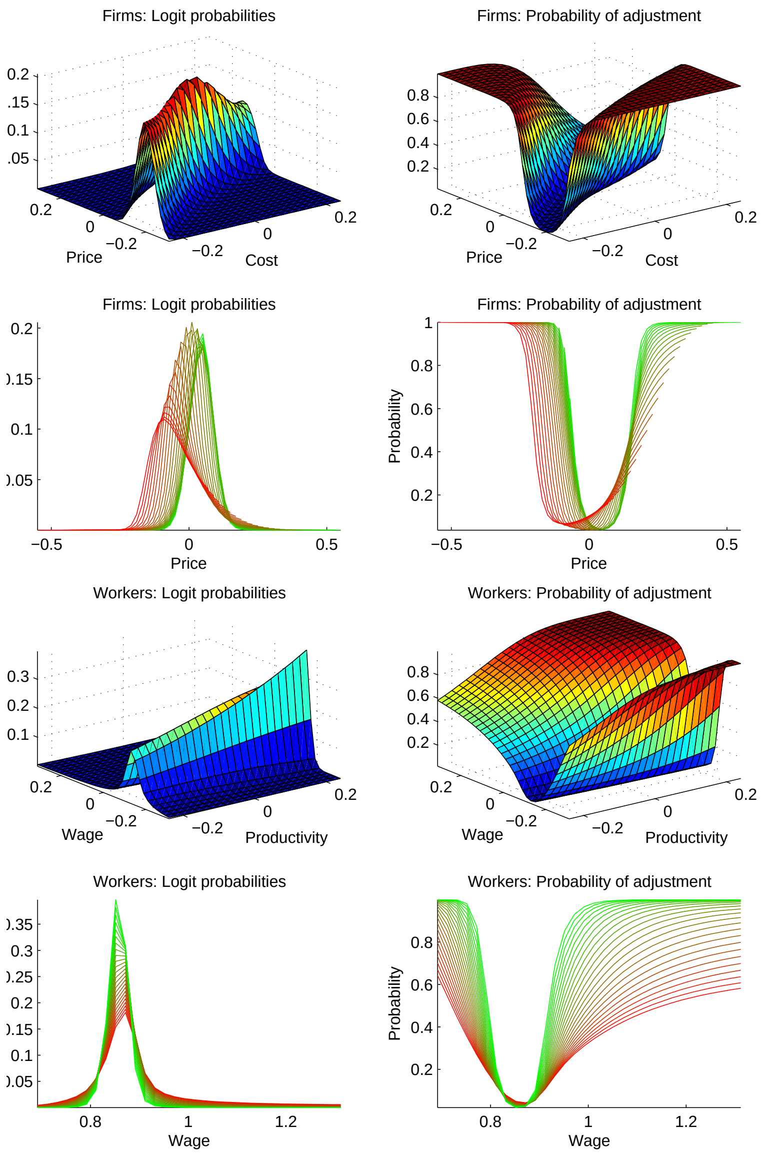
Green: prices flexible, wages sticky (V3).

Figure 3: Distribution of nonzero wage changes: varying wage stickiness.



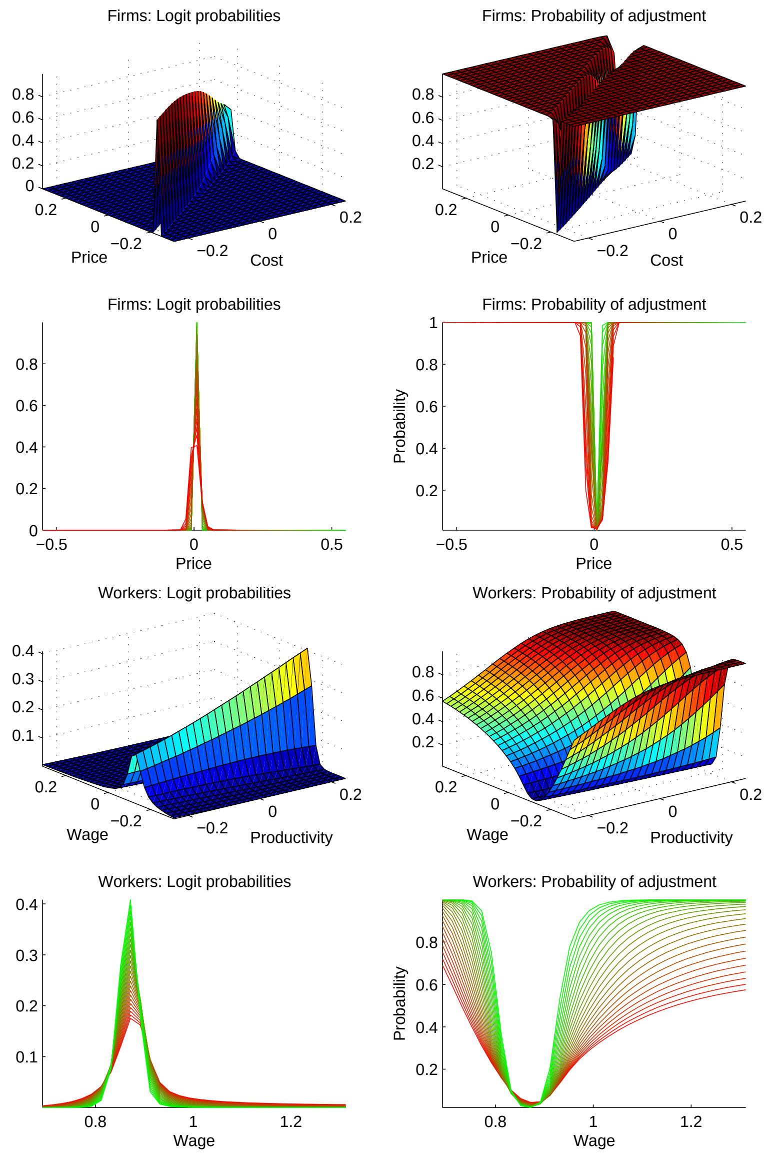
Notes:
Effect of wage stickiness on histogram of nonzero wage adjustments.
Red: both prices and wages sticky (V1).
Blue: prices sticky, wages semi-flexible (V4).
Green: prices sticky, wages flexible (V5).

Figure 4: Adjustment behavior. Benchmark model (V1) with sticky prices and sticky wages.



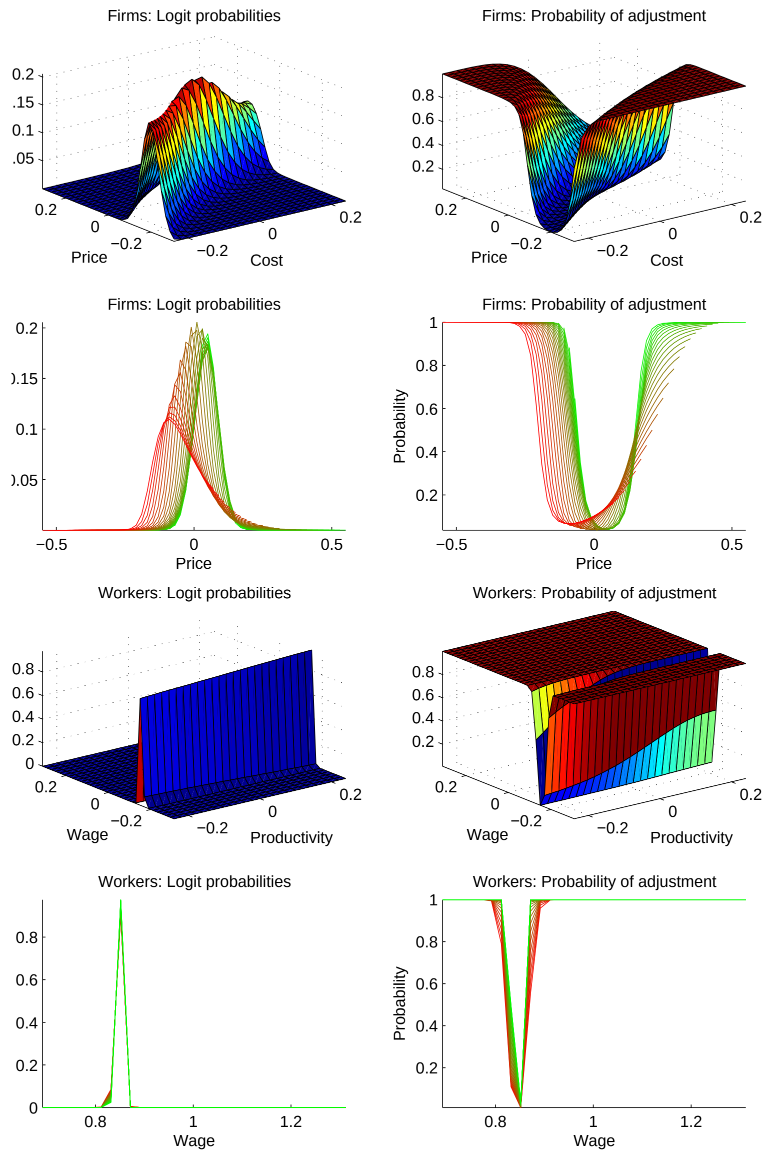
Notes: Distribution of adjustments and adjustment probability for prices (top four panels) and wages (bottom four panels).
Left panels: 3d plots show logit probabilities of each price (wage), conditional on cost.
Left panels: 2d plots show logit probabilities of each price (wage), conditional on cost, from green (low cost) to red (high cost).
Right panels: 3d plots show adjustment probabilities, conditional on current price (wage) and cost.
Right panels: 2d plots show adjustment probabilities, conditional on current price (wage) and cost, from green (low cost) to red (high cost).

Figure 5: Adjustment behavior. Model V3: flexible prices and sticky wages.



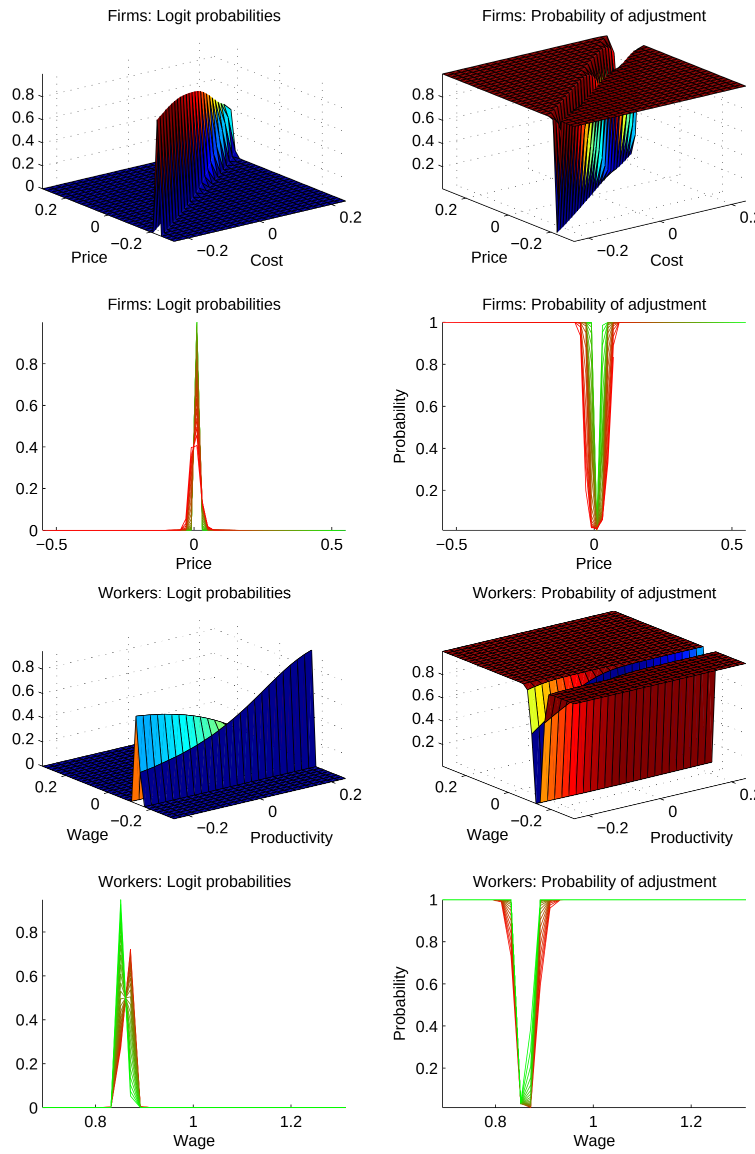
Notes: Distribution of adjustments and adjustment probability for prices (top four panels) and wages (bottom four panels).
Left panels: 3d plots show logit probabilities of each price (wage), conditional on cost.
Left panels: 2d plots show logit probabilities of each price (wage), conditional on cost, from green (low cost) to red (high cost).
Right panels: 3d plots show adjustment probabilities, conditional on current price (wage) and cost.
Right panels: 2d plots show adjustment probabilities, conditional on current price (wage) and cost, from green (low cost) to red (high cost).

Figure 6: Adjustment behavior. Model V5: sticky prices and flexible wages.



Notes: Distribution of adjustments and adjustment probability for prices (top four panels) and wages (bottom four panels).
Left panels: 3d plots show logit probabilities of each price (wage), conditional on cost.
Left panels: 2d plots show logit probabilities of each price (wage), conditional on cost, from green (low cost) to red (high cost).
Right panels: 3d plots show adjustment probabilities, conditional on current price (wage) and cost.
Right panels: 2d plots show adjustment probabilities, conditional on current price (wage) and cost, from green (low cost) to red (high cost).

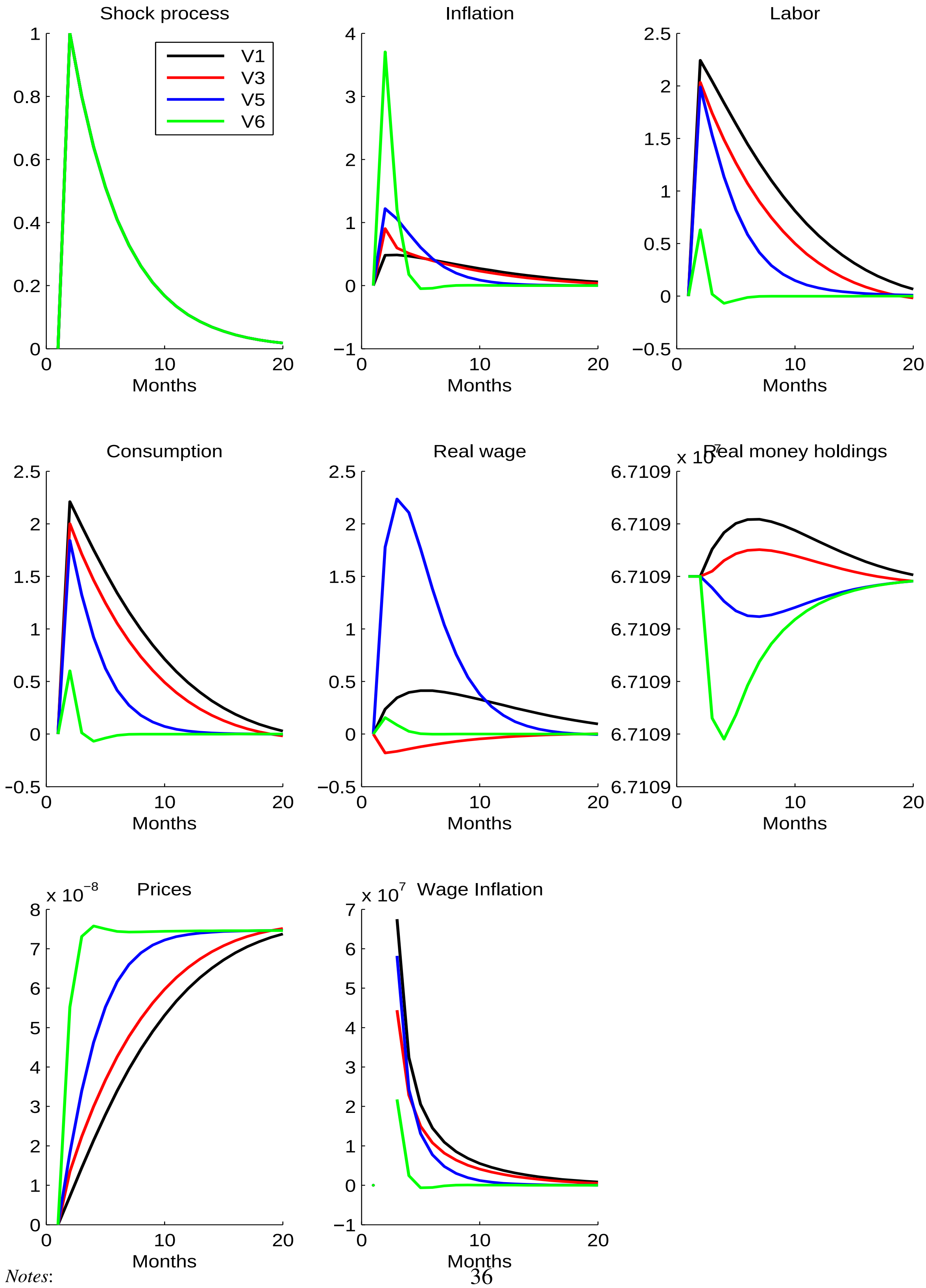
Figure 7: Adjustment behavior. Model V6: flexible prices and flexible wages.



Notes: Distribution of adjustments and adjustment probability for prices (top four panels) and wages (bottom four panels).
Left panels: 3d plots show logit probabilities of each price (wage), conditional on cost.
Left panels: 2d plots show logit probabilities of each price (wage), conditional on cost, from green (low cost) to red (high cost).
Right panels: 3d plots show adjustment probabilities, conditional on current price (wage) and cost.
Right panels: 2d plots show adjustment probabilities, conditional on current price (wage) and cost, from green (low cost) to red (high cost).

3.5 Results: money supply and TFP shocks.

Figure 8: Impulse responses to money growth shock: effects of nominal rigidity.



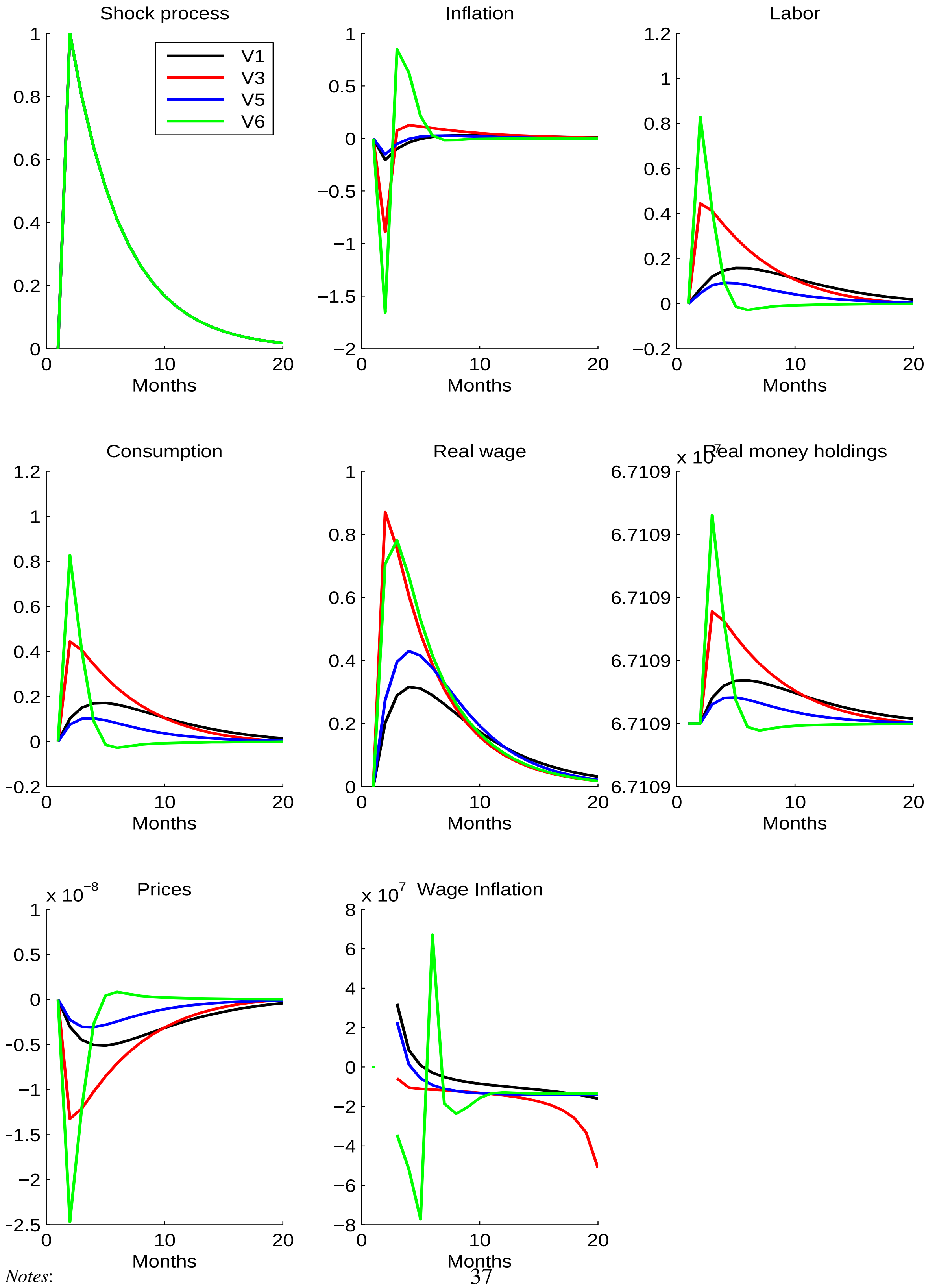
Notes:

Impulse responses of inflation and consumption to money growth shock with autocorrelation 0.8 (monthly).

Black: Benchmark (V1), both prices and wages sticky. Red: V3, flexible prices and sticky wages.

Blue: V5, sticky prices and flexible wages. Green: V6: both prices and wages flexible.

Figure 9: Impulse responses to TFP shock: effects of nominal rigidity.



Notes:

Impulse responses of inflation and consumption to TFP shock with autocorrelation 0.95 (monthly).

Black: Benchmark (V1), both prices and wages sticky. Red: V3, flexible prices and sticky wages.

Blue: V5, sticky prices and flexible wages. Green: V6: both prices and wages flexible.

4 Nonlinear disutility

As we have seen, generating a nontrivial wage distribution will require nonlinear disutility of labor. Here I will spell out several possible formulations. Some are tractable, others not. We will not include all these formulations in the paper. I'm just planning to state them here temporarily so that we can see why some setups are tractable and others are not.

Here are the main points I want to make.

1. When disutility is linear, the marginal cost of labor x appears in both steps of the decision problem: the choice of the adjustment hazard, and the choice of the wage distribution

2. Under nonlinear disutility, the marginal cost of labor $x(h) \equiv X'(h)$ varies; so if we write the decision problem in two steps, then the quantity of labor already used up in the first step becomes a state variable in the second-step problem. So the idiosyncratic state in the second step will be something like (w, z, h), where h is time already used up. Adding a third idiosyncratic dimension is unlikely to be tractable with our current methodology (MATLAB will run out of memory).

3. An alternative is to concatenate the two steps of the decision problem, assuming that the adjustment hazard and the price distribution are chosen simultaneously; I have explored this in my bargaining paper. By doing both decisions simultaneously, the intermediate step where h is a state variable is eliminated; the idiosyncratic state reduces to (w, z). However, maintaining the assumption (imposed in this paper) that today's decision can alter today's wage *immediately*, we find that the system of first-order conditions is no longer tractable.

4. The reason the simultaneous decision setup remains tractable in my bargaining paper is that I made a small change in the timing assumptions (motivated by my intention to consider the continuous-time limit). Instead of assuming that the time t wage decision affects the time t wage, I assumed that the time t wage decision affects the wage at time $t + 1$ (or $t + \Delta$, for some arbitrary time step Δ). This makes the first-order conditions tractable– I show that conditional on the current state (w, z) we can solve for the accuracy β, time use h, arrival rate ρ, and wage distribution $\vec{\pi}$ by solving a system of two equations. Hence a backwards induction step involves looping over all possible (w, z) to solve for the decision in all states.

5. That calculation is feasible, but at a monthly frequency it would be restrictive to assume you cannot adjust the current wage, only the next month's wage. Therefore if we do this specification I would prefer to move to a weekly frequency.

6. Calculation will be slower than because: (1) higher frequency; (2) two-dimensional root finding problem instead of explicit solution at each state (w, z); (3) GE requires solving wage problem and price problem simultaneously. Nonetheless, we are talking about minutes or maybe hours, not days.

(((SUBSECTION COMPARING VARIOUS NONLINEAR SPECIFICATIONS HAS BEEN DELETED)))

4.1 Agenda

I propose that we solve the model following the setup described in section **??**. This would involve the following steps.

1. In my bargaining paper I showed (Prop. 8) that the backwards induction step at a given state (w, z) could be reduced to a system of two equations, one sloping up and the other sloping down, implying a unique interior solution. The calculation here is either exactly identical, or very similar. It would be convenient (but not essential) to repeat that argument using the notation of this paper.

2. For my bargaining paper I wrote a MATLAB program called `twomargins_ns.m` to calculate that unique solution. We should check whether we can use exactly the same program, or whether it needs some reprogramming in order to apply it to this model.

3. When `twomargins_ns.m` is running, we can rewrite our backwards induction routine, starting from Borja's `firmsproblem.m` or Anton's `Viter.m`. The difference is that the previous problems implicitly loop over all (w, z) implicitly by using MATLAB matrix notation. Now we will instead need to write the loop explicitly. We will need to loop over all idiosyncratic states (w, z), calling `twomargins_ns.m` inside the loop to calculate the endogenous variables associated with each value of (w, z).

4. I don't think we need to loop inside Borja's `dynsys.m` (similar to Anton's `dyneqklein.m`) because those equations in effect define a system of functional equations (which we solve by linear approximation, without trying to isolate each state (w, z) from the others). But we will need to make changes in the program to reflect the altered timing suggested in section **??**, so that the choice at t affects prices at $t + \Delta$.

5. In this discussion I have concentrated on how nonlinear utility changes the structure of the partial equilibrium backwards induction calculation. But linear utility also implied major simplification in the structure of the steady-state general equilibrium fixed point calculation. For example, in LPD we argued that steady-state GE could be written as a fixed point in w, the steady state wage. With nonlinear utility that would be more complex; see footnote 22 in section 2.5 of LPD. In the current paper (numerical appendix) we described steady-state GE as a fixed point problem in C, N, and the *aggregate* wage w. We need to check whether that form of the problem is still valid under nonlinear utility, and we have to write `dynsys.m` accordingly.

6. Once we have those programs running, we can calibrate the model using the price data that we already used for LPD, and the wage data that Anton proposed recently. I suggest calibrating and simulating at weekly frequency.

7. We can then run simulations like the ones already reported in the current version of the paper.

8. When we finally write up the paper, we will eliminate most of the specifications reported in this version.

5 Conclusions

References

[1] Álvarez, Fernando; Francesco Lippi, and Luigi Paciello (2014), "Monetary shocks in models with inattentive producers." NBER Working Paper 20817.

[2] Álvarez, Fernando; Martín González-Rozada; Andy Neumeyer; and Martín Beraja (2011), "From hyperinflation to stable prices: Argentina's evidence on menu cost models." Manuscript, Univ. of Chicago.

[3] Blundell, Richard, and Stephen Bond (2000), "GMM Estimation with persistent panel data: an application to production functions." *Econometric Reviews* 19 (3): 321-40.

[4] Calvo, Guillermo (1983), "Staggered prices in a utility-maximizing framework." *Journal of Monetary Economics* 12: 383-98.

[5] Caplin, Andrew, and Daniel Spulber (1987), "Menu costs and the neutrality of money." *Quarterly Journal of Economics* 102 (4): 703-25.

[6] Costain, James (2016), "Costly decisions and sequential bargaining. " Manuscript, Banco de España.

[7] Costain, James, and Anton Nakov (2011), "Distributional dynamics under smoothly state-dependent pricing." *Journal of Monetary Economics* 58 (6-8): 646-65.

[8] Costain, James, and Anton Nakov (2015), "Logit price dynamics." CEPR Discussion Paper #10731.

[9] Cover, Thomas, and Joy Thomas (2006), *Elements of Information Theory*, 2nd ed. Wiley Interscience.

[10] Dotsey, Michael; Robert King; and Alexander Wolman (2013), "Inflation and real activity with firm-level productivity shocks." FRB Philadelphia Working Paper 13-35.

[11] Eichenbaum, Martin; Nir Jaimovich, and Sergio Rebelo (2011), "Reference prices, costs and nominal rigidities." *American Economic Review* 101 (1): 234-62.

[12] Erceg, Christopher, Dale Henderson, and Andrew Levin (2000), "Optimal monetary policy with staggered wage and price contracts." *Journal of Monetary Economics* 46: 281-313.

[13] Gagnon, Etienne (2009), "Price setting under low and high inflation: evidence from Mexico." *Quarterly Journal of Economics* 124: 1221-63.

[14] Golosov, Mikhail, and Robert E. Lucas, Jr. (2007), "Menu costs and Phillips curves." *Journal of Political Economy* 115 (2): 171-99.

[15] Kehoe, Patrick, and Virgiliu Midrigan (2012), "Prices are sticky after all." FRB Minneapolis Staff Report 413.

[16] Klenow, Peter, and Oleksiy Kryvtsov (2008), "State-dependent or time-dependent pricing: does it matter for recent US inflation?", *Quarterly Journal of Economics* 123: 863-904.

[17] Klenow, Peter, and Benjamin Malin (2010), "Microeconomic evidence on price setting." In *Handbook of Monetary Economics*, Benjamin Friedman and Michael Woodford, eds., Vol. 3, no. 3.

[18] Klein, Paul (2000), "Using the generalized Schur form to solve a multivariate linear rational expectations model." *Journal of Economic Dynamics and Control* 24 (10): 1405-23.

[19] Matejka, Filip, and Alisdair McKay (2011), "Rational inattention to discrete choices: a new foundation for the multinomial logit model." CERGE-EI Working Paper Series 442.

[20] Matejka, Filip (2010), "Rationally inattentive seller: sales and discrete pricing." CERGE-EI Working Paper 408.

[21] Mattsson, Lars-Goram, and Jorgen Weibull (2002), "Probabilistic choice and procedurally bounded rationality." *Games and Economic Behavior* 41: 61-78.

[22] Midrigan, Virgiliu (2011), "Menu Costs, Multi-Product Firms and Aggregate Fluctuations." *Econometrica* 79 (4): 1139-80.

[23] Nakamura, Emi, and Jón Steinsson (2008), "Five facts about prices: a reevaluation of menu cost models." *Quarterly Journal of Economics* 123 (4): 1415-64.

[24] Reiter, Michael (2009), "Solving heterogeneous-agent models by projection and perturbation." *Journal of Economic Dynamics and Control* 33 (3): 649-65.

[25] Stahl, Dale (1990), "Entropy control costs and entropic equilibrium. " *International Journal of Game Theory* 19: 129-38.

[26] Takahashi, Shuhei (2015), "State dependency in price and wage setting." KIER Working Paper #918, Kyoto University.

[27] Van Damme, Eric (1991), *Stability and Perfection of Nash Equilibrium*, 2nd edition. Springer Verlag.

[28] Woodford, Michael (2008), "Inattention as a source of randomized discrete adjustment." Manuscript, Columbia University.

[29] Woodford, Michael (2009), "Information-constrained state-dependent pricing." *Journal of Monetary Economics* 56: S100-S124.

Computational appendix

Outline of algorithm

Computing this model is challenging due to heterogeneity. At any time t, firms will face different productivity shocks A_{jt} and will be stuck at different prices P_{jt}; likewise productivity and wages will vary across workers. The Calvo model is popular because, up to a first-order approximation, only the average price matters for equilibrium. But this property does not hold in most models; here we must treat all equilibrium quantities as functions of the time-varying distribution of prices and productivity across firms.

We address this problem by implementing Reiter's (2009) solution method for dynamic general equilibrium models with heterogeneous agents and aggregate shocks. As a first step, the algorithm calculates the steady-state general equilibrium in the absence of aggregate shocks. Idiosyncratic shocks are still active, but are assumed to have converged to their ergodic distribution, so the real aggregate state of the economy is a constant, Ξ. The algorithm solves for a discretized approximation to this steady state, restricting all idiosyncratic state variables to discrete grids. That is, real log prices p_{jt} lie at all times in a fixed grid $\gamma^p \equiv \{p^1, p^2, ...p^{\#^p}\}$; real log wages w_{it} lie in $\gamma^w \equiv \{w^1, w^2, ...w^{\#^w}\}$; and likewise for log productivities of firms and workers: $a_{jt} \in \gamma^a \equiv \{a^1, a^2, ...a^{\#^a}\}$ and $z_{it} \in \gamma^z \equiv \{z^1, z^2, ...z^{\#^z}\}$. The four grids γ^p, γ^w, γ^a, and γ^z are all assumed to have constant step sizes (in logs) between grid points. Moreover, we assume (only for numerical convenience) that the step size in γ^w equals that in γ^p, and also that the number of grid points is the same in these two grids: $\#^w = \#^p$.

We can then view firms' steady state value function as a matrix $\mathbf{V}$ of size $\#^p \times \#^a$, comprising the values $v^{jk} \equiv v(p^j, a^k, \Xi)$ associated with prices and productivities $\left(p^j, a^k\right) \in \gamma^p \times \gamma^a$.[22] Similarly, the distribution of firms can be viewed as a $\#^p \times \#^a$ matrix $\mathbf{\Psi}$ in which the row j, column k element Ψ^{jk} represents the fraction of firms in state (p^j, a^k) at the end of any given period. Likewise, the workers' steady-state value function and the end-of-period distribution of workers can be represented by matrices $\mathbf{L}$ and $\mathbf{\Psi^w}$ of size $\#^w \times \#^z$. While these matrices are large objects, we can nonetheless solve for a steady-state general equilibrium as a low-dimensional root-finding problem. By guessing the steady-state values of C and N, we can set up the Bellman equations of the workers and firms, and solve for their fixed points $\mathbf{L}$ and $\mathbf{V}$; given optimal policies, we can describe the dynamics of the distributions, and thus solve for the steady-state distributions $\mathbf{\Psi^w}$ and $\mathbf{\Psi}$; knowing the distributions, we will show that we can construct two scalar equations that suffice to check the values of C and N.

In a second step, Reiter's method constructs a linear approximation to the dynamics of the discretized model, by perturbing it around the steady state general equilibrium on a point-by-point basis. That is, the firms' value function is represented by a $\#^p \times \#^a$ matrix $\mathbf{V}_t$ with row j, column k element $v_t^{jk} \equiv v(p^j, a^k, \Xi_t)$, thus summarizing the time t values at all grid points $(p^j, a^k) \in \gamma^p \times \gamma^a$. Then, instead of viewing the Bellman equation as a functional equation that defines $v(p, a, \Xi)$ for all possible

[22] In this appendix, bold face indicates matrices, and (most) superscripts represent indices of matrices or grids.

idiosyncratic and aggregate states p, a, and Ξ, we think of it as an expectational relation between the matrices $\mathbf{V}_t$ and $\mathbf{V}_{t+1}$. This amounts to a (large!) system of $\#^p\#^a$ first-order expectational difference equations that determine the dynamics of the $\#^p\#^a$ variables v_t^{jk}. In addition, there will be a relation between the workers' values $\mathbf{L}_t$ and $\mathbf{L}_{t+1}$ at times t and $t+1$, which can also be seen as a system of $\#^w\#^z$ scalar equations in $\#^w\#^z$ unknowns. Finally, the distribution of firms at time $t+1$, $\mathbf{\Psi}_{t+1}$ is derived from the distribution at time t, $\mathbf{\Psi}_t$, which amounts to $\#^p\#^a$ scalar equations; and the distributional dynamics of workers links the distributions $\mathbf{\Psi}_t$ and $\mathbf{\Psi^w}_{t+1}$ with a matrix equation that is equivalent to a system of $\#^w\#^z$ scalar equations.

We linearize these equations numerically (together with a handful of scalar equations, including first-order conditions for some aggregate variables). We then solve for the saddle-path stable solution of the linearized model using the QZ decomposition, following Klein (2000). It is crucial to note here that our problem is tractable because we have separated the two sticky decisions in our model between two different classes of decision-makers. In a model where a single decision-maker adjusted p and w in response to the shocks a and z, the value function and distributional dynamics would both have to be evaluated over $\#^p\#^w\#^a\#^z$ grid points. Solving for dynamic general equilibrium would require solving a system of slightly more than $2\#^p\#^w\#^a\#^z$ equations. Instead, since we have assumed prices and wages are set by different agents, we will have to solve slightly more than $2\#^p\#^a + 2\#^w\#^z$ equations, which is a vastly smaller problem.[23]

The discretized model

Firms' values are summarized by matrices $\mathbf{V}_t$ and $\mathbf{V}_t^e$, of size $\#^p \times \#^a$, and the vector $\tilde{\mathbf{v}}_t$, of length $\#^a$. Workers' values are described by the matrices, $\mathbf{L}_t$, $\mathbf{L}_t^e$, and $\widetilde{\mathbf{L}}_t$, of size $\#^w \times \#^z$. The elements of $\mathbf{V}_t$ are $v_t^{jk} \equiv v(p^j, a^k, \Xi_t)$, and the elements of $\mathbf{V}_t^e$ are $v_t^{e,jk} \equiv v^e(p^j, a^k, \Xi_t)$, for $\left(p^j, a^k\right) \in \gamma^p \times \gamma^a$. Likewise, $\mathbf{L}_t$ has elements $l_t^{jk} \equiv l(w^j, z^k, \Xi_t)$, and $\mathbf{L}_t^e$ has elements $l_t^{e,jk} \equiv l^e(w^j, z^k, \Xi_t)$, for $\left(w^j, z^k\right) \in \gamma^w \times \gamma^z$. The expected values of setting a new price or wage are given by vectors $\tilde{\mathbf{v}}_t$ and $\tilde{\mathbf{L}}_t$, with elements $\tilde{v}_t^k \equiv \tilde{v}(a^k, \Xi_t)$ and $\tilde{l}_t^{jk} \equiv \tilde{l}(w^j, a^k, \Xi_t)$.

Related matrices include the adjustment values of firms and workers, $\mathbf{D}_t$ and $\mathbf{D^w}_t$; the probability matrices of firms and workers, $\mathbf{\Lambda}_t$ and $\mathbf{R}_t$; and the option values of firms and workers, $\mathbf{O}_t$ and $\mathbf{O^w}_t$. The

[23]In other words, computational complexity under our approach scales exponentially with the number of sticky decisions if these decisions are all taken by the same agent, but scales linearly in the number of sticky decisions if different decisions are controlled by different agents. (Actually, the same principle is true in models of fully flexible decisions, but the issue is more relevant here because *stickiness creates heterogeneity*— while prices and wages are jump variables in flexible models, in the presence of nominal rigidity they become state variables.)

(j,k) elements of these matrices are given by[24]

$$d_t^{jk} \equiv \tilde{v}_t^k - v_t^{jk}\,, \quad d_t^{w,jk} \equiv \tilde{l}_t^k - l_t^{jk}\,; \tag{104}$$

$$\lambda_t^{jk} \equiv \lambda\left(d_t^{jk}/(\kappa_\lambda w_t)\right), \quad \rho_t^{jk} \equiv \rho\left(d_t^{jk}/(\kappa_\rho \xi_t^{jk})\right); \tag{105}$$

Finally, we also define the logit probabilities $\mathbf{\Pi}_t$ (a matrix) and $\mathbf{\Pi^w}_t$ (a 3d array). The elements of these matrices are

$$\pi_t^{jk} = \pi_t(p^j|a^k) \equiv \frac{\eta^j \exp\left(v_t^{jk}/(\kappa_\pi w_t)\right)}{\sum_{n=1}^{\#p} \eta^n \exp\left(v_t^{nk}/(\kappa_\pi w_t)\right)}\,, \tag{106}$$

$$\pi_t^{w,jkn} = \pi_t^w(w^n|w^j, z^k) \equiv \frac{\eta^{w,n} \exp\left(l_t^{nk}/(\kappa_w \xi_t^{jk})\right)}{\sum_{m=1}^{\#w} \eta^{w,m} \exp\left(l_t^{mk}/(\kappa_w \xi_t^{jk})\right)}\,. \tag{107}$$

Here π_t^{jk} is the probability that a firm which has decided to adjust its price at time t chooses real log price p^j, conditional on log productivity a^k; $\pi_t^{w,jkn}$ is a worker's corresponding probability of choosing the real log wage w^n, conditional on current log real wage w^j and log productivity z^k. The default probabilities for log real prices $p \in \gamma^p$ are $\boldsymbol{\eta} \equiv (\eta^1, \ldots, \eta^{\#^p}) \equiv (\eta(p^1), \ldots, \eta(p^{\#^p}))$, and $\boldsymbol{\eta}^{\mathbf{w}} \equiv (\eta^{w,1}, \ldots, \eta^{w,\#^w}) \equiv (\eta^w(w^1), \ldots, \eta^w(w,^{\#^w}))$ is the analogous vector for log real wages $w \in \gamma^w$.

In this discrete representation, the productivity processes (77) and (85) can be summarized by matrices $\mathbf{S}$ and $\mathbf{S^z}$ of size $\#^a \times \#^a$ and $\#^z \times \#^z$. The (m,k) elements of these matrices represent the following transition probabilities, respectively:

$$S^{mk} = prob(a_{jt} = a^m | a_{j,t-1} = a^k), \quad S^{z,mk} = prob(z_{it} = z^m | z_{i,t-1} = z^k). \tag{108}$$

It is helpful to introduce analogous Markovian notation to describe the deflation of real prices and wages as the aggregate price level rises. Let $\mathbf{T}_t$ be a $\#^p \times \#^p$ Markov matrix in which the row m, column l element represents the probability that firm j's beginning-of-period log real price $\tilde{p}_{jt}$ equals $p^m \in \gamma^p$ if its log real price at the end of the previous period was $p^l \in \gamma^p$:

$$T_t^{ml} \equiv prob(\widetilde{p}_{jt} = p^m | p_{j,t-1} = p^l). \tag{109}$$

Generically, the deflated log price $\tilde{p}_{jt} \equiv p_{j,t-1} - i_t \equiv p_{j,t-1} - i(\Xi_t, \Xi_{t-1})$ will fall between two grid points; then the matrix $\mathbf{T}_t$ must round up or down stochastically. Also, if $p_{j,t-1} - i_t$ lies below the smallest or above the largest element of the grid, then $\mathbf{T}_t$ must round up or down to keep prices on the

[24] Actually, (105) is a simplified description of λ_t^{jk}. While (105) implies that λ_t^{jk} represents the function $\lambda(\bullet)$ evaluated at the log price grid point p^j and log productivity grid point a^k, in our computations λ_t^{jk} actually represents the *average* of $\lambda(\bullet)$ over all log prices in the interval $\left(\frac{p^{j-1}+p^j}{2}, \frac{p^j+p^{j+1}}{2}\right)$, given log productivity a^k. Calculating this average requires interpolating the function $d_t(p, a^k)$ between price grid points. Defining λ_t^{jk} this way ensures differentiability with respect to changes in the aggregate state Ξ_t.

grid.[25] Therefore we construct $\mathbf{T}_t$ according to

$$T_t^{ml} = prob(\widetilde{p}_{jt} = p^m | p_{j,t-1} = p^l, i_t) = \begin{cases} 1 & \text{if } p^l - i_t \leq p^1 = p^m \\ \frac{p^l - i_t - p^{m-1}}{p^m - p^{m-1}} & \text{if } p^1 < p^m = \min\{p \in \Gamma^p : p \geq p^l - i_t\} \\ \frac{p^{m+1} - p^l + i_t}{p^{m+1} - p^m} & \text{if } p^1 \leq p^m = \max\{p \in \Gamma^p : p < p^l - i_t\} \\ 1 & \text{if } p^l - i_t > p^{\#^p} = p^m \\ 0 & \text{otherwise} \end{cases} \tag{110}$$

Furthermore, recall that we have assumed that the price and wage grids γ^p and γ^w have the same step size, and the same number of grid points. Note that in this case, the transition probabilities mapping real log wages from one period to the beginning of the next are the same as those for real log prices. In other words, for all m and l,

$$prob(\widetilde{w}_{it} = w^m | w_{j,t-1} = w^l) \;=\; prob(\widetilde{p}_{jt} = p^m | p_{j,t-1} = p^l) \;=\; T_t^{ml}. \tag{111}$$

Thus we can describe the distributional dynamics of wages using exactly the same matrix $\mathbf{T}_t$ that we used from prices.

Given this notation, we can now write the distributional dynamics in a more compact form. The time t distributions of firms and workers are derived from the distributions at the end of $t-1$ as follows:

$$\mathbf{\Psi}_t = \mathbf{T}_t \widetilde{\mathbf{\Psi}}_{t-1} \mathbf{S}', \qquad \mathbf{\Psi}_t^w = \mathbf{T}_t \widetilde{\mathbf{\Psi}}_{t-1}^w (\mathbf{S}^{\mathbf{z}})'. \tag{112}$$

Note that exogenous shocks are represented from left to right in the matrices $\widetilde{\mathbf{\Psi}}_t$ and $\widetilde{\mathbf{\Psi}}_t^w$, so that their transitions can be treated by right multiplication, while sticky decision variables are represented vertically, so that transitions related to choice variables can be described by left multiplication. Next, to calculate the effects of price adjustment on the distribution, let $\mathbf{E}_{pp}$, $\mathbf{E}_{pa}$, $\mathbf{E}_{ww}$, and $\mathbf{E}_{wz}$ be matrices of ones of size $\#^p \times \#^p$, $\#^p \times \#^a$, $\#^w \times \#^w$, and $\#^w \times \#^z$, respectively. After production occurs at time t, as new real prices are set, the price distribution adjusts as follows:

$$\widetilde{\mathbf{\Psi}}_t \;=\; (\mathbf{E}_{pa} - \mathbf{\Lambda}_t) \odot \mathbf{\Psi}_t + \mathbf{\Pi}_t \odot (\mathbf{E}_{pp}(\mathbf{\Lambda}_t \odot \mathbf{\Psi}_t)). \tag{113}$$

where the operator $\odot$ represents element-by-element multiplication (the Hadamard product). Equivalently, we can write the dynamics in summation notation as follows:

$$\tilde{\psi}_t^{jk} \;=\; (1 - \rho_t^{jk}) \psi_t^{jk} + \sum_i \pi_t^{j|k} \rho_t^{ik} \psi_t^{ik}. \tag{114}$$

[25] In other words, we assume that any nominal price that would have a real log value less than p^1 after inflation is automatically adjusted upwards to the real log value p^1 (and when computing examples with deflation we must adjust down any real log price exceeding $p^{\#^p}$). This assumption is made for numerical purposes only, and has a negligible impact on the equilibrium as long as we choose a sufficiently wide grid γ^p.

The matrix notation does not carry over to the wage dynamics, because the distribution of new wages varies with the *current* wage, so here the summation notation is unavoidable:

$$\tilde{\psi}_t^{w,jk} \;=\; (1-\rho_t^{jk})\psi_t^{w,jk} + \sum_{i,k} \pi_t^{j|ik}\rho_t^{ik}\psi_t^{w,ik}. \tag{115}$$

The same transition matrices $\mathbf{T}_t$, $\mathbf{S}$, and $\mathbf{S^z}$ show up when we write the Bellman equations in matrix form. The discounted values of choosing each possible real price $\tilde{p}$ are

$$\mathbf{V}_t^e = \beta E_t \left\{ \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} \mathbf{T}'_{t+1}\mathbf{V}_{t+1}\mathbf{S} \right\}, \qquad \mathbf{L}_t^e = \beta E_t \left\{ \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} \mathbf{T}'_{t+1}\mathbf{L}_{t+1}\mathbf{S^z} \right\}. \tag{116}$$

Here the expectation E_t refers only to the effects of the time $t+1$ aggregate shock g_{t+1}, because the dynamics of the idiosyncratic states (p_{jt}, a_{jt}) and (w_{it}, z_{it}) are completely described by the matrices $\mathbf{T}'_{t+1}$, $\mathbf{S}$, and $\mathbf{S^z}$.

The discounted values of choosing each possible price and wage are given by:

$$\tilde{v}_t^k \;\equiv\; \kappa_\pi w_t \ln\left(\sum_{j=1}^{\#^p} \eta^j \exp\left(\frac{v_t^{e,jk}}{\kappa_\pi w_t}\right)\right), \qquad \tilde{l}_t^{jk} \;\equiv\; \kappa_w x_t \ln\left(\sum_{n=1}^{\#^w} \eta^{w,n} \exp\left(\frac{l_t^{e,nk}}{\kappa_w \xi_t^{jk}}\right)\right). \tag{117}$$

The elements of the matrices $\mathbf{O}$ and $\mathbf{O}^w$ can then be calculated as

$$o_t^{jk} = v_t^{e,jk} + \kappa_\lambda w_t \ln\left(1-\bar{\lambda}+\bar{\lambda}\exp\left(\frac{d^{jk}}{\kappa_\lambda w_t}\right)\right) \tag{118}$$

$$o_t^{w,jk} = l_t^{e,jk} + \kappa_\rho \xi_t^{jk} \ln\left(1-\bar{\rho}+\bar{\rho}\exp\left(\frac{d^{w,jk}}{\kappa_\rho \xi_t^{jk}}\right)\right) \tag{119}$$

$$\tag{120}$$

Now, let $\mathbf{U}_t$ be the $\#^p \times \#^a$ matrix of current payoffs to the firm, with elements

$$u_t^{jk} \equiv \left(\exp(p^j) - \frac{w_t}{\exp(a^k)}\right)\frac{C_t}{\exp(\epsilon p^j)} \tag{121}$$

for $\left(p^j, a^k\right) \in \gamma^p \times \gamma^a$. The corresponding payoff matrix for the workers can be written $\mathbf{U^w}_t$. of size $\#^w \times \#^z$. To define its elements, let $h_t^{jk} \equiv h_t(w^j, z^k)$ be labor demand in state (w^j, z^k, Ξ_t). Then the elements of $\mathbf{U^w}_t$ are

$$u_t^{w,jk} \equiv \exp(w^j)h_t^{jk} - \frac{X(h_t^{jk}+\tau_t^{jk}+\mu_t^{jk})}{u'(c(\Xi_t))} \tag{122}$$

for $\left(w^j, z^k\right) \in \gamma^w \times \gamma^z$. We also define the real marginal value of time as $\xi_t^{jk} = \frac{X'(h_t^{jk}+\tau_t^{jk}+\mu_t^{jk})}{u'(c(\Xi_t))}$.

Then we can calculate the value functions as

$$\mathbf{V}_t = \mathbf{U}_t + \mathbf{O}_t \qquad \mathbf{L}_t = \mathbf{U^w}_t + \mathbf{O}_t^w \tag{123}$$

In order to check labor market clearing it will be helpful to define several summary statistics related to labor time use. First, let K_t^λ and K_t^π and be total time use for choosing the timing of the price decision, and actually choosing prices:

$$K_t^\lambda = \sum_{j=1}^{\#p}\sum_{k=1}^{\#a} \psi_t^{jk} \left(\lambda_t^{jk} \ln\left(\frac{\lambda_t^{jk}}{\bar{\lambda}}\right) + (1-\lambda_t^{jk}) \ln\left(\frac{1-\lambda_t^{jk}}{1-\bar{\lambda}}\right)\right), \tag{124}$$

$$K_t^\pi = \sum_{j=1}^{\#p}\sum_{k=1}^{\#a} \psi_t^{jk} \lambda_t^{jk} \left(\sum_{i=1}^{\#p} \pi_t^{ik} \ln\left(\frac{\pi_t^{ik}}{\eta^k}\right)\right), \tag{125}$$

$$\Delta_t = \sum_{j=1}^{\#p}\sum_{k=1}^{\#a} \psi_t^{jk} \exp(-\epsilon p^j - a^k). \tag{126}$$

Note that in the second equation, the time K_t^π devoted to choosing prices is weighted by the fraction adjusting, λ_t^{jk}. In the third equation, Δ_t represents a price dispersion measure that relates time devoted to production to total goods produced.

Next, we discuss how we apply the two steps of Reiter's (2009) method to this discrete model.

Step 1: steady state

In the aggregate steady state, aggregate shocks are zero; the distribution of firms takes some unchanging value $\mathbf{\Psi}$, and the distribution of workers takes some unchanging value $\mathbf{\Psi^w}$. Thus the aggregate state of the economy is constant: $\Xi_t \equiv (g_t, \Psi_{t-1}, \Psi_{t-1}^w) = (0, \mathbf{\Psi}, \mathbf{\Psi^w}) \equiv \Xi$. We indicate the steady state of all equilibrium objects by dropping the time subscripts and the function argument Ξ, so the steady state value function $\mathbf{V}$ has elements $v^{jk} \equiv v(p^j, a^k, \Xi)$.

Long run monetary neutrality in steady state implies that the rate of nominal money growth equals the rate of inflation:

$$\mu = \exp(i).$$

Thus, the steady-state transition matrix $\mathbf{T}$ is known, since it depends only on steady state inflation i. Morever, the Euler equation reduces to

$$\exp(i) = \beta R,$$

which simply serves to determine the nominal interest rate R.

We can then calculate general equilibrium as a three-dimensional root-finding problem, by guessing consumption C, labor demand N, and the aggregate wage level w. On one hand, knowing $c(\Xi)$ and $w(\Xi)$ we can construct the firm's profit function $u(p,a,\Xi) = (e^p - w(\Xi)e^{-a})c(\Xi)e^{-\epsilon p}$. Knowing the profit function, we can solve the firm's problem by backwards induction, which yields the value functions v, v^e, and $\tilde{v}$, and the policy functions λ and π. Given the firm's policy functions, we can calculate the distributional dynamics to find the steady-state distribution of prices and productivities, $\Psi(p,a)$. From the firm's problem and the steady-state distribution we can also calculate the time firms devote to decision-making (K_t^λ and K_t^π), and the efficiency wedge Δ.

On the other hand, knowing $n(\Xi)$ and $w(\Xi)$ we can construct the labor demand function $h(w,z,\Xi) = e^{z(\epsilon_n-1)}n(\Xi)w(\Xi)^{\epsilon_n}e^{-\epsilon_n w}$, and given $c(\Xi)$ we can also calculate worker's utility value of labor income, $u'(c(\Xi))e^w h(w,z,\Xi)$. We can then solve the worker's Bellman equation by backwards induction. This yields the vaue functions l, l^e, and $\tilde{l}$, and the policy functions ρ, and π^w, as well as the time use function τ and μ, and the worker's marginal value of time ξ. Given the worker's policy functions, we can calculate the distributional dynamics to find the steady-state distribution of wages and productivities, $\Psi^w(w,z)$.

With these distributions in hand, we can then check whether the guessed values of C, N, and w are consistent with an equilibrium. Then we check the following three scalar equations:

$$1 = \sum_{j=1}^{\#^p}\sum_{k=1}^{\#^a}\psi^{jk}\exp\left((1-\epsilon)p^j\right), \tag{127}$$

$$w = \left\{\sum_{j=1}^{\#^p}\sum_{k=1}^{\#^a}\psi^{w,jk}\exp\Big((1-\epsilon_n)(w^j-z^k)\Big)\right\}^{\frac{1}{1-\epsilon_n}}. \tag{128}$$

$$N = \Delta C + \kappa_\pi K^\pi + \kappa_\lambda K^\lambda. \tag{129}$$

The first two equations are the aggregate price and wage identities; the last is the labor market clearing condition. If these three equations are satisfied with sufficient accuracy, then a steady-state general equilibrium has been found.

Step 2: linearized dynamics

(((EVIDENTLY THE GE CALCULATION IS DIFFERENT NOW, SINCE P_t IS NOW PREDETERMINED. PRESUMABLY THE PROBLEM INVOLVES CALCULATING EXPECTED INFLATION $E_t i_{t+1}$.)))

Given the steady state, the general equilibrium dynamics can be calculated by linearization. To do so, we eliminate as many variables from the equation system as we can, reducing it to

$$g_{t+1} = \phi_g g_t + \epsilon_{t+1}^g \tag{130}$$

$$\frac{\mu\exp(g_t)}{\exp i_t} = \frac{m_t}{m_{t-1}} \tag{131}$$

$$1 - \frac{\nu}{m_t C_t^{-\gamma}} = \beta E_t\left(\frac{C_{t+1}^{-\gamma}}{i_{t+1}C_t^{-\gamma}}\right) \tag{132}$$

$$\widetilde{\mathbf{\Psi}}_t = (\mathbf{E}_{pa} - \mathbf{\Lambda}_t) \odot \mathbf{\Psi}_t + \mathbf{\Pi}_t \odot (\mathbf{E}_{pp}(\mathbf{\Lambda}_t \odot \mathbf{\Psi}_t)) \tag{133}$$

$$\widetilde{\mathbf{\Psi}}_t^{\mathbf{w}} = (\mathbf{E}_{wz} - \mathbf{R}_t) \odot \mathbf{\Psi^w}_t + \mathbf{\Pi^w}_t \odot (\mathbf{E}_{ww}(\mathbf{R}_t \odot \mathbf{\Psi^w}_t)) \tag{134}$$

$$\mathbf{V}_t^e = \beta E_t \left\{ \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} \mathbf{T}'_{t+1} \mathbf{V}_{t+1} \mathbf{S} \right\} \tag{135}$$

$$\mathbf{L}_t^e = \beta E_t \left\{ \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} \mathbf{T}'_{t+1} \mathbf{L}_{t+1} \mathbf{S^z} \right\} \tag{136}$$

$$\mathbf{V}_t = \mathbf{U}_t + \mathbf{O}_t \tag{137}$$

$$\mathbf{L}_t = \mathbf{U^w}_t + \mathbf{O^w}_t \tag{138}$$

$$1 = \sum_{j=1}^{\#^p} \sum_{k=1}^{\#^a} \Psi_t^{jk} \exp((1-\epsilon)p^j) \tag{139}$$

$$w_t^{1-\epsilon_n} = \sum_{j=1}^{\#^p} \sum_{k=1}^{\#^a} \Psi_t^{w,jk} \exp((1-\epsilon_n)(w^j - z^k)) \tag{140}$$

$$N_t \;=\; \Delta_t C_t + \kappa_\pi K_t^\pi + \kappa_\lambda K_t^\lambda \tag{141}$$

If we now collapse all the endogenous variables into a single vector

$$\overrightarrow{X}_t \equiv \left(vec\left(\mathbf{\Psi}_{t-1}\right)',\ vec\left(\mathbf{\Psi^w}_{t-1}\right)',\ vec\left(\mathbf{V}_t\right)',\ vec\left(\mathbf{L}_t\right)',\ C_t,\ N_t,\ w_t,\ m_{t-1},\ i_t \right)'$$

then the whole set of expectational difference equations (130)-(139) governing the dynamic equilibrium becomes a first-order system of the following form:

$$E_t \mathcal{F}\left(\overrightarrow{X}_{t+1}, \overrightarrow{X}_t, g_{t+1}, g_t \right) = 0 \tag{142}$$

where E_t is an expectation conditional on g_t and all previous shocks.

The first three equations, (130)-(132), state that the nominal money growth rate g is consistent with money supply and demand; all variables in these equations are included explicitly in $\overrightarrow{X}_t$ or $\overrightarrow{X}_{t+1}$, except for g_t and g_{t+1}. But some variables that appear in the next equations are not counted explicitly in the $\overrightarrow{X}$ vectors, because they have been substituted out of the list. These variables can be calculated using some simple intratemporal relations that are not counted in the equation list (130)-(139). To see that the variables in vector $\overrightarrow{X}_t$ do in fact suffice to evaluate the equation system, we will describe the substitutions, step by step.

First, note that given i_t and i_{t+1} we can construct the matrices $\mathbf{T}_t$ and $\mathbf{T}_{t+1}$. Given $\mathbf{T}_t$, we can construct $\widetilde{\mathbf{\Psi}}_t = \mathbf{T}_t \mathbf{\Psi}_{t-1} \mathbf{S}'$ and $\widetilde{\mathbf{\Psi}}_t^{\mathbf{w}} = \mathbf{T}_t \mathbf{\Psi}_{t-1}^w (\mathbf{S^z})'$ and from $\mathbf{\Psi}_{t-1}$ and $\mathbf{\Psi^w}_{t-1}$.

Given $\mathbf{V}_t$ and $\mathbf{L}_t$, we can calculate $\tilde{\mathbf{v}}_t$ and $\tilde{\mathbf{l}}_t$ from (117). We can then calculate $\mathbf{D}_t$ and $\mathbf{D^w}_t$, from (104) and the policy functions $\mathbf{\Lambda}_t$, $\mathbf{R}_t$, $\mathbf{\Pi}_t$, and $\mathbf{\Pi^w}_t$ from (105) and 106. We then have all the variables needed to evaluate the distributional dynamics (133)-(134).

Likewise, given $\mathbf{V}_{t+1}$ and $\mathbf{L}_{t+1}$, we can calculate $\tilde{\mathbf{v}}_{t+1}$ and $\tilde{\mathbf{l}}_{t+1}$ from (117), then $\mathbf{D}_{t+1}$, $\mathbf{D^w}_{t+1}$, $\mathbf{O}_{t+1}$ and $\mathbf{O^w}_{t+1}$ from (104) and (**??**). Given N_t, C_t, w_t and $x_t = \chi C_t^{\gamma}$, we can construct $\mathbf{U}_t$ and $\mathbf{U^w}_t$ from (121) and (122). Thus we have all the variables needed to evaluate the Bellman equations (137)-(138).

Finally, given $\widetilde{\mathbf{\Psi}}_t$ and $\mathbf{\Psi}_t$, we can calculate the formulas on the right-hand side of the price and wage identities (139)-(140). Given also $\mathbf{\Lambda}_t$, and $\mathbf{R}_t$, we can calculate K_t^{λ}, K_t^{π}, and Δ_t from (124)-(126), so we can evaluate equation (141). Therefore the variables in $\overrightarrow{X}_t$ and $\overrightarrow{X}_{t+1}$, together with g_t and g_{t+1}, are indeed sufficient to evaluate the system (130)-(141).

Finally, if we linearize system $\mathcal{F}$ numerically with respect to all its arguments to construct the Jacobian matrices $\mathcal{A} \equiv D_{\overrightarrow{X}_{t+1}}\mathcal{F}$, $\mathcal{B} \equiv D_{\overrightarrow{X}_t}\mathcal{F}$, $\mathcal{C} \equiv D_{g_{t+1}}\mathcal{F}$, and $\mathcal{D} \equiv D_{g_t}\mathcal{F}$, then we obtain the following first-order expectational difference equation system:

$$E_t \mathcal{A} \Delta \overrightarrow{X}_{t+1} + \mathcal{B} \Delta \overrightarrow{X}_t + E_t \mathcal{C} g_{t+1} + \mathcal{D} g_t = 0 \tag{143}$$

where Δ represents a deviation from steady state. This system has the form considered by Klein (2000), so we solve our model using his QZ decomposition method.[26]

[26] Alternatively, the equation system can be rewritten in the form of Sims (2001). We chose to implement the Klein method because it is especially simple and transparent to program.