THE MACROECONOMICS OF STICKY PRICES WITH GENERALIZED HAZARD
FUNCTIONS

Fernando E. Alvarez
Francesco Lippi
Aleksei Oskolkov

The Macroeconomics of Sticky Prices with Generalized Hazard Functions
Fernando E. Alvarez, Francesco Lippi, and Aleksei Oskolkov
NBER Working Paper No. 27434
June 2020
JEL No. C41,C61,E31

ABSTRACT

We give a thorough analytic characterization of a large class of sticky-price models where the firm's price setting behavior is described by a generalized hazard function. Such a function provides a tractable description of the firm's price setting behavior and allows for a vast variety of empirical hazards to be fitted. This setup is microfounded by random menu costs as in Caballero and Engel (1993) or, alternatively, by information frictions as in Woodford (2009). We establish two main results. First, we show how to identify all the primitives of the model, including the distribution of the fundamental adjustment costs and the implied generalized hazard function, using the distribution of price changes or the distribution of spell durations. Second, we derive a sufficient statistic for the aggregate effect of a monetary shock: given an arbitrary generalized hazard function, the cumulative impulse response to a once-and-for-all monetary shock is given by the ratio of the kurtosis of the steady-state distribution of price changes over the frequency of price adjustment times six. We prove that Calvo's model yields the upper bound and Golosov and Lucas' model the lower bound on this measure within the class of random menu cost models.

Fernando E. Alvarez
University of Chicago
Department of Economics
1126 East 59th Street
Chicago, IL 60637
and NBER
f-alvarez1@uchicago.edu

Francesco Lippi
Luiss University and EIEF
francescolippi@gmail.com

Aleksei Oskolkov
Department of Economics
University of Chicago
Saieh Hall
1126 E. 59th St.
Chicago, IL 60637
aoskolkov@uchicago.edu

1 Introduction and Summary of Results

Several sticky price models employ a "generalized hazard function", a function relating the firm's price adjustment probability to its state. Such a function provides a tractable description of the firm's price setting behavior and allows for a vast variety of empirical hazards to be fitted. Compared to the workhorse Calvo (1983) model, where the adjustment probability is constant, a generalized hazard function $\Lambda(x)$ allows this probability to depend on the state x, the firm's desired adjustment, like the markup deviation from the desired level. Such state dependence is appealing theoretically, see e.g. Barro (1972); Dixit (1991); Caplin and Spulber (1987); Golosov and Lucas (2007), and has been found to be relevant empirically, see e.g. Eichenbaum, Jaimovich, and Rebelo (2011); Gautier and Saout (2015). The notion of a generalized hazard function, and its derivation from first principles, were developed in seminal "menu-cost" papers by Caballero and Engel (1993a, 1999, 2007) and Dotsey, King, and Wolman (1999), and later revisited using information theoretical foundations by Woodford (2009) and Costain and Nakov (2011b). Several authors have since employed the generalized hazard function in applications and empirical work.[1]

This paper mostly follows the setup introduced by Caballero and Engel (2007) to frame a broad class of sticky price models where the firm's pricing decisions are represented by a generalized hazard function $\Lambda(|x|)$. The symmetry of the function arises since we focus on economies where the state is driftless and is subject to idiosyncratic productivity shocks, an accurate benchmark for low-inflation economies.[2] A large number of models are nested by this framework: starting with two "extreme" versions such as the canonical Calvo model with a constant hazard $\Lambda(x) = \lambda$ and unbounded x, and the Golosov and Lucas (2007) model with x bounded by the adjustment threshold $\pm X$ and a zero hazard on $|x| \in (0, X)$ with a spike that can be thought of as an "infinite hazard" at the adjustment thresholds,[3] and other intermediate cases, such as the so called Calvo-

---

[1]For recent applications see e.g. Costain and Nakov (2011a); Carvalho and Kryvtsov (2018); Sheremirov (2019); for empirical work see e.g. Berger and Vavra (2018); Petrella, Santoro, and de la Porte Simonsen (2018), and for related theoretical work Baley and Blanco (2019).

[2]See proposition 7 in Alvarez, Le Bihan, and Lippi (2016) for a result explaining why inflation has no first order effects on the propagation of monetary shocks in this class of models. See Baley and Blanco (2019) for extensions to large drift, as in the case of large inflation.

[3]The infinite hazard at the threshold should be thought of as an approximation of the behaviour at an sS barrier. In our results we make the different behaviour precise. In Proposition 4 we produce a rigorous approximation of the behaviour at barrier as a very high hazard rate, which justifies this analogy.

plus model by Nakamura and Steinsson (2010), the random menu cost problem of Dotsey and Wolman (2020), as well as the cases discussed above which explicitly use a generalized hazard function.

We employ this setup to prove five new analytical results which gives a thorough understand the workings of sticky price models, their mapping to the data, and the propagation of monetary shocks. First, we establish an *invertible* mapping between the fundamental cost of price adjustment (menu cost or information cost) and the "reduced form" generalized hazard function. We consider two possible foundations underlying this mapping. One, introduced in the seminal work by Caballero and Engel (1993b), assumes the firm can change its price upon paying a fixed (menu) cost ψ that is drawn every period from an *unrestricted* distribution of costs $G(\psi)$. We prove that the mapping between *any* given menu cost distribution $G(\psi)$ and the generalized hazard function $\Lambda(x)$ is invertible. This means that *any* non-decreasing generalized hazard function can be rationalized by a *unique* choice of the distribution of the random fixed costs $G(\psi)$.

While the non-decreasing nature of the generalized hazard function was established by Caballero and Engel (1993a), we prove the invertibility of the mapping and give an explicit formula to recover $G(\psi)$ from any $\Lambda(x)$ non-decreasing in $|x|$. We also provide an identical result for an alternative foundation with the firm optimally selecting the "probability" of adjustment opportunity in every period subject to a cost $c(\cdot)$, a simplified version of Woodford (2009), where the cost is modeled in a rational inattention framework.[4] We show that every non-decreasing generalized hazard rate Λ can be rationalized by a convex cost function $c(\cdot)$.

Second, we derive a mapping between the theory and the data on the size-distribution of price changes, which have been heavily used to discipline sticky price models over the past two decades (see e.g. Klenow and Malin (2010); Cavallo and Rigobon (2016)). We provide an invertible mapping between the (observed) distribution of price changes $Q(\Delta p)$ with density $q(\Delta p)$, frequency of price changes N_a, and the generalized hazard function $\Lambda(x)$. A price change $\Delta p = -x$ is chosen by a firm with desired adjustment x that is given the option to adjust. A straightforward relation links the density of price changes $q(-x)$ to the hazard function and the cross-sectional distribution of

---

[4]See also Costain and Nakov (2011b) for a sticky price model where firms must pay a cost to increase the probability of a price change.

desired adjustments, $f(x)$, namely $q(-x)N_a = \Lambda(x)f(x)$. Previous contributions such as Berger and Vavra (2018) retrieve $f(x)$ by postulating a parametric form for $\Lambda(x)$ and then using this relation. We show that, surprisingly to us, $\Lambda(x)$ and $f(x)$ are both fully encoded in $q(\Delta p)$ and N_a, and that it is possible to identify *both* functions using the distribution and frequency of price changes alone. We derive the expression to retrieve $f(x)$ and $\Lambda(x)$ from $q(\Delta p)$ and N_a in closed form. The recovery of the function $f(x)$ from observables relates to the work by Baley and Blanco (2019) who obtain all the moments of $f(x)$ even in the presence of drift and asymmetries. Using our first result, we can then recover the entire distribution of random menu cost, $G(\psi)$. We propose an estimator for such distributions that is consistent with the theory, and allows for unobserved heterogeneity among products.

We complement these results by establishing that, under regularity conditions, the survival function $S(t)$, measuring the distribution of *durations* of unchanged prices, uniquely identify $\Lambda(x)$ as well. Summarizing, the generalized hazard rate Λ is identified either by the distribution of price changes Q and one temporal statistic (the frequency), or by the distribution of durations S and one statistic on the size of price changes (the variance).

To illustrate our proposed procedure, we estimate the underlying distribution of price changes, measure its kurtosis, and recover $\Lambda(x)$, $f(x)$, and $G(\psi)$ using publicly available scraped-price data of Cavallo's (2015). Interestingly, accounting for measurement error and aggregation, and correcting for unobserved heterogeneity, we find values around 2, much smaller values of kurtosis than those typically reported in the literature, and roughly consistent with the quadratic generalized rate.

Furthermore, we define a statistic $\mathcal{C}$ (for "Calvo-ness") that measures the fraction of price changes happening *independently* of the state of the firm. Using our characterization of the relationship between the observed distribution of price changes and the generalized hazard rate, we show that $\mathcal{C}$ is proportional to $q(0)$, the density of price changes near zero. We estimate $\mathcal{C}$ in Cavallo's (2015) data set and found it to be about 6%, i.e. about 94% of price changes show some state dependence.

Third, we establish that the cumulative impulse response (CIR) of output to a once-and-for-all monetary shock in *any* model characterized by a generalized hazard function $\Lambda(x)$ is a simple

function of two steady state statistics: the Kurtosis of the distribution of price changes divided by six times the frequency of price changes. The CIR, namely the area under the output impulse response function, is a convenient summary measure of the non-neutrality of monetary shocks. The notion of CIR was introduced in Alvarez, Le Bihan, and Lippi (2016), which showed that the "kurtosis result" holds for a Calvo-plus model (and multi-product firms), a class that implies a constant hazard function $\Lambda(x) = \lambda$ in the inaction region. Alvarez, Lippi, and Paciello (2016) showed that the kurtosis results holds in a large class of rational inattention models, as proposed by Reis (2006), which are purely time dependent.[5] This paper provides a substantive extension to the previous cases: we establish that the kurtosis result holds for *any* symmetric $\Lambda(x)$ function, allowing for both finite and infinite X.[6] This includes decreasing or non-monotone hazard functions (which are not rationalized by random menu cost models), and hazards with discontinuities corresponding to mass points in the distribution of menu cost. For instance, this provides a rigorous (negative) answer to Dotsey and Wolman (2020) who conjecture, based on numerical simulations of a different model, that the "kurtosis result" may fail to apply in a model with random menu costs.[7]

The fourth contribution establishes that within the class of non-decreasing generalized hazard functions the largest Kurtosis is six, attained by the constant hazard rate model, like the pure Calvo (1983) case. The smallest one, equal to one, corresponds to the pure (non-random) menu cost model of Golosov and Lucas (2007). This result is interesting because non-decreasing hazard rates are an implication of either the random menu cost or the information gathering setup, and thus it establishes Calvo as an upper bound within this broad classes of models. Indeed, due to our first result, the class of non-decreasing generalized hazard functions describes the entire set of models with random menu costs or the entire class of information gathering models. For a Kurtosis higher than six, the value of the pure Calvo model, one would need to come up with an economic foundation for a (locally) *decreasing* hazard function.

---

[5]Our initial use of the cumulative response was restricted to problems with zero drift (inflation). In Alvarez and Lippi (2019, 2020) we extend the result to the cumulated output response for problems featuring a non-zero but small inflation. Baley and Blanco (2019) extend our result, characterizing the CIR in terms of objects related to the distribution of price gaps, to any moment of interest for the CIR, and allowing for non-zero drift and asymmetries.

[6]The only case where this result was already established is the Calvo$^+$ model with $\Lambda(x)$ constant on $(0, X)$.

[7]We conjecture that the reason their numerical simulations fail to establish this property is due to the presence of a fraction of firms that have flexible prices. In such settings it is essential to properly aggregate across the different types of firms. Failing to do so will obfuscate the result, which holds for each firm's type.

The fifth contribution studies the scope of the Flexibility Index, $\mathcal{F}$, a notion introduced by Caballero and Engel (2007) to analyze the link between the microeconomic behavior and aggregate stickiness. Subsequently this measure has been used by some authors as a summary measure of monetary non-neutrality, see e.g. Berger and Vavra (2018); Petrella, Santoro, and de la Porte Simonsen (2018). We show that the index corresponds to the slope of the impulse response on impact (right after the shock). For models with barriers, where X is bounded, the flexibility index is not informative about the non-neutrality since it is always infinite. For models without barriers, where X is unbounded, $\mathcal{F}$ is finite, and can be computed in closed form for isoelastic hazards Λ, including the widely used quadratic case. We present non-pathological examples where the same $\mathcal{F}$ corresponds to different values of the *cumulative* impulse response, and where $\mathcal{F}$ is not even an accurate summary of the short-term response of output.

Structure of the paper. The next section provides two foundations for the generalized hazard function. In the first one (Section 2.1), firms choose when to change prices subject to random menu costs, distributed according to CDF G. In the second one (Section 2.2), firms choose the intensity with which they can change prices, subject to a cost function c. In both models, the optimal decision rule is summarized by a generalized hazard function Λ. We show that in both models, given Λ, one can recover the primitive cost, either the G or c. Section 3 characterizes the steady-state statistics of a model where the firms' decisions follow a generalized hazard function. Section 4 shows how to recover Λ starting from an observed distribution of the size of price changes, and estimate it in Cavallo's (2015) data set. Section 5 establishes that, under some conditions, the information encoded by the size distribution of price changes can equivalently be obtained from the distribution of spell durations. Section 6 discusses the propagation of a once-and-for-all small aggregate shock in an economy characterized by a generalized hazard function, and proves that its effect can be summarized by a simple sufficient statistic. Section 7 discusses scope and limitations of the flexibility index, a summary measure for the non-neutrality of monetary shock that is often used in the literature.

2 Foundations of the Generalized hazard function $\Lambda(x)$

The generalized hazard function is a building block of several macro models featuring sticky prices. It is a function that maps the state of the firm, x, e.g. the deviation of the current markup from the profit maximizing one, into the *likelihood* of a price adjustment $\Lambda(x)$. Such a function is appealing to scholars because it allows for substantive flexibility in fitting cross-sectional data on price setting behavior while, at the same time, having explicit microeconomic foundations. This section presents two alternative settings for such foundations, and provides an invertible mapping that allows one to recover the foundations from a given hazard function.

Our first setup uses a random menu cost model, first proposed by Caballero and Engel (1993a) and elaborated in Caballero and Engel (1999, 2007). A particular case, the Calvo-plus model, was analyzed by Nakamura and Steinsson (2010). The second setup relates to models of inattention as in Woodford (2009), where firms choose the arrival rate of opportunities to change prices.[8]

Both setups feature a firm that maximizes the present discounted value of a per-period profit function given by $-Bx^2$, a second order approximation of the profit function, where x is the price gap, and the parameter $B > 0$ measures the curvature of the profit function. If prices are not changed, the price gap x evolves as a standard Brownian Motion with zero drift and variance σ^2. The lack of drift indicates that the economy under consideration has no inflation.[9] The two setups differ in the friction that prevents the firm from setting $x = 0$ at all times. In the first, the friction is due to the presence of random fixed costs of price adjustment; in the second the friction is due to an information cost.

2.1 The Random Menu Cost Model

The Calvo-plus model supplements the traditional Calvo model with the possibility that the firm can change its price by paying a fixed menu cost at any time. The advantage of this model is

---

[8]In Woodford (2009) the form of the firm's problem and the specification of $c(\cdot)$ are derived assuming constraints on information flows.

[9]See The Online appendix B in Alvarez and Lippi (2014) for a detailed microfoundation of this model. The focus on a model with zero inflation provides an accurate approximation for economies where inflation is low, as the effects on decision rules are of second order when inflation is close to zero, as shown theoretically and validated empirically in Alvarez et al. (2019); Alvarez and Lippi (2020).

to eliminate a long tail of delayed adjustments that seems counterfactual. The generalized model allows the firm to draw a fixed menu cost ψ from a distribution G at random times – arriving at a Poisson rate $\kappa > 0$.

As in Caballero and Engel (1993a), we call the difference between the current price of the firm and its ideal price a "price gap". We will specify the process for the demand and cost of the firm, so that the price gap is the state of the firm's problem. The menu costs drawn by the firm can be zero or strictly positive. If the cost is zero the firm changes its price to the ideal one (i.e. it "closes its price gap"). If the firm draws a strictly positive cost, it will either ignore it or change its price depending on the value of the "price gap" relative to the realization of the fixed cost. In particular, the optimal decision rule will be characterized by a threshold rule that gives the maximum adjustment cost that the firm is willing to pay for adjustment. For all fixed costs smaller than the threshold the firm changes its price, while for larger costs it keeps the price unchanged.

We also allow the firm to have a price change at any time by paying a (relatively large) fixed cost, which we denote by $\Psi > 0$ and refer to as the "deterministic fixed cost". If $\Psi = \infty$, then the firm has no such alternative. We can write the value function of the firm, $v(x)$, as:

$$rv(x) = \min\left\{ Bx^2 + \frac{\sigma^2}{2}v''(x) + \kappa \int_0^\Psi \min\left\{\psi + \min_{\tilde{x}} v(\tilde{x}) - v(x),\, 0\right\}dG(\psi),\, r\left(\Psi + \min_{\tilde{x}} v(\tilde{x})\right)\right\}$$

Two points are worth making. First, given the symmetry of Bx^2, the value function is *symmetric around* $x = 0$. A proof can be constructed by a simple guess and verify argument. Second, if $\Psi = \infty$ then $X = \infty$, and thus there is no second branch in the Bellman equation. Note that as long that either $r > 0$ and/or that $\kappa > 0$, the value function v is finite and well defined in the case of $\Psi = \infty$.

The term $\min_{\tilde{x}} v(\tilde{x})$ is the value right after adjustment, and given the symmetry of the return function, we have $v(0) = \min_{\tilde{x}} v(\tilde{x})$. Thus we can simply write that for all x

$$rv(x) = \min\left\{ Bx^2 + \frac{\sigma^2}{2}v''(x) + \kappa \int_0^\Psi \min\left\{\psi + v(0) - v(x),\, 0\right\}dG(\psi),\, r\left(\Psi + v(0)\right)\right\}$$

7

For the case where $\Psi < \infty$ we can use that the optimal decision rule has a threshold $X < \infty$ such that if $|x| \geq X$ the firm pays the fixed cost Ψ. Thus we can write

$$rv(x) = \begin{cases} Bx^2 + \frac{\sigma^2}{2}v''(x) + \kappa \int\limits_{0}^{v(x)-v(0)} \psi dG(\psi) + \kappa \left[v(0) - v(x)\right] G\left(v(x) - v(0)\right), & |x| \leq X \\ r\left(v(0) + \Psi\right), & |x| > X \end{cases} \tag{1}$$

Note that we can define the threshold function $\bar{x} : [0, \Psi] \to [0, X]$ as solving

$$v(\bar{x}(\psi)) = v(0) + \psi \ \text{ for all } \psi \in [0, \Psi] \tag{2}$$

It is easy to see that v is increasing in $|x|$, since the period cost Bx^2 is strictly increasing in $|x|$, the uncontrolled process is a brownian motion, and the adjustment cost is independent of x. Since v is strictly increasing in $[0, X]$, then $\bar{x}'(\psi) = 1/v'(\bar{x}(\psi)) > 0$. We can let the function $\bar{\psi}(x)$ be the inverse of $\bar{x}(\psi)$.

For simplicity, in the characterization of the problem that follows we will assume a distribution function G with a continuous density. We require G to be continuously differentiable at all points, with the possible exception of $\psi = 0$. For completeness, Appendix C considers the case of a discrete distribution G, where ψ takes finitely many values.[10] In either case we have the following smooth pasting and optimal return point conditions:

$$v'(-X) = v'(X) = v'(0) = 0 \tag{3}$$

We are now ready to define the *generalized hazard rate*, $\Lambda : (-X, X) \to \mathbb{R}_+$, which gives the probability (per unit of time) that a firm with $x \in (-X, X)$ will change its price. It is defined by the optimal decision rule, or the value function, as well the Poisson arrival rate $\kappa > 0$ and the distribution of fixed cost G:

$$\Lambda(x) = \kappa \, G\left(v\left(x\right) - v\left(0\right)\right) \ \text{ for all } x \in (-X, X) \ . \tag{4}$$

---

[10]The two cases differ on whether $\bar{x}(\cdot)$ is a continuous function, and on whether the value function $v(\cdot)$ is twice differentiable everywhere or it has jump discontinuities on finitely many values. Indeed in the latter case we need to rewrite the value function since $v''(x)$ is not defined at all points.

The function Λ is symmetric around $x = 0$ and weakly increasing in $|x|$. It is continuous at x if G is continuous at $\psi = v(x) - v(0)$, and bounded above by κ. While the function Λ is not defined at $x = \pm X$, we abuse notation and let $\Lambda(X) = \lim_{x \to X} \Lambda(x) = \kappa G(\Psi) = \kappa$.

2.1.1 Rationalizing a given generalized hazard Λ

We next show that any increasing, differentiable, symmetric and bounded hazard rate Λ can be rationalized as the solution to the firm problem in equation (1) by a unique menu cost distribution G and two parameters $\{\kappa, \Psi\}$. Our proof is constructive: we provide an algorithm to compute $\{G, \kappa, \Psi\}$ from Λ, proving existence and uniqueness. Indeed G is obtained by solving a linear ordinary differential equation of the second order. Section 2.2 describes an alternative problem of the firm that also generates a non-decreasing generalized hazard function. We find this interesting because it allows us to relate to setups costly information collection, as in Woodford (2009).

The main result in this section shows how to recover the distribution G, with a density $G' = g$, given Λ and the values of three parameters: r, B, and σ^2. Three remarks are in order. First, the values of the fixed costs ψ are measured relative to B, and thus the optimal decision rules depend only on the distribution of ψ/B. Second, we show that σ^2, while in principle unobservable, is encoded in the frequency and variance of price changes. Thus, once Λ is given, we can recover all the parameters of the firm's problem, except the discount rate r. Third, while in this section we consider the case where G is differentiable for $\psi > 0$ to simplify the exposition, Appendix C considers discrete distributions of costs which imply an hazard Λ that is a step function. In this case we can recover G starting from Λ by solving a system of linear equations.

Assume the firm faces a distribution G of the menu costs with a density g for all $\psi > 0$, and possibly a mass point at $\psi = 0$. In this case its Bellman equation solves

$$rv(x) = Bx^2 + \frac{\sigma^2}{2}v''(x) + \int_0^{v(x)-v(0)} \kappa[\psi + v(0) - v(x)]g(\psi)d\psi + \kappa[v(0) - v(x)]G(0) \qquad (5)$$

for all $x \in [-X, 0]$, and we can use the symmetry of v to define it as $v(x) = v(-x)$. The boundary conditions are $v'(X) = 0$ and $v(X) = v(0) + \Psi$, the smooth pasting and value matching. Note that

9

in the interior $(0, X)$ the function v solves a *non-linear* ordinary differential equation.

Before the main result on the existence of a unique invertible mapping between Λ and G, we state an intermediate result that provides a solution for the value function v and a new auxiliary function that will be used to solve the general problem. Consider the function Λ describing the probability per unit of time of a price adjustment if the price gap is $|x| < X$. We have the following:

LEMMA 1. Let the function u solve the linear ordinary differential equation

$$[r + \Lambda(x)] u(x) = 2Bx + \frac{\sigma^2}{2} u''(x) \text{ for } x \in [0, X] \tag{6}$$

with boundary conditions $u(0) = u(X) = 0$. The solution for u is unique. Moreover, v is given by

$$v(x) = u'(0)\frac{\sigma^2}{2r} + \int_0^x u(z)dz \quad \text{for } x \in [0, X] . \tag{7}$$

The auxiliary function u can readily be used to compute the value function and, as shown below, to characterize the distribution of costs that rationalizes the postulated hazard function. We now state the main result of this section:

PROPOSITION 1. Fix a discount rate $r > 0$, the curvature of the profit function $B > 0$, the volatility of shocks $\sigma > 0$, and the threshold X, with $X \in \mathbb{R}_+ \cup \{+\infty\}$. Consider a generalized hazard function $\Lambda(\cdot) : (-X, X) \to \mathbb{R}_+$ that is symmetric around zero, increasing in $|x|$, differentiable on $(0, X)$, and bounded. There exist real numbers $\{\kappa > 0, \Psi > 0\}$, both positive, and a cost distribution $G(\cdot) : [0, \Psi] \to [0, 1]$ with a density $g(\cdot)$, continuous on $(0, \Psi)$, and possibly a mass point $G(0) > 0$, that uniquely rationalizes Λ with a value function that solves equation (5). Using the auxiliary function u in Lemma 1 and $U(x) = \int_0^x u(z)dz$ for $x \in (0, X)$,

$$\kappa = \lim_{x \uparrow X} \Lambda(x) \quad , \quad \Psi = U(X) \quad , \quad G(0) = \frac{\Lambda(0)}{\kappa} \tag{8}$$

$$g(U(x)) = \frac{\Lambda'(x)}{u(x)\kappa} \text{ for all } x \in (0, X) \text{ with } \psi = U(x) \tag{9}$$

10

The proposition allows us to retrieve the primitives of a fully specified price setting problem starting from any given non-decreasing hazard function Λ. Note that whenever $\Lambda(0) > 0$ the model implies a mass point at $\psi = 0$. Intuitively, rationalizing a non-zero probability of adjustment when the gap is small requires a mass point of zero menu costs. Also note that $g(\cdot) > 0$ requires $\Lambda'(\cdot) > 0$.

Application: a quadratic hazard function. We conclude with an application to a quadratic generalized hazard function $\Lambda(x) = \Lambda_0 + \Lambda_2 x^2$ where $\Lambda_0 \geq 0, \Lambda_2 \geq 0$ and $|x| \in [0, X]$.[11] We can solve for the auxiliary function $u(x)$ using Lemma 1. This yields a polynomial:

$$u(x) = \sum_{i=0}^{\infty} a_{2i+1} \, x^{2i+1} \tag{10}$$

satisfying the ODE in equation (6) and the boundary conditions $u(0) = u(X) = 0$. Straightforward application of the method of undetermined coefficients gives the recursive relation

$$a_3 = \frac{(r + \Lambda_0)a_1 - 2B}{3\sigma^2} \tag{11}$$

$$a_{2i+1} = \frac{(r + \Lambda_0)a_{2i-1} + \Lambda_2 a_{2i-3}}{\sigma^2 \, i(2i+1)} \quad \text{for} \quad i = 2, 3, ... \tag{12}$$

All coefficients are determined as a function of a_1, which is pinned down by the boundary condition $u(X) = 0$. Application of Proposition 1 gives $U(x) = \sum_{i=1}^{\infty} \frac{a_{2i-1}}{2i} x^{2i}$, the value function $v(x) = a_1 \frac{\sigma^2}{2r} + U(x)$, the arrival rate $\kappa = \Lambda_0 + \Lambda_2 X^2$, the distribution function $G(\psi)$ with

$$G(0) = \frac{\Lambda_0}{\Lambda_0 + \Lambda_2 X^2} \tag{13}$$

Note that if $\Lambda_0 > 0$ the proposition implies a mass point at $\psi = 0$. For $|x| \in [0, X]$, the proposition gives the menu cost density function

---

[11]If Λ is symmetric and smooth, it often admits a quadratic approximation close to zero. This feature, mentioned by Caballero and Engel (2007) and Berger and Vavra (2018), makes quadratic generalized hazard functions especially appealing. In Appendix I we show that if Λ does not admit a quadratic approximation around $x = 0$, the underlying density g exhibits non-generic behavior.

$$g(U(x)) = \frac{2\Lambda_2 x}{\kappa u(x)} \tag{14}$$

The limit of the density is finite and positive, $\lim_{\psi \downarrow 0} g(\psi) = \frac{2\Lambda_2}{\kappa a_1}$. This happens because $v(x)$ is smooth and symmetric, so $u(x) = v'(x)$ admits a linear approximation close to zero.

2.2 An Optimal Adjustment Intensity Model

In this section we describe an alternative setup that yields a similar mapping between an underlying cost function and the generalized hazard. Now the firm does not face random menu cost. Instead, it directly controls the arrival rate of a free opportunity to change prices. At each moment the firm must pay a flow cost $c(\ell)$ to obtain an arrival rate ℓ. We assume that the flow cost is increasing and convex. This will give rise to the choice of the optimal rate of price changes as a function of the price gap, leading to a generalized hazard function Λ. As in the previous case, we also allow the firm to pay a deterministic menu cost Ψ to change its price with certainty. This Ψ will give rise to a barrier X, and we allow $\Psi = \infty$, in which case this will never be used, so $X = \infty$.

The main result of this section is that, analogously to the previous setup, any increasing symmetric function Λ can be rationalized by some increasing and convex cost function c. One difference between the two setups is that the resulting Λ in this setup does not need to be bounded above. This justifies the use of some of our examples later on. Additionally, this setup imposes fewer constraints on the tails of the implied distribution of price changes.

The firm's problem is:

$$rv(x) = \min \left\{ Bx^2 + \frac{\sigma^2}{2} v''(x) + \min_{\ell \geq 0} \{ \ell \left(v(0) - v(x) \right) + c(\ell) \}, \, r \left(\Psi + v(0) \right) \right\} \tag{15}$$

We assume that the cost function $c : \mathbb{R}_+ \to \mathbb{R}_+$ is increasing and convex in ℓ, and that $c(\ell) \to \infty$ as $\ell \to \infty$. We can also allow c to have finitely many flat segments, and do not assume that c is continuously differentiable. The possibility of kinks in c may be needed to rationalize constant segments on Λ. Allowing for flat segments in c implies that the minimizer ℓ^* may be an interval

for some x, which we can represent with a discontinuity in Λ at that value of x. We can now state a result that echoes the one in Proposition 1:

PROPOSITION 2. Fix a discount rate, curvature, variance, and a value of the threshold (r, B, σ^2, X), all positive. Let $\Lambda(\cdot) : (-X, X) \to \mathbb{R}_+$. Assume that $\Lambda(\cdot)$ is symmetric around zero, $\Lambda(x) = \Lambda(-x)$, increasing in $|x|$, and differentiable on $(0, X)$. Then, there exists an increasing convex cost function $c(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ that uniquely rationalizes the postulated decision rule as in equation (15). Moreover, the marginal cost $c'(\cdot)$ can be constructed by solving a second order linear ordinary differential equation.

The proof of the statement follows the same logic used in the proof of Proposition 1. Appendix D provides more details on the solution of this model. Note that observation on the frequency and size of price changes cannot in general distinguish between the random menu cost model of Section 2.1 and the optimal intensity of price adjustment of this section. In this sense, the generalized hazard function Λ is a more fundamental object. Furthermore, as explained above, the model of this section allows a slightly larger set of generalized hazard functions Λ.

3 Steady State observable statistics

In this section we show how to use the hazard function Λ to derive several observable statistics produced by our model in the steady state. In particular, we solve for the implied invariant distribution of price gaps, with density $f(x)$, the number of price changes per unit of time, N_a, and the distribution of price changes, with density $q(\Delta p)$. We focus on two moments of this distribution, the variance and the Kurtosis, denoted $Var(\Delta p)$ and $Kurt(\Delta p)$. The setup allows for $\Psi \in \bar{\mathbb{R}}_+ \equiv \mathbb{R}_+ \cup \{\infty\}$. If Ψ is finite then the inaction range is bounded, $X < \infty$. Otherwise, the support is unbounded, $X = \infty$. Both cases are encompassed by the analysis of this section. The starting point of this section is the function Λ that summarizes the firm's optimal decisions:

ASSUMPTION 1. Let $\Lambda : (-X, X) \to \mathbb{R}_+$, be non-negative, piece-wise continuous, symmetric, i.e. $\Lambda(x) = \Lambda(-x)$ for all x, with at most finitely many discontinuities $x_k \geq 0$, and let $\mathbb{J} \equiv \{x_k\}$. If $X = \infty$, we assume that there is a $\lambda > 0$ and $0 < x_H < \infty$ such that $\Lambda(x) \geq \lambda$ for all $|x| > x_H$.

Note that if Λ is the solution to the firm problem studied in Section 2, then $\Lambda(x)$ must be weakly increasing for $x > 0$, although Assumption 1 does not impose that.

Next we define the invariant distribution of price gaps, with density $f(\cdot) : (-X, X) \to \mathbb{R}_+$. Importantly, f must be continuous everywhere, continuously differentiable at $|x| \in (0, X)$, twice continuously differentiable at all $|x| \in (0, X)/\mathbb{J}$, and symmetric around $x = 0$. Given the symmetry, we only define f on positive real values. It solves the following equations:

$$f(x)\Lambda(x) = \frac{\sigma^2}{2} f''(x) \text{ for all } x \in [0, X), \ x \neq 0 \text{ and } x \notin \mathbb{J} \tag{16}$$

$$\text{with boundary conditions: } \frac{1}{2} = \int_0^X f(x)dx \quad \text{and} \quad \lim_{x \to X} f(x) = 0 \ . \tag{17}$$

Note that if $\Psi < \infty$, then $f(X) = 0$ is an implication of X being an exit point, i.e. a barrier. Otherwise it is a requirement for integrability. Figure 1 plots three examples of the invariant distribution of price gaps which solves equation (16)-(17) for a hazard function with power form $\Lambda(x) = \kappa \left(\frac{x}{X}\right)^{\nu}$. The quadratic case, $\nu = 2$, has been considered for instance by Caballero and Engel (1993a); Berger and Vavra (2018).

Figure 1: Density function $f(x)$ for the invariant distribution of gaps

Power hazard function: $\Lambda(x) = \kappa \left(\frac{x}{X}\right)^{\nu}$, $X < \infty$

14
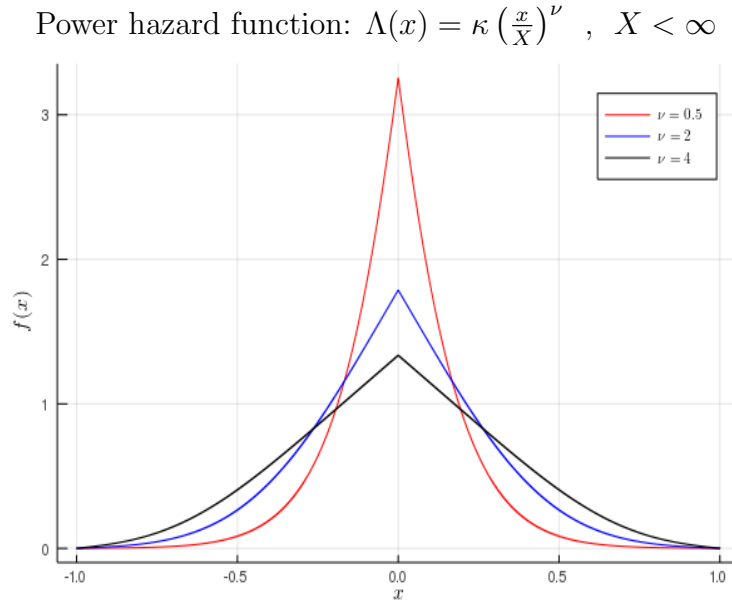
Frequency of price changes N_a. There are two types of price changes: those that occur when x reaches X, if it is finite, and those that occur when the firm draws a low enough fixed cost. Since X is an exit point, the number of price changes of the first type is given by $-2\frac{\sigma^2}{2}f'(X)$. The sign is negative because $f'(X)$ is negative. The 2 in front is becaus the same number of price changes happens when x reaches X as when x reaches $-X$. Note that if $X = \infty$ then $f'(X) = 0$. The second type of price changes occurs when $|x| < X$, which happens with density $f(x)$, and draws a sufficiently low fixed cost, which happens with probability $\Lambda(x)$ per unit of time. This gives

$$N_a = 2\left[\int_0^X f(x)\Lambda(x)dx - \frac{\sigma^2}{2}f'(X)\right] \ . \tag{18}$$

We remark for future reference that, as shown in Alvarez, Le Bihan, and Lippi (2016) for a very wide class of models that includes the ones in this paper, the following relation holds for any feasible policy in this class of menu cost problems:[12]

$$N_a \, Var\left(\Delta p\right) = \sigma^2 \tag{19}$$

This equation will be useful later in applications. We will use s for the fraction of price changes that occur before hitting the boundary $\pm X$. We can use equation (18) to replace the Kolmogorov forward equation for f, and integrate by parts to obtain that:

$$s \equiv \frac{\int_{-X}^X \Lambda(x)f(x)dx}{N_a} = 1 - \frac{\sigma^2|f'(X)|}{N_a} = 1 - \frac{|f'(X)|}{|f'(0)|} \text{ since } N_a = \sigma^2|f'(0)| \tag{20}$$

where $|f'(0)|$, with a slight abuse of notation, is the absolute value of either the right or left derivative of $f(x)$ evaluated at $x = 0$.

Distribution of price changes. Recall that upon any price change the firm "closes" its gap x, i.e. the size of the adjustment is $\Delta p = -x$. If $X < \infty$ then the distribution of price changes has a mass point at $\Delta p = -X$. The mass of such price changes is equal to $\frac{\sigma^2}{2}|f'(X)|$. There are also

---

[12]The key assumption for this result to hold is that the price gap is closed upon adjustment. This assumption is not true in e.g. models with high inflation or models with price plans, see Alvarez and Lippi (2020).

price changes of size $|\Delta p| < X$ that occur when a firm has $x < X$ and draws a sufficiently low fixed cost. This occurs with probability $\Lambda(x)$ per unit of time for a firm with price gap x. Recall also that at steady state, there is a density $f(x)$ of firms with price gap x. This density is symmetric around zero. The distribution of price changes is thus symmetric around zero as well. It has the following form:

$$\Delta p = \begin{cases} -x & \text{w/ density } q(-x) \equiv \frac{\Lambda(x)f(x)}{N_a} \text{ for } x \in (0, X) \\[2ex] -X & \text{w/ probability } \frac{\frac{\sigma^2}{2}|f'(X)|}{N_a} \end{cases} \tag{21}$$

Note that $1-s$, as defined in equation (20), is also twice the size of the mass point at the boundary of the support of this distribution.

Figure 2: Density function $q(\Delta p)$ of the distribution of price changes

Quadratic Hazard function: $\Lambda(x) = \kappa x^2$, $X = \infty$, shape parameter: $\eta \equiv \left(\frac{2\kappa}{\sigma^2}\right)^{\frac{1}{4}}$
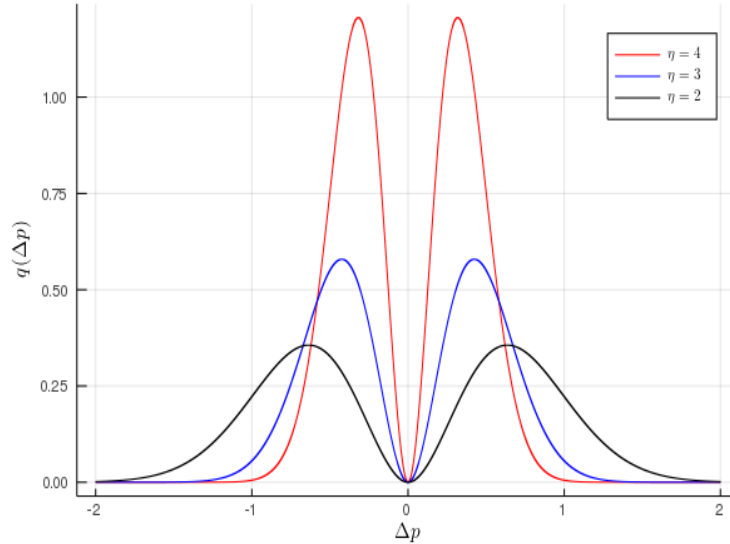
Figure 2 plots a few examples of the density of price changes implied by a quadratic hazard function $\Lambda(x) = \kappa x^2$ with an unbounded support $X = \infty$. This uses the definition $q(-x) \equiv \frac{\Lambda(x)f(x)}{N_a}$ from equation (21), where the density of price gaps f solves the Kolmogorov forward equation equation (16). The generalized hazard function and frequency of price changes alone are sufficient to construct both f and q.

We note that in the quadratic case the distribution of price changes is indexed by a single

parameter $\eta \equiv \left(\frac{2\kappa}{\sigma^2}\right)^{\frac{1}{4}}$ determining its shape, and features no mass points at the boundary of the inaction region since $X = \infty$. This means $s = 1$ in terms of equation (20). The parameter η is recurring in the class of generalized hazard functions of the power form, and generally determines the shape of the distributions of price changes.

For future reference we define two useful moments. The variance and the Kurtosis of the price changes $Kurt(\Delta p)$ can be defined using the distribution in equation (21):

$$Var(\Delta p) = \frac{2\left[\int_0^X x^2 \Lambda(x) f(x) dx - X^2 \frac{\sigma^2}{2} f'(X)\right]}{N_a} \tag{22}$$

$$Kurt(\Delta p) = \frac{2\left[\int_0^X x^4 \Lambda(x) f(x) dx - X^4 \frac{\sigma^2}{2} f'(X)\right]}{N_a} \frac{1}{[Var(\Delta p)]^2} \tag{23}$$

Standardization. It is useful to rescale the firm's decision rule to isolate the role of the shape of Λ and of other parameters. Standardization clarifies which objects matter conceptually, and also helps to bring the model to the data, as shown in Section 4. Let's start with a price-setting problem represented by the triplet $\{X, \Lambda, \sigma^2\}$ with $\sigma^2 > 0$ and $\Lambda : (-X, X) \to \mathbb{R}_+$ satisfying Assumption 1. Given the triplet $\{X, \Lambda, \sigma^2\}$, we can compute the corresponding density of price changes $q(\cdot) : (-X, X) \to \mathbb{R}$, the variance of price changes $Var(\Delta p)$, the frequency of price changes N_a, and the share of price changes away from the boundaries s. We have the following result:

PROPOSITION 3. Consider an economy characterized by $\{X, \Lambda, \sigma^2\}$, and associated $q, Var(\Delta p), N_a$ and s. For any $b > 0$ define another economy $\{\tilde{X}, \tilde{\Lambda}, \tilde{\sigma}^2,\}$ where $\tilde{X} = b X$, $\tilde{\Lambda}(z) = \Lambda(z/b)$ for all $z \in (-\tilde{X}, \tilde{X})$, and $\tilde{\sigma} = b \sigma$. These economies feature: (i) the same frequency of price changes, $\tilde{N}_a = N_a$, (ii) the same fraction of price changes away from the boundaries, $\tilde{s} = s$, and (iii) the same shape of the density of price changes, namely: $\tilde{q}(z) = q(z/b)/b$ for all $z \in (-X b, X b)$.

Note that we can choose $b^2 = 1/Var(\Delta p)$, for instance, so that the variance of price changes in the rescaled economy is one, $\tilde{V}ar(\Delta p) = 1$. This new economy can then be referred to as "standardized". The proposition shows that $Kurt(\Delta p)$ and the share s only depend on the shape of Λ, described by $\hat{\Lambda}$. In general, the shape cannot be summarized by a finite number of parameters, but in some situations a single parameter will suffice. For instance, below we consider a case where

Λ is a power function and $\hat{\Lambda}$ is described by a single parameter.

In addition to the standardization described above, we can also consider transformations akin to changing the time units, thus only affecting the frequency of adjustment, but not the distribution of price changes. In particular, consider a scalar $k > 0$, and define $\hat{\Lambda}(x) = \Lambda(x)/k$ for all x, $\hat{\sigma}^2 = \sigma^2/k$ and $\hat{X} = X$. It easy to see that $\{\hat{X}, \hat{\Lambda}, \hat{\sigma}^2\}$ has $\hat{N}_a = kN_a$ and $\hat{Q}(x) = Q(x)$ for all x.[13]

A useful approximation. We conclude with a proposition showing that for the case in which $\Psi < \infty$, so that $X < \infty$, the invariant distribution can be accurately approximated by one corresponding to a generalized hazard function Λ with unbounded support and arbitrarily large values for $x > X$. This approximation is useful because the case with unbounded support is somewhat simpler to analyze, since it does not involve discussing the mass points at the boundary of the inaction region.

PROPOSITION 4. Let $X < \infty$ and let $\Lambda : [0, X) \to \mathbb{R}_+$ be a continuous generalized hazard function, where $f : [0, X] \to \mathbb{R}_+$ is its corresponding invariant density, assumed to be symmetric. Let $\Lambda_k : [0, \infty) \to \mathbb{R}_+$ be defined as $\Lambda_k(x) = \Lambda(x)$ if $x < X$ and $\Lambda_k(x) = k$ otherwise. Let also $f_k : [0, \infty) \to \mathbb{R}_+$ be the invariant density associated with Λ_k, also assumed symmetric for negative $x's$. Then f_k converges uniformly to f in $[0, X]$ as $k \to \infty$.

4   From Price Changes to Price Gaps and Hazards

In this section we show how to recover the invariant density of price gaps f and the adjustment hazard Λ from the observable distribution of price changes. These two objects then allow us to recover the underlying distribution G of adjustment costs ψ in a random menu cost model of Section 2.1. We apply the algorithm to data taken from Cavallo (2015), fitting the distribution of price changes Q and recovering f, Λ, and G. For future reference, we pay particular attention to estimate the kurtosis of the distribution of price changes.

To do this, we first characterize the restrictions that an increasing hazard function Λ imposes on Q and establish a mapping from the observables (price changes) to the distribution of price

---

[13]In Appendix H we consider an alternative normalization, suitable for the case where $X < \infty$.

gaps f and adjustment hazard Λ. We then propose a non-parametric identification strategy to identify the distribution of price changes which takes into account unobserved heterogeneity across different products consistent with the theory as described in Proposition 3 . We illustrate these results fitting a flexible functional form to the distributions Q for several product categories in the dataset from Cavallo (2015). Interestingly, using this data set which have no time aggregation, arguably minimum measurement error, and accounting for unobserved heterogeneity, we find distributions with much smaller Kurtosis than in the literature. From this estimated distribution, we then recover f and Λ, and from them obtain the distribution of random menu cost G using the characterization in Proposition 1.

Identification of f and Λ. We start with a lemma that describes the properties of the distribution of price changes generated by a generalized hazard function. It only requires Assumption 1:

PROPOSITION 5. Let Q be the CDF of price changes corresponding to a generalized hazard function Λ satisfying Assumption 1. Then, Q is absolutely continuous on $(-X, X)$, so that $Q(x) = Q(-X) + \int_{-X}^{x} q(s)ds$ for $x < X$. The density $q(\cdot) : (-X, X) \mapsto \mathbb{R}_+$ is symmetric around zero, $q(x) = q(-x)$, and continuous at $x \notin \mathbb{J}$. Q has mass points if and only if $X < \infty$, in which case they are at $-X$ and X, and is fully identified by the collection of all its moments.

The next proposition, which is one of the main results of the paper, obtains the density f of price gaps from the distribution of price changes. The idea is simple: we integrate the Kolmogorov forward equation twice and replace σ^2 as in equation (19). Once we have f, it is straightforward to get Λ using $f(x)\Lambda(x) = q(x)N_a$.

PROPOSITION 6. Let Δp be price changes, and let Q and q be the CDF and corresponding density of price changes corresponding to a generalized hazard function Λ satisfying Assumption 1. Let N_a be the frequency of price changes. The density for the invariant distribution $f(x)$ is given by

$$f(x) = \frac{2}{Var(\Delta p)} \left[\int_x^X (1 - Q(z))\, dz \right] \quad \text{for all } x \in (0, X) \tag{24}$$

and $f(-x) = f(x)$, where $Var(\Delta p)$ is the variance of the price changes computed using Q. The

generalized adjustment hazard $\Lambda(x)$ is given by

$$\Lambda(x) = \frac{N_a \, Var(\Delta p)}{2} \frac{q(x)}{\int_x^X (1 - Q(z)) \, dz} \quad \text{for all } x \in [0, X) \tag{25}$$

and $\Lambda(-x) = \Lambda(x)$.

Recall that the function Λ implied by the models of Section 2.1 and Section 2.2 is increasing in $x \in (0, X)$. If Λ is increasing in $(0, X)$, the right hand side of equation (25) must be increasing. At any x where Λ is differentiable,

$$\frac{\Lambda'(x)}{\Lambda(x)} = \frac{q'(x)}{q(x)} + \frac{1 - Q(x)}{\int_x^X (1 - Q(z)) \, dz} \geq 0 \quad \text{for all } x \in (0, X), x \notin \mathbb{J} \tag{26}$$

The model of Section 2.1 also implies that $\Lambda(\cdot)$ is bounded above on $(0, X)$. If this is the case, the right hand side of equation (25) must be bounded. If Λ is increasing, this is equivalent to

$$\lim_{x \to X} \frac{q(x)}{\int_x^X (1 - Q(z)) \, dz} = \lim_{x \to X} \frac{q'(x)}{-(1 - Q(x))} \leq C \tag{27}$$

for some constant C. Moreover, if $X < \infty$, then $\lim_{x \to X}(1 - Q(x)) > 0$ and hence $\lim_{x \to X} q'(x)$ must be finite. If $X = \infty$, then $\lim_{x \to X} \frac{q''(x)}{q(x)} \leq C$ by L'Hopital rule. Note that if q has exponential tails, equation (27) is satisfied even if $X = \infty$. Moreover, since the model of Section 2.2 does *not* imply a bounded Λ it does not require equation (27).

A simple measure of state dependence. The expression for Λ in Proposition 6 evaluated at $x = 0$ can be used to measure a simple index of the lack of state dependence in pricing. We label it as $\mathcal{C}$, for "Calvo-ness":

$$\mathcal{C} \equiv \frac{\Lambda(0)}{N_a} \tag{28}$$

The index $\mathcal{C}$ measures the fraction of price changes that happen independently of the price gap x. In terms of the random menu cost model, it measures the fraction of price changes with no adjustment cost paid. Alvarez, Le Bihan, and Lippi (2016) use the same statistic to index multi-

product version of the Calvo$^+$ model. In that special case, the function Λ is constant, equal to $\Lambda(0)$ for all $|x| < X$, and then it jumps to infinity, i.e. there is a barrier in this case $\mathcal{X}$ is simply equal to the fraction of price changes that do not occur at the barriers $\pm X$. Clearly, the setup here is much more general, and the definition captures all the price changes that are unrelated to the value of the price gap. Hence, $\mathcal{C}$ is a broad measure of lack of state dependence. The next corollary of Proposition 6 shows that $\mathcal{C}$ can be measured using data on the distribution of price changes.

COROLLARY 1. The fraction of price changes independent of the price gap $\mathcal{C}$ defined in equation (28), is given by

$$\mathcal{C} = \frac{Var(\Delta p)}{2\,\mathbb{E}[|\Delta p|]}\, q(0)\,. \tag{29}$$

Moreover, using equation (18) for N_a, $\mathcal{C} \leq 1$ if Λ is increasing.

The expression in Corollary 1 is intuitive: the fraction of price changes independent of the state is proportional to the density of price changes at zero, a magnitude that can be estimated. The constant of proportionality is a ratio of two easily measurable statistics. The importance of Corollary 1 is that the right hand side of equation (29) involves three observable quantities which depend exclusively on the distribution of price changes: the density at zero, $q(0)$, and two of its moments: the variance $Var(\Delta p)$ and the expected absolute value $\mathbb{E}[|\Delta p|]$.

Unobserved Heterogeneity. Armed with Proposition 6 we can recover f, Λ, and the model primitives, like the distribution of menu cost, using Proposition 1. As an intermediate step, we discuss how to account for a simple, yet pervasive, form of unobserved heterogeneity in the estimation of q. We assume that products in a narrowly defined category have the same distribution of price changes up to an (unobserved) shift in the size, i.e. the distributions have the same shape but different scale. Proposition 3 discusses exactly this type of transformation that changes the scale without affecting shape. The reason we want to account for this form of unobserved heterogeneity is that, as is well known, a mixture of distributions with identical kurtosis but different variances has itself a higher kurtosis.[14] The setup is similar to a random effect model,

---

[14]See Appendix G for the formal treatment of this result.

yet without assuming any functional form for the distributions. In particular we use a variation of Kotlarski (1967)'s lemma. The products (within a category) are indexed by i, and t is the chronological number of adjustment. Let I be the set of all products and $T(i)$ be the set of adjustment instances for a product $i \in I$. We use the following specification:

$$\Delta p_{it} = b_i \Delta \tilde{p}_t \text{ for } i \in I \text{ and } t \in T(i) \tag{30}$$

Here b_i corresponds to the scaling factor b in Proposition 3. The six identification assumtpions are

1. $\#T(i) > 1$, so there are at least two price changes for each i

2. $\Delta \tilde{p}_t$ are drawn from a distribution Q, described by Proposition 5, for all $t \in \bigcup_{i \in I}$

3. $\Delta \tilde{p}_t$ and $\Delta \tilde{p}_s$ are statistically independent for all $t, s \in \bigcup_{i \in I} T(i)$

4. $b_i \geq 0$ are drawn from a distribution H for all $i \in I$

5. $\Delta \tilde{p}_t$ and b_i are statistically independent for all $i \in I$ and $t \in \bigcup_{i \in I} T(i)$

6. $\mathbb{E}[(\Delta \tilde{p}_t)^2] = 1$ for all $t \in \bigcup_{i \in I} T(i)$

That the distribution Q is described by Proposition 5 means, in particular, that it is symmetric around zero. The last assumption is a normalization, since the variances of H and Q are not identified together. We can show the following result:

PROPOSITION 7.   Consider two pairs of integer numbers (j, k) and (j', k') such that $j + k = j' + k'$. Under the assumptions stated above we have:

$$\frac{\mathbb{E}[(\Delta \tilde{p}_t)^j]\mathbb{E}[(\Delta \tilde{p}_t)^k]}{\mathbb{E}[(\Delta \tilde{p}_t)^{j'}]\mathbb{E}[(\Delta \tilde{p}_t)^{k'}]} = \frac{\mathbb{E}[(\Delta p_{it})^j(\Delta p_{is})^k]}{\mathbb{E}[(\Delta p_{it})^{j'}(\Delta p_{is})^{k'}]} \tag{31}$$

for any (t, s) with $t \neq s$.

This proposition has two important implications. First, we can establish a recursive expression for the even moments of the distribution of $\Delta \tilde{p}_t$:

$$\mathbb{E}[(\Delta \tilde{p}_t)^{2k+2}] = \mathbb{E}[(\Delta \tilde{p}_t)^{2k}] \cdot \frac{\mathbb{E}[(\Delta p_{it})^{2k+2}]}{\mathbb{E}[(\Delta p_{it})^{2k}(\Delta p_{is})^2]} \text{ for all } k \geq 0 \tag{32}$$

which only uses equation (31) and the normalization assumed above that $\mathbb{E}[(\Delta\tilde{p}_t)^2] = 1$. Starting from the normalized second moment, we can construct all even moments recursively using equation (32), thus obtaining a *non-parametric identification* of the density q.[15] Second, for future reference we display an expression for the Kurtosis of $\Delta\tilde{p}$:

$$Kurt(\Delta\tilde{p}_t) = \frac{\mathbb{E}[(\Delta p_{it})^4]}{\mathbb{E}[(\Delta p_{it})^2(\Delta p_{is})^2]} = \frac{Kurt(\Delta p_{it})}{1 + \mathrm{corr}(\Delta p_{it}^2, \Delta p_{is}^2)CV(\Delta p_{it}^2)CV(\Delta p_{is}^2)} \text{ for } t \neq s \quad (33)$$

The first equality is how we estimate kurtosis, correcting for this unobserved heterogeneity. The second equality shows how our method to measure kurtosis amounts to a correction of the kurtosis computed by pooling different goods without accounting for heterogeneity. Whenever the squares of price changes of individual products are positively correlated, as we have systematically found in the data, the correction leads to a substantial downward adjustment of the estimated kurtosis.

Data and estimation. We use the open access data from Billion Prices Project presented by Cavallo (2015).[16] We have chosen scraped price data to reduce the measurement error present in other data sets for example due to time aggregation using average revenue. It is important to avoid this form of measurement error to accurately estimate the kurtosis of the distribution of price changes, one of the goals of this section. The time span of our sample is between May 2008 and June 2010. From daily data on prices we construct the series of spells together with the size of the price change at the end of each spell. We trim the sample at price changes larger than 150 log points size in absolute value.[17] To fit a symmetric density, for each value Δp in the sample we use points in the band around it and around $-\Delta p$ as well. The left panel of Figure 9 plots the histogram of price changes for a narrowly defined product category. The right panel presents the symmetrized histogram with a fitted density. This fitted density is *not* the underlying density q, since it is confounded by the unobserved values of b_i.

Table 1 presents summary statistics on the seven categories we use, as well as the estimated kurtosis. The kurtosis is estimated in two ways: first by (incorrectly, according to our assumptions) pooling different products in the same category (p), and second by accounting for product

---

[15]This fully characterizes the distribution of $\Delta\tilde{p}_t$, since its odd moments are equal to zero due to symmetry.

[16]Link: http://www.thebillionpricesproject.com/datasets/. We use the US store number 1.

[17]We remove 87 (larger than 150 log points) out of 326,570 price changes for products with at least three spells.

Figure 3: Distribution of price changes in a narrow category

Histogram for Δp	Symmetrized histogram with a fitted density
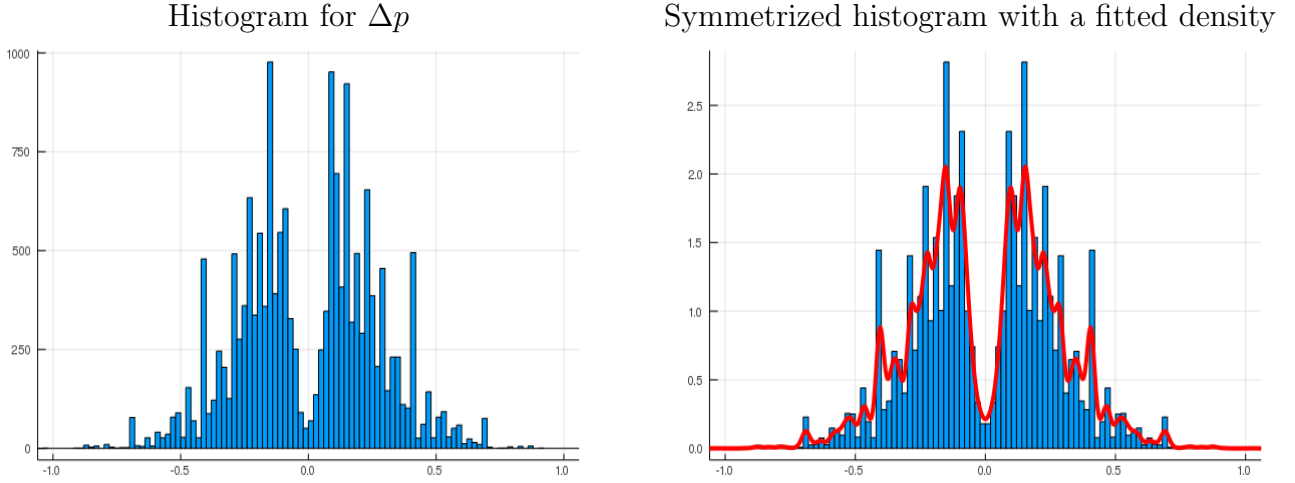
Pooling all products for category 561, COICOP label "Non-durable household goods"

heterogeneity (u). To implement the latter procedure, we use equation (33). This equation is a particular case of equation (31) with $j = 4$, $k = 0$, and $j' = k' = 2$. Importantly, the expectation is taken over $t \neq s$, and any such pair (t, s) can be taken to estimate it. Our estimator is constructed as follows: for any pair (j, k), we estimate $\mathbb{E}[|\Delta p_{it}|^j |\Delta p_{is}|^k]$ by

$$\frac{1}{\#I} \sum_{i \in I} \frac{1}{\#T(i)(\#T(i) - 1)} \sum_{t,s \in T(i), t \neq s} |\Delta p_{it}|^j |\Delta p_{is}|^k \tag{34}$$

where $\#$ denotes the number of elements in the set. Note that this estimator includes all available price changes $t, s \in T(i)$ for every product, maximizing the use of the data. Individual products in the sample have around 20 price changes each. Distributions are centered around zero, with the mean being around one hundredth of the standard deviation. It is evident from the table that properly accounting for heterogeneity reduces the estimated kurtosis to about half. This points to a substantial correlation in absolute values of the consecutive squared price changes. In Appendix B we tabulate implied correlations recovered from equation (33). These turn out to be in the range between 0.29 and 0.45.

To estimate Q we use a Gamma distribution. In principle, the distribution could be estimated non-parametrically, since every moment is identified. In practice, this would require estimating a

Table 1: Summary statistics and kurtosis estimates

| Category	Number Products	Number P. changes	$\hat{\mathbb{E}}(\Delta p_{it})$	$\hat{\sigma}(\Delta p_{it})$	Kurtosis Pooled	Kurtosis w/Unobs. Heterog.	$\mathcal{C}_{pooled}$	$\mathcal{C}$ w/Unobs. Heterog.
111	3437	74464	0.002	0.341	3.418 (0.162)	1.656 (0.071)	0.077	0.071
119	3225	56527	0.002	0.328	3.831 (0.092)	1.955 (0.050)	0.085	0.064
1212	2551	30343	-0.001	0.245	3.524 (0.272)	2.052 (0.162)	0.040	0.039
122	1401	27321	0.002	0.342	2.956 (0.089)	1.677 (0.051)	0.118	0.091
118	1388	30111	0.003	0.308	3.624 (0.240)	2.044 (0.118)	0.080	0.078
117	1154	20995	0.007	0.309	3.487 (0.135)	1.989 (0.047)	0.071	0.058
561	1032	17724	0.002	0.260	3.324 (0.221)	1.778 (0.133)	0.034	0.030

Categories legend: 111 "bread and cereals", 119 "other food products", 1212 "electric appliances for personal care", 122 "soft drinks", 118 "sugar, honey, and confectionary", 117 "vegetables", 561 "non-durable household goods"

large number of moments, substantially decreasing precision. Instead, we estimate the Kurtosis and use the unit variance restriction to fit the scale and size. The fitted density q is presented on the left panel of Figure 4 together with the underlying generalized hazard function Λ and the density of price gaps f. In the Appendix B, we detail our algorithm and show extended results from fitting a mixture of two Gamma distributions, for which we estimate moments using Proposition 7 to fit five parameters: scale and size of the distributions and the weight.

Finally, the right panel contains the distribution of menu cost recovered from the resulting hazard Λ. The units on the horizontal axis correspond to the annual profit of the firm. There is no mass point at zero, since the recovered generalized hazard function has $\Lambda(0) = 0$. Note that the model with random menu cost can only rationalize a bounded generalized hazard function. Gamma distribution is convenient, since by Proposition 6 it implies a bounded Λ. In the Appendix B, we show the procedure to recover the cost function c corresponding to the model in Section 2.2.

This model allows for an unbounded Λ, so we use a power specification $\Lambda(x) = \kappa x^\nu$, deriving the moments of Q and using the analytical expressions to fit the parameters.

Figure 4: Estimated distribution of price changes and implied cost functions

Estimated $q(\cdot)$, recovered $f(\cdot)$ and $\Lambda(\cdot)$ Recovered CDF and density of menu costs
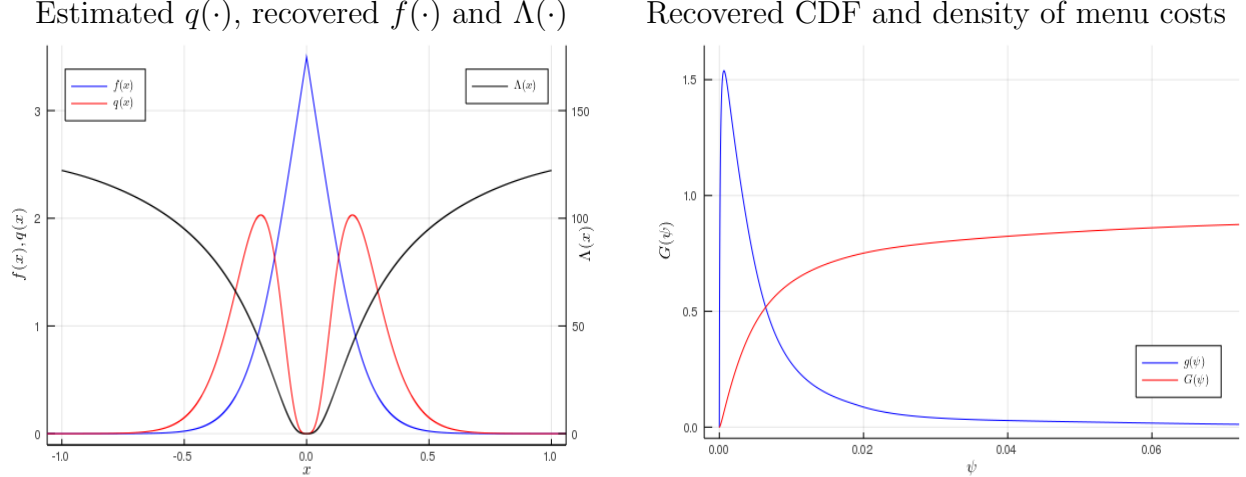
Estimating the degree of state dependence. We now turn to measuring $\mathcal{C}$. We do this in two ways. First, we ignore the unobserved heterogeneity and assume that the price data for the narrowest category of goods all come from the same primitives of the model. These primitives are the generalized hazard function Λ (including the value of X, the barrier) and σ^2. Recall from Proposition 6 that these objects fully describe the data-generating process, and this mapping is injective. In this exercise we just estimate all the objects on the right hand side of equation (29). The results are shown in the column labelled "$\mathcal{C}_{pooled}$" in Table 1. Their average across categories is 0.072, i.e. just above 7% of price changes are independent of the state. This small number is due to the small value of the density q at $\Delta p = 0$, which is apparent from the right hand side panel of Figure 9.

Second, we account for unobserved heterogeneity of the type described above. If there is heterogeneity of this type across products in the narrowest category, using the simple expression for $\mathcal{C}$ from equation (29) produces an upward bias in the estimate. We derive an unbiased estimator, the result being analogous to the one in Proposition 7. We express this estimator as a function of the pooled estimator $\mathcal{C}_{pooled}$ and a correction due to the unobserved heterogeneity:

PROPOSITION 8. Under the assumptions 1-5 stated above,

$$\mathcal{C} = \mathcal{C}_{pooled}\left(1 + \frac{Cov(b_i^{-1}, b_i^2)}{\mathbb{E}[b_i^{-1}]\mathbb{E}[b_i^2]}\right) < \mathcal{C}_{pooled} \tag{35}$$

where the two components are given by

$$\mathcal{C}_{pooled} = \frac{q(0)Var(\Delta p_{it})}{2\,\mathbb{E}[|\Delta p_{it}|]} \text{ and } 1 + \frac{Cov(b_i^{-1}, b_i^2)}{\mathbb{E}[b_i^{-1}]\mathbb{E}[b_i^2]} = \frac{\mathbb{E}[b_i]}{\mathbb{E}[b_i^{-1}]\mathbb{E}[b_i^2]} = \frac{\mathbb{E}[|\Delta p_{it}|^{-1}|\Delta p_{is}|^2]}{\mathbb{E}[|\Delta p_{it}|^{-1}]\mathbb{E}[|\Delta p_{it}|^2]} \text{ for } t \neq s$$

The estimate for $\mathcal{C}_{pooled}$ is obtained from the pooled data, and the correction for unobserved heterogeneity is measured using the (short) time dimension of the panel. The last column, labelled "$\mathcal{C}$ w/Unobs. Heterogeneity", in Table 1 contains the estimation results. Averaging across categories, the fraction of price changes independent of the state is 0.062 or just above 6%.

The correction multiplier is smaller than one because $1/b_i$ and b_i^2 are negatively correlated. Of course, if data for all products i in the narrowest category come from the same model primitives (Λ, X, σ^2), then there is no variation in b_i, and $\mathcal{C} = \mathcal{C}_{pooled}$.

5   Duration Analysis and Generalized Hazard Rate

In this section we consider the Survival and the Hazard Rate as functions of the *duration* of the price spells. Duration-based functions are often used in sticky price models. It is interesting to know whether the information encoded in them is different from that encoded in the size-distribution of price changes used above. We establish conditions for a non-trivial equivalence result: the distribution of durations and the variance of price changes together contain the same information about the fundamentals of the model as the distribution of price changes and frequency of adjustment. The distribution of spells with one statistic on the size of changes (the variance) is as informative as the size-distribution of changes and one temporal statistic (the frequency).

Denote by $S(t)$ the Survival function, the probability that a price spell lasts at least t units of time. We will show that, when $X = \infty$, an analytical Survival Function S uniquely identifies an

analytical Generalized Hazard Rate function Λ. When $X = \infty$, the Survival function is given by

$$S(t) = \mathbb{E}\left[e^{-\int_0^t \Lambda(x(s))ds} \,\big|\, x(0) = 0\right] \text{ for all } t \geq 0 \tag{36}$$

where the expectation is taken with respect to the paths of the drift-less Brownian motion x with variance per unit of time equal to σ^2. The value of $S(t)$ is the Feynman-Kac formula evaluated at $x = 0$. The hazard rate $h(t) = -S'(t)/S(t)$ measures the probability per unit of time of a price spell ending conditional on lasting at least t. For example, the Survival function and its associated hazard rate for the case of a quadratic generalized hazard rate $\Lambda(x) = \Lambda(0) + \kappa x^2$ are:

$$S(t) = \frac{e^{-t\Lambda(0)}}{\left(\cosh\left(t\sqrt{2\kappa\sigma^2}\right)\right)^{\frac{1}{2}}} \text{ and } h(t) = \Lambda(0) + \sqrt{\kappa\frac{\sigma^2}{2}}\tanh\left(t\sqrt{2\kappa\sigma^2}\right) \text{ for all } t \geq 0 \tag{37}$$

This was obtained by Kac in his seminal study of what we now know as the Kac formula. The next lemma gives the main technical result to establish the link between the Survival function, which can in principle be measured in the data, and the generalized hazard function $\Lambda(x)$.

LEMMA 2. Fix a value of $\sigma^2 > 0$, and assume that $X = \infty$. Assume that S is related to Λ by equation (36). The derivatives of the Survival function S a time $t = 0$ and the derivatives of Λ at $x = 0$ are related by the recursively generated functions $\{F_n\}$ as follows:

$$\left.\frac{\partial^n S(t)}{\partial t^n}\right|_{t=0} = F_n(0) \text{ and all } n = 1, 2, \ldots \text{ where } F_n(\cdot) \text{ are given by} \tag{38}$$

$$F_{n+1}(x) = \frac{\sigma^2}{2}\frac{\partial^2 F_n(x)}{\partial x^2} - \Lambda(x)F_n(x) \text{ for all } x \in \mathbb{R} \text{ and } n = 1, 2, \ldots \text{ and} \tag{39}$$

$$F_1(x) = -\Lambda(x) \text{ for all } x \in \mathbb{R} \tag{40}$$

Lemma 2 is the base of an algorithm to compute the derivatives of S at $t = 0$ given Λ and the derivatives of Λ at $x = 0$ given S. Using this lemma, we obtain the main result of this section:

PROPOSITION 9. Assume that $\sigma^2 > 0$, $X = \infty$, and Λ satisfies Assumption 1. Let S be the Survival function of Λ, as in equation (36). If the generalized hazard function Λ is analytical, then

the Survival function S uniquely identifies Λ. Likewise, if the Survival function S is analytical, then the generalized hazard function Λ uniquely identifies S.

As remarked before, Lemma 2 gives an algorithm to recursively compute an expansion of S based on the derivatives of Λ, or an expansion of Λ based on the derivatives of S. An implication of Lemma 2 and Proposition 9 is that the hazard rate and its first three derivatives at zero duration $(t = 0)$ are given by particularly simple expressions involving the level and first two even derivatives of the generalized hazard function evaluated at zero price gap, i.e. $x = 0$:

$$h(0) = \Lambda(0) \geq 0\,, \frac{\partial h(t)}{\partial t}\Big|_{t=0} = \frac{\sigma^2}{2}\frac{\partial^2 \Lambda(x)}{\partial x^2}\Big|_{x=0}\,, \frac{\partial^2 h(t)}{\partial t^2}\Big|_{t=0} = \left(\frac{\sigma^2}{2}\right)^2 \frac{\partial^4 \Lambda(x)}{\partial x^4}\Big|_{x=0}\,,$$

$$\text{and } \frac{\partial^3 h(t)}{\partial t^3}\Big|_{t=0} = \left(\frac{\sigma^2}{2}\right)^3 \frac{\partial^6 \Lambda(x)}{\partial x^6}\Big|_{x=0} - 4\left(\frac{\sigma^2}{2}\frac{\partial^2 \Lambda(x)}{\partial x^2}\Big|_{x=0}\right)^2$$

These formulas give a simple connection between the local behavior of Λ around $x = 0$ and h around $t = 0$. Note that if $\Lambda(x)$ is, in addition of being symmetric and differentiable in x, increasing in $|x|$ around $x = 0$, then $\Lambda''(0) > 0$, and hence the hazard rate as function of duration, $h(t)$, must be increasing in duration, at least for small durations t. Likewise, if $\Lambda(x)$ were decreasing in $|x|$ around $x = 0$, then $\Lambda''(0) < 0$ and hence $h(t)$ must be locally decreasing in duration.

Comparing with the case of Proposition 6, in this case we use much more restrictive conditions for Λ, and obtain a more cumbersome representation — an infinite expansion instead of a closed-form expression involving an integral. In spite of this Proposition 6 and Proposition 9 have the same flavor: they show that if Λ is analytical and $X = \infty$, then Λ can be fully identified either using the information contained in the Survival function, i.e duration on price changes, and σ^2, which can be recovered from N_a and the variance of price changes with equation (19). Of course, this also means that the information on the survival function and the size distribution of price changes can be used as an over-identifying test of the model.

Finally, we can also estimate $\mathcal{C} \equiv \Lambda(0)/N_a$, the fraction of price changes independent of the state, by using duration data. Given the results above, $\mathcal{C}$ can be estimated as $h(0)/N_a$. This can be an alternative to the estimates presented in Table 1 using the size distribution of price changes. As in Section 4, a correction of unobserved heterogeneity may be important.

6 A Sufficient Statistic for Monetary Shocks

This section characterizes the real output effect of monetary shocks using a simple summary statistic, the cumulative output generated by a once and for all monetary shock. This is the area under the output's impulse response function. It combines in a single value the persistence and the size of the output response. The key result we present is that for small monetary shocks, like the ones typically considered in the literature, the area is completely encoded by the kurtosis and the frequency of price changes. These two moments are thus sufficient to compare different models.

We also find that, among the models with non-decreasing adjustment hazards, the kurtosis of price changes is maximized in the Calvo model. As was established above, only a non-decreasing generalized hazard function can be rationalized by random menu costs. Calvo model is the limiting case with no randomness and no option to adjust, so it minimizes the amount of selection and hence maximizes the output response. To establish this, we develop a general result that compares kurtoses generated by two different hazard functions.

The contribution to the cumulative impulse response of a firm with price gap x is

$$m(x) = -\mathbb{E}\left[\int_0^\tau x(t)dt \,\middle|\, x(0) = x\right] \tag{41}$$

where τ is the stopping time defined as the first time when $x(t)$ hits $\pm X$ or a reduction in adjustment costs causes the firm to change price. This stopping time is stochastic, so the expectation accounts for both the diffusion of the firm's price gap and the possible event of adjustment that happens with a Poisson intensity $\Lambda(x(t))$. In words, $m(x)$ is the expected (cumulative) price gap of a firm that starts with a gap x. Notice that in the Calvo case, where $\Lambda(x) = \lambda$ is independent of x, we immediately obtain $m(x) = -x/\lambda$, where $1/\lambda$ is the expected duration of a price spell.

The definition above uses the steady state decision rule $\Lambda(x)$, thus ignoring the general equilibrium feedback effect of the shock on the firm's decision. In Proposition 7 of Alvarez and Lippi (2014) it is shown that, given a combination of the general equilibrium setup in Golosov and Lucas (2007) and the lack of the strategic complementarities, these general equilibrium effects are of second order. In addition, we use the fact that after the first price change the expected contribution

to output of each firm is zero, since positive and negative output contributions are equally likely, so $m(0) = 0$. This allows us to characterize the propagation of the monetary shocks without tracking the time evolution of the whole price gap distribution.

The expectation in the right hand side of equation (41) is with respect to the process for x, a jump-diffusion with jump intensity $\Lambda(x)$, diffusion variance σ^2, and zero drift. The function $m : [-X, X] \to \mathbb{R}$ is once continuously differentiable, antisymmetric around $x = 0$, and satisfies:

$$m(x)\Lambda(x) = -x + \frac{\sigma^2}{2}m''(x) \text{ for all } x \text{ at which } \Lambda \text{ is continuous} \tag{42}$$

$$0 = m(X) \text{ if } X < \infty \text{ and } \lim_{x \to \infty} \frac{|m(x)|}{x} \leq \frac{1}{\inf_y \Lambda(y)} \text{ if } X = \infty \quad . \tag{43}$$

Now we can define the cumulative impulse response to a monetary shock of size δ as

$$\mathcal{M}(\delta) = \int_{-X}^{X} m(x - \delta)f(x)dx \quad . \tag{44}$$

This is simply the aggregate contribution of the firms to the cumulative impulse response. The response of a firm with the price gap x before the shock is $m(x - \delta)$.

Let $\{X, \Lambda, \sigma^2\}$ characterize an economy, with its corresponding invariant density f and firm's contribution to CIR, m. Let $\{\tilde{X}, \tilde{\Lambda}, \tilde{\sigma}^2\}$ be the standardized economy, defined as in Proposition 3, that has its associated $\{\tilde{f}, \tilde{m}\}$ with $\tilde{m}$ defined as $\tilde{m}(z) = m(z/b)/b$ for $b^2 = 1/Var(\Delta p)$ and satisfying the corresponding ODE with the boundary conditions for $\tilde{\sigma}^2$ and $\tilde{X}$. Define $\tilde{\mathcal{M}}(\delta)$, the cumulative impulse response of output to a monetary shock for the standardized economy, as

$$\tilde{\mathcal{M}}(\delta) = \int_{-\tilde{X}}^{\tilde{X}-\delta} \tilde{m}(x) \, \tilde{f}(x + \delta) \, dx \tag{45}$$

The next proposition relates the CIR of output in an economy to the one of its standardized version by scaling the monetary shock with the steady-state standard deviation of price changes. In words, for small monetary shocks the dispersion of price changes is immaterial, although in general the size of monetary shocks should be measured relative to the steady-state dispersion of price changes.

PROPOSITION 10. Let $\mathcal{M}$ and $\tilde{\mathcal{M}}$ be the cumulative impulse responses of an economy $\{X, \Lambda, \sigma^2\}$ with $Std(\Delta p) = Var(\Delta p)^{1/2}$ and the corresponding standardized economy $\{\tilde{X}, \tilde{\Lambda}, \tilde{\sigma}^2\}$. Then

$$\mathcal{M}(\delta) = \tilde{\mathcal{M}}\left(\frac{\delta}{Std(\Delta p)}\right) Std(\Delta p) \tag{46}$$

and thus $\mathcal{M}'(0) = \tilde{\mathcal{M}}'(0)$.

The proof is immediate, using the properties of $\tilde{m}$ and $\tilde{f}$ established above, differentiating equation (45), and evaluating at $\delta = 0$. Summarizing, Proposition 10 says that for small monetary shocks, the steady state standard deviation of the price changes is not important. For large shocks it clearly is. For example, take the case $X < \infty$. For $\delta \geq 2\bar{X}$, we have $\mathcal{M}(\delta) = 0$, because the shock displaces all the firms far enough, and they adjust immediately. Since the standardized version has $\tilde{X} = X/Std(\Delta p)$, this shows the importance of the size of the shock for large values.[18] The marginal version of this cumulative impulse response is

$$\mathcal{M}'(0) = -\int_{-X}^{X} m'(x) f(x) dx \tag{47}$$

This term can be used for a linear approximation of $\mathcal{M}$ around zero. Our main result is that it can be expressed as a function of two sufficient statistics: $Kurt(\Delta p)$, the kurtosis of the steady state distribution of price changes, and N_a, frequency of price changes.

PROPOSITION 11. Let $\Lambda(x)$ be any function satisfying Assumption 1. Then the cumulative impulse response for a small monetary shock is given by the ratio of two steady state statistics:

$$\mathcal{M}(\delta) = \frac{Kurt(\Delta p)}{6 N_a} \delta + o(\delta^2) \tag{48}$$

The approximation is accurate up to second order terms, so the remainder is of order δ^3. This happens since $\mathcal{M}''(0)$ is zero, which follows from $\mathcal{M}$ being an antisymmetric function, because m

---

[18]A similar result was shown in Alvarez and Lippi (2014) for the case of multiproduct firms, which only overlap with the current set up for the Golosov and Lucas case — with one product per firm.

is antisymmetric and f is symmetric.

Our results from Proposition 1 and Proposition 17 show that only weakly increasing Λ can be rationalized by the solution of a firm problem subject to random menu costs. But Assumption 1 allows for a very large class of generalized hazard functions, including decreasing and non-monotone ones. Proposition 11 holds for such functions too. It makes no reference to the micro-foundations behind Λ and hence also applies to setups where firms's behaviour is not described by an increasing hazard. An example is the model in Woodford (2009), where firms conduct costly reviews and have imperfect recall and access to their state. Also Costain and Nakov (2011b) use generalized hazard functions, without linking them to random menu costs.

Aggregation across heterogenous firms. We briefly discuss how the above results can be applied to economies composed of heterogenous firms. Assume that there are S groups of firms with different parameters, each with an expenditure weight $e(s) > 0$, $N(s)$ price changes per unit of time, and a distribution of price changes with kurtosis $Kurt(s)$. In this case, after repeating the arguments above for each group and aggregating, we obtain that the area under the IRF of *aggregate* output for a small monetary shock δ is

$$\mathcal{M}(\delta) = \delta\mathcal{M}'(0) + o(\delta^2) = \frac{\delta}{6} \sum_{s \in S} \frac{e(s)}{N_a(s)} Kurt(s) + o(\delta^2) = \frac{\delta}{6} D \sum_{s \in S} d(s)\, Kurt(s) + o(\delta^2) \quad (49)$$

where D is the expenditure-weighted average duration of prices $D \equiv \sum_{s \in S} \frac{e(s)}{N_a(s)}$, and $d(s) \equiv \frac{e(s)}{N_a(s)D}$ are weights that take into account both relative expenditures and durations. When all groups have the same durations, then $d(s) = e(s)$ and $\mathcal{M}$ is proportional to the average of the kurtosis of the sectors. As explained in Section 4, and shown in Proposition 18 in Appendix G, this average is also different from the kurtosis of the pooled data. This applies even if all the groups have the same kurtosis.[19] However, if groups are heterogenous in duration (or expenditures), then the kurtoses of the groups with longer duration (or higher expenditures) receive a higher weight in the computation of $\mathcal{M}$. Suppose for instance that a fraction of firms have flexible prices (zero duration in our model, or infinitely many price changes per unit of time), as in Dotsey and Wolman (2020).

---

[19]The effect of heterogeneity in $N_a(\Delta p_i)$ on aggregation is well known for the Calvo model: D is different from the average of $N_a(\Delta p_i)$'s, see for example Carvalho (2006) and Nakamura and Steinsson (2010).

The above formula implies that the group of the flexible price firms are excluded (zero duration yields a zero weight), and that the cumulative impulse response (CIR) is computed on the mass of firms with sticky prices. Notice that this is different from computing the CIR as the ratio of the cross-sectional average kurtosis and the average frequency. Since the latter is diverging because of the firms with flexible prices, the CIR computed this way would be zero, while obviously it is not.

Kurtosis. The next proposition shows the properties of generalized hazard functions that determine the Kurtosis of price changes. We will concentrate on the case where we will hold the adjustment frequency constant. Recall that that fixing the frequency of price changes can be accomplished as fixing the units of time. This procedure allows us to isolate the effect of a change in Λ on selection from its effect on the frequency. Moreover, with the frequency fixed, the kurtosis of price changes directly maps into the approximate cumulative impulse response.

PROPOSITION 12. Fix N_a and consider two hazard functions $\Lambda_1(x)$ and $\Lambda_2(x)$ with the corresponding boundaries X_1 and X_2, where $0 < X_2 \leq X_1 \leq \infty$. Let $\Lambda_1(0) > \Lambda_2(0)$ and let the function $\Lambda_1(x) - \Lambda_2(x)$ change sign at most once. Then, $\Lambda_1(x)$ generates a higher kurtosis of price changes.

The condition that $\Lambda_1 - \Lambda_2$ changes sign only once means that it is positive at first and maybe negative for x far from zero. This is to say that Λ_1 generates more adjustment for smaller x, and Λ_2 generates more for larger ones. Selection is therefore more pronounced with Λ_2, and the kurtosis of price changes is lower. There are two interesting corollaries of this result. The first is that for a fixed X the highest kurtosis is attained by the constant generalized hazard function. This corresponds to the Calvo$^+$ case:

COROLLARY 2. Fix N_a, the number of adjustments per unit of time, and $X < \infty$. The function $\Lambda(x)$ that is constant on $(-X, X)$ maximizes the kurtosis of the price changes over all functions $\Lambda(x)$ that are weakly increasing on $(0, X)$ and satisfy Assumption 1.

Second, a constant hazard function in combination with the infinite boundary X maximizes the kurtosis of price changes over *all* weakly increasing hazards. This is the pure Calvo case:

COROLLARY 3. Fix N_a. The constant function $\Lambda(x) \equiv \lambda$ maximizes the kurtosis of the price changes over all weakly increasing functions $\Lambda(x)$ satisfying Assumption 1.

By Proposition 11, it also means a constant Λ maximizes $\mathcal{M}'(0)$ for a fixed N_a. This highlights the role of selection. A strictly increasing rate of adjustment Λ implies positive selection, so the firms with larger deviations are more likely to adjust. When Λ is flat, there is no selection, so the price changers are drawn randomly from the population. Shocks are accommodated more slowly in this case, because the adjustment frequency does not depend on how much a firm needs to adjust, so the response of price takes longer, and hence the response of output is larger.

Finally, Proposition 12 sheds some light on the relationship between the strength of state dependence and the magnitude of output response. As we noted before, one measure of the strength of state dependence is the index $\mathcal{C} \equiv \Lambda(0)/N_a$, the share of adjustment happening independently of the price gap. We can show that, holding constant the shape of Λ (captured by its curvature) and adjustment frequency, this index co-moves with the Kurtosis. Hence, a higher share of adjustment independent of x means a stronger output response for the same shape of the hazard.

Define the curvature of the function Λ as

$$k(x) = \frac{\Lambda''(x)x}{\Lambda'(x)} \tag{50}$$

To understand what it means for two functions to have the same curvature, take some arbitrary Λ and decompose it into two parts, the intercept and the rest: $\Lambda(x) = \Lambda(0) + (\Lambda(x) - \Lambda(0))$. Now consider two simple linear transformations of the two parts of the hazard:

$$\Lambda_1(x) = a_1\Lambda(0) + b_1(\Lambda(x) - \Lambda(0)), \quad \Lambda_2(x) = a_2\Lambda(0) + b_2(\Lambda(x) - \Lambda(0)) \tag{51}$$

The transformation scales the intercept and the rest with different numbers, changing the strength of state dependence but broadly preserving the shape (it is easy to see that both Λ_1 and Λ_2 have the same curvature as Λ). When Λ_1 and Λ_2 generate the same adjustment frequency, the one with a weaker state dependence (higher $\mathcal{C}$) corresponds to a higher Kurtosis.

COROLLARY 4. Consider two generalized hazard functions $\Lambda_1(x)$ and $\Lambda_2(x)$ with the same boundary $X \leq \infty$. Furthermore, assume that they have the same curvature k everywhere and the frequency of adjustment N_a. Then $Kurt_1(\Delta p) > Kurt_2(\Delta p)$ if and only if $\mathcal{C}_1 > \mathcal{C}_2$.

An immediate implication of the Corollary 4 is that for two economies with the same frequency of price changes and the same curvature of the generalized hazard function, the one with higher value of $\mathcal{C}$ has a higher cumulative impulse response after a monetary shock.

6.1 Illustration with a power hazard function

In this section we describe the case where the generalized hazard function is a power function with the power parameter ν. In particular, we let $\Lambda(x) = \kappa |x/X|^\nu$ on $(-X, X)$ for some $\nu \geq 0$. This functional form nests Calvo-plus models with $\nu = 0$ and quadratic generalized hazard functions with $\nu = 2$.

We use this example to illustrate how the parameters affecting the shape of Λ determine the Kurtosis of price changes. To do this, we first show that, with ν fixed, the Kurtosis of price changes varies one-to-one with the share of adjustments from strictly between the barriers, s. This highlights the role of selection: the output response is weaker when fewer firms reach the boundaries, because fewer firms are close to adjustment right before a monetary shock happens. Second, we show that for any s the Kurtosis of price changes decreases monotonically with the power ν, which governs the shape of Λ.

We now describe the invariant density f. Upon a renormalization, we can solve for a symmetric density $\hat{f}(z)$ defined by $\hat{f}(z) = Xf(zX)$. The function $\hat{f}$ satisfies

$$\rho z^\nu \hat{f}(z) = \hat{f}''(z) \text{ with } \hat{f}(1) = 0 \text{ and } \int_0^1 \hat{f}(z)dz = \frac{1}{2} \tag{52}$$

where $\rho \equiv 2\kappa X^2/\sigma^2$. The solution to equation (52) is given by

$$\hat{f}(z) = c_1 \sqrt{z} \, \mathcal{I}_{\frac{1}{\nu+2}}\left(\frac{2\sqrt{\rho}}{\nu+2} z^{\frac{\nu+2}{2}}\right) + c_2 \sqrt{z} \, \mathcal{K}_{\frac{1}{\nu+2}}\left(\frac{2\sqrt{\rho}}{\nu+2} z^{\frac{\nu+2}{2}}\right) \text{ for all } z \in [-1, 1] \tag{53}$$

This is a combination of modified Bessel functions of the first and second kind $\mathcal{I}_{\frac{1}{\nu+2}}$ and $\mathcal{K}_{\frac{1}{\nu+2}}$ of order $\frac{1}{\nu+2}$, where the constants c_1, c_2 are chosen to satisfy the two boundary conditions described in equation (52). The form $\hat{f}$ depends on the parameters (ρ, ν) that capture adjustment coming from random menu costs. Note that if $X = \infty$ then $c_1 = 0$.

From the previous result, we can see that if two models have the same (ν, ρ), then the distribution of price changes in one is a rescaling of that in the other. The dimensionless statistics such as the kurtosis, the fraction of adjustment strictly within the boundaries s, or equivalent the mass of Q on $\pm X$, are the same. We can summarize this result as follows:

PROPOSITION 13. Let $\Lambda(x) = \kappa |x/X|^\nu$. The Kurtosis of price changes, the share of adjustments strictly between the boundaries, and the frequency of price changes satisfy: $Kurt(\Delta p) = \hat{K}(\rho, \nu)$, $s = \hat{S}(\rho, \nu)$, and $N_a = \frac{\sigma^2}{X^2}\hat{N}(\rho, \nu)$ respectively, where these functions have no other parameters. For fixed ν, the function $\hat{S}(\cdot, \nu)$ is increasing in ρ, and $\hat{K}(\cdot, \nu)$ is decreasing in ρ.

Using this proposition we can fix s, say to $s = 1$, or $X = \infty$, then $Kurt(\Delta p)$ is just a function of ν only, displayed in Figure 5. Alternatively, fixing ν we have that $Kurt(\Delta p)$ is only a function of the fraction of price changes strictly between barriers, s as displayed in Figure 6 below.

Figure 5: Kurtosis of power hazard function as ν varies

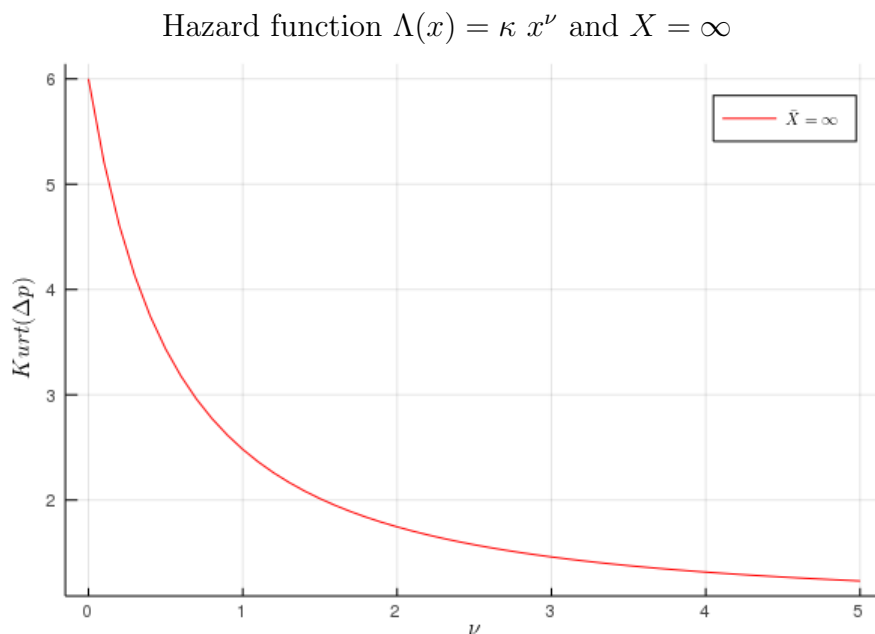Hazard function $\Lambda(x) = \kappa\ x^\nu$ and $X = \infty$

Figure 5 displays the value of $Kurt(\Delta p)$ for the case where $X = \infty$ as a function of ν. Note that Kurtosis goes from 6, corresponding to $\nu = 0$, or pure Calvo, to a value of 1, corresponding to $\nu \to \infty$ which approximates Golosov and Lucas. Increases in ν clearly change the shape of Λ, making it more convex, which is reflected in lower kurtosis of price changes. This illustrates how the shape of Λ determines the selection effect on price changes. Note that for $\nu = 2$, the quadratic case, $Kurt(\Delta p) \approx 1.75$.

Figure 6: Kurtosis behavior with a power hazard function as s varies

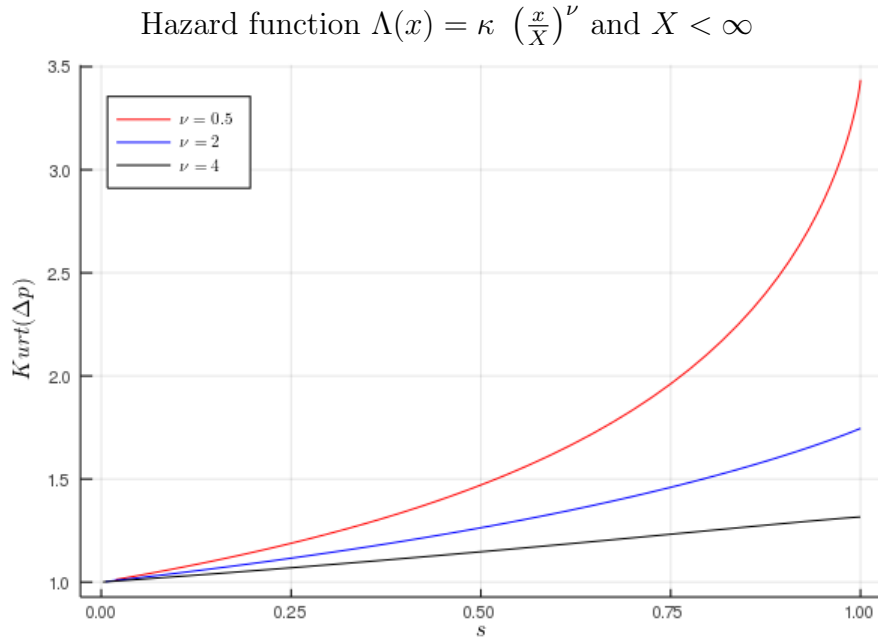Hazard function $\Lambda(x) = \kappa \left(\frac{x}{X}\right)^{\nu}$ and $X < \infty$

Figure 6 displays the value of $Kurt(\Delta p)$ for the case of $X < \infty$ as a function of s, the fraction of price changes strictly between the boundaries. We display such relationship for three values of ν. Note that fixing ν, as we change ρ, and obtain a larger share of price changes strictly inside the barriers s, which corresponds to a lower $Kurt(\Delta p)$. This illustrate the smaller selection effect of price changes when barriers are hit less often. Recall that $s = 0$ is equivalent to $X = \infty$. For each value of s, the difference lines shows that curvature ν corresponds to lower $Kurt(\Delta p)$.

7 Flexibility Index: scope and limitations

Caballero and Engel (2007) introduced the concept of the Flexibility Index ($\mathcal{F}$), subsequently used in several studies such as Berger and Vavra (2018), as an inverse measure of monetary nonneutrality. We show below that $\mathcal{F}$ measures the slope of the impulse response of prices right after a small once-and-for-all monetary shock. Below we define $\mathcal{F}$ in terms of the model, using Caballero and Engel's (2007) formula, and study the extent to which it is an accurate summary of the model's non-neutrality. We show that for models with barriers, where $X < \infty$, the flexibility index is always infinite. This prompts us to focus on the cases without barriers, $X = \infty$, where $\mathcal{F}$ is finite. In such cases we can compare $\mathcal{F}$ with the summary measure given by the cumulative impulse response defined in equation (44). We display non-pathological simple examples where the $\mathcal{F}$ is *not* an accurate summary of the effect on output, neither of its cumulative response, nor of its short term response.

The IRF of the aggregate price level after a shock δ can be written as

$$\mathcal{P}(t, \delta) = \Omega(\delta) + \int_0^t \omega(s, \delta) \, ds \tag{54}$$

where $\omega(s, \delta)$ is the flow contribution to the IRF at time $s > 0$, and $\Omega(\delta)$ is the time $t = 0$ jump in the price level. By definition $\frac{\partial}{\partial t}\mathcal{P}(t, \delta) = \omega(t, \delta)$. The flow value of the IRF of the aggregate price level at time $t > 0$ is given by

$$\omega(t, \delta) = -\int_{-X}^{X} x\Lambda(x)f(x,t)dx + X\sigma^2 \left[f'(-X,t) - f'(X,t) \right]$$

where $f(x, t)$ is the distribution of the price gaps among the firms that have not adjusted prices t units of time after the monetary shock. The first term is the change of prices across the distribution of price gaps at time t, with $f(x, t)$ solving the time dependent Kolmogorov Forward Equation:

$$\partial_t f(x,t) = -\Lambda(x)f(x,t) + \frac{\sigma^2}{2}\partial_{xx}f(x,t) \text{ for all } x \in [-X, X] \text{ and } t \geq 0, \tag{55}$$

$$f(X, t) = f(-X, t) = 0 \text{ for all } t > 0, \text{ and } f(x, 0) = f_0(x) \text{ for all } x \in [-X, X] \tag{56}$$

The initial jump is given by

$$\Omega(\delta) = \int_{-X}^{-X+\delta} (-x + \delta) \; f_0(x) \, dx \tag{57}$$

The initial distribution f_0 that we consider is a uniform shift by δ of some distribution $\hat{f}$:

ASSUMPTION 2.   The initial condition is $f_0(x) = \hat{f}(x + \delta)$, where $\hat{f}$ i) equals zero at the bounds, $0 = \hat{f}(-\bar{X}) = \hat{f}(\bar{X})$, ii) increases close to the lower bound, $0 < \hat{f}'(-\bar{X}) < \infty$, and iii) is differentiable on $(-\bar{X}, 0)$.

We write $f_0(x) = \hat{f}'(x)\delta + o(\delta)$ and consider the case of small δ. Note that the assumptions allow $\hat{f}$ to be the invariant distribution corresponding to $\{X, \Lambda, \sigma^2\}$, but they do not require it. In particular, $\hat{f}$ can be any distribution that has for any strictly positive time evolved according to equation (55) and equation (56). The Flexibility index is defined as $\mathcal{F} \equiv \frac{\partial}{\partial \delta}\omega(0, \delta)|_{\delta=0}$, which is equivalent to the definition in equation (17) in Caballero and Engel (2007).

PROPOSITION 14.   Let Ω and ω be the jump and flow values of the IRF of prices at $t = 0$. Let $X < \infty$, let Λ satisfy Assumption 1, and assume that the initial distribution f_0 satisfies Assumption 2. Then $\Omega(0) = \Omega'(\delta)|_{\delta=0} = 0$. Moreover, $\partial_\delta\omega(0, \delta)|_{\delta=0} = \infty$ and $\omega(0, 0) = 0$. Thus, if $X < \infty$, the flexibility index is infinite for any Λ.

Because of this result we will move to analyze the flexibility index for models with $X = \infty$, where it is finite. We will will do so for a family of hazard functions which is a slight generalization of the one treated in Section 6.1.

7.1   Power plus family of generalized hazard functions

We consider a simple four parameter family of models where $\Lambda(x) = \Lambda(0) + \kappa x^\nu$. We label this case as power-plus, because it adds a constant to the power case. Besides $\Lambda(0)$, κ, and ν, the other parameter of the model is σ^2. We introduce the parameter η and let α be the adjusted intercept:

$$\eta = \left(\frac{2\kappa}{\sigma^2}\right)^{\frac{1}{\nu+2}} \quad , \quad \alpha = \frac{\Lambda(0)\eta^\nu}{\kappa} \quad . \tag{58}$$

The quadratic case is $\nu = 2$ and $\alpha = 0$. This adjusted intercept measures the relative magnitude of $\Lambda(0)$ and the slope κ, increasing in the former and decreasing in the latter. We will show that for a fixed power the Kurtosis, adjustment frequency, and the flexibility index only depend on α.

PROPOSITION 15. Fix σ^2 and let $\Lambda(x)$ be a power-plus hazard function parameterized by $(\kappa, \Lambda(0), \nu)$. The adjustment frequency, the kurtosis of price changes, and the flexibility index are

$$N_a = \frac{\eta^2 \sigma^2}{2} \tilde{N}(\nu, \alpha) \tag{59}$$

$$\frac{Kurt(\Delta p)}{6 N_a} = \frac{1}{\eta^2 \sigma^2} \tilde{K}(\nu, \alpha) \tag{60}$$

$$\mathcal{F} = \frac{\eta^2 \sigma^2}{2} (\tilde{N}(\nu, \alpha)(1 + \nu) - \nu\alpha) \tag{61}$$

where $\tilde{N}(\nu, \alpha)$ and $\tilde{K}(\nu, \alpha)$ only depend on ν and α; $\tilde{N}(0, \alpha) \equiv 1 + \alpha$, and $\tilde{K}(0, \alpha) \equiv 2/(1 + \alpha)$.

With no intercept, the flexibility index and adjustment frequency are related by a simple formula via the elasticity of the hazard:

$$\mathcal{F} = N_a(1 + \nu) \tag{62}$$

If two models have the same (ν, α), the density of price changes in one is a rescaling of that in the other. This implies that kurtosis (and other dimensionless statistics) is the same. If η also coincides in the two models, the distributions of price changes are identical.

The power-plus parameterization allows us to illustrate substantial disconnect between the CIR and the flexibility index. In one example where we vary one parameter at time: in this case the flexibility index and the cumulative IRF move in the same direction. In the second example we change three parameters at a time and show how for the same flexibility index cumulative IRF can vary substantially, even keeping the adjustment frequency fixed.

PROPOSITION 16. Assume that Λ is given by a power-plus function. Fix (ν, σ^2) and take two different power plus generalized hazard functions Λ_1 and Λ_2. If they generate the same frequency N_a, then $Kurt_1(\Delta p) > Kurt_2(\Delta p)$ if and only if $\mathcal{F}_1 < \mathcal{F}_2$.

This result is not surprising, since we are varying one parameter only. This comparative static exercise is very far away from the idea of a "sufficient statistic", where one finds a statistic that summarizes significant outputs of a class of models. Even the simple power-plus parameterization affords much more flexibility than varying one parameter can offer.

Now we turn to the second case, where we argue that, however intuitive this might be, relying on the flexibility index can be quite misleading. In Figure 7 we display a number of economies with the same adjustment frequency N_a, and with the same Flexibility Index $\mathcal{F}$, but with very different cumulative response to a monetary shock. That is, we vary the parameters in such a way that both $\mathcal{F}$ and N_a stay constant, while $\mathcal{M}'(0)$ varies substantially. This is done by increasing the power parameter ν and finding the pairs $(\Lambda(0), \kappa)$ that keep N_a and $\mathcal{F}$ constant. We solve this problem numerically and find that for the same N_a and $\mathcal{F}$ the Kurtosis of price changes varies by 90% when ν increases from 2 to 20, as plotted on the Figure 7. The slope of the impulse response at $t = 0$ does not capture the area under it quite well.
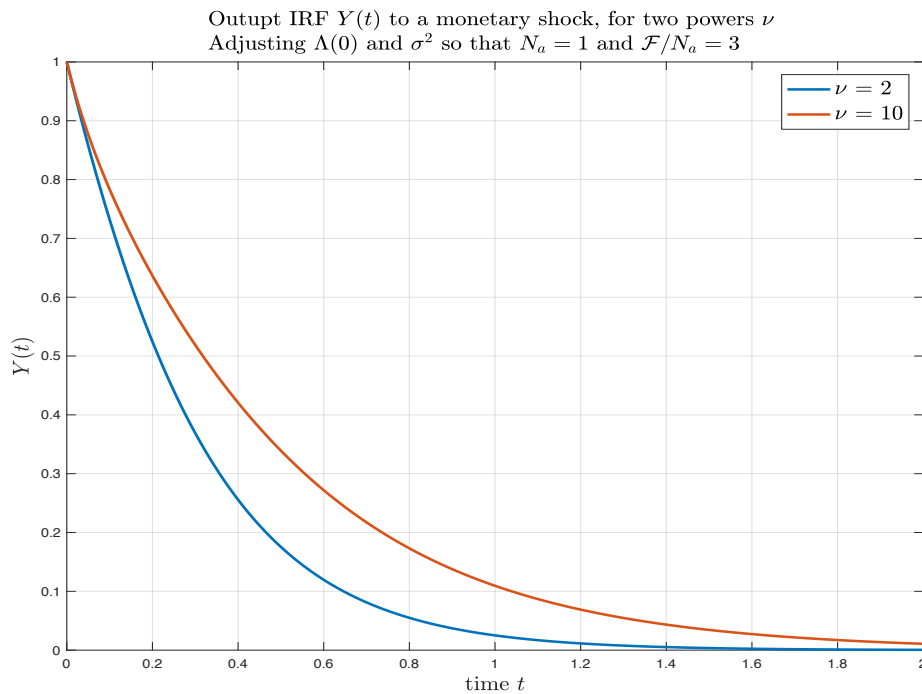
Figure 7: Values of $Kurt(\Delta p)$ or CIR relative to the case of $\nu = 2$, all cases have $\mathcal{F} = 3$, $N_a = 1$
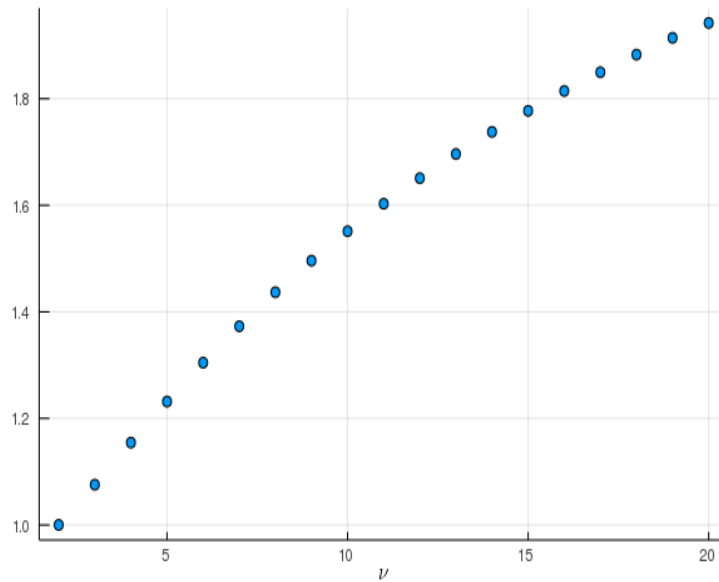
In Figure 8 we take two examples from the previous plot, one with $\nu = 2$ and the other with $\nu = 10$, and display the entire output impulse response function $Y(t)$ as a function of time t. Thus, both IRF's have the same frequency N_a and flexibility index $\mathcal{F}$. The areas under both IRF's are

clearly different, the one for $\nu = 10$ is at least 50% larger than the one for $\nu = 2$, consistent with the values displayed on Figure 7. By construction the slope of $Y(\cdot)$ at $t = 0$ is the same for both cases (i.e. for $\nu = 2$ and $\nu = 10$), since both IRF's have the same Flexibility index $\mathcal{F}$. Yet, the slopes of both impulse responses starts to differ substantially even for low values of t. Since in both cases $N_a = 1$, the values of time in the horizontal axis can be measured in terms of expected adjustment time. For instance, if prices change on average three times a year, meaning $N_a = 3$, then $t = 1$ represents 4 months. The ratio of the two IRF evaluated at $t = 1$ is higher than 4, namely $Y_{10}(1)/Y_2(1) \approx 4.4$. This example shows that even the short run output effect can be substantially different with the same flexibility index.

Figure 8: Impulse Responses for power plus case, both cases with same $\mathcal{F}$ and N_a

Outupt IRF $Y(t)$ to a monetary shock, for two powers ν
Adjusting $\Lambda(0)$ and σ^2 so that $N_a = 1$ and $\mathcal{F}/N_a = 3$

8 Conclusion

We discuss the economic foundations of the generalized hazard function, a flexible modeling block used in several sticky-price setups, and map it to the primitives underlying the firm's optimiza-

tion. We show how to identify the generalized hazard function using observable objects, such as the distribution of price changes, and provide a procedure to recover the hazard and the other microeconomic primitives from data. On the analytical side, we extend the "Kurtosis result" of Alvarez, Le Bihan, and Lippi (2016) to a considerably larger class of models and prove that the Calvo model yields the maximum amount of monetary non-neutrality within the broad class of models we consider. Within a narrower class where the inaction region is bounded, the upper bound is Calvo$^+$. Our empirical strategy addresses unobserved heterogeneity in different products, which we show can be important in estimating the Kurtosis of price changes, and hence in quantifying the real effects of monetary shocks. Finally, we show that another convenient accounting measure for sticky price models, the flexibility index, can sometimes be a poor summary of monetary non-neutrality, both over the long-term as well as over the short-term.

References

Alvarez, Fernando, Martin Beraja, Martin Gonzalez-Rozada, and Pablo Andres Neumeyer. 2019. "From Hyperinflation to Stable Prices: Argentinas Evidence on Menu Cost Modles." *The Quarterly Journal of Economics* 143 (1):451–505.

Alvarez, Fernando and Francesco Lippi. 2020. "Temporary Price Changes, Inflation Regimes, and the Propagation of Monetary Shocks." *American Economic Journal: Macroeconomics* 12 (1):104–52.

Alvarez, Fernando E., Herve Le Bihan, and Francesco Lippi. 2016. "The real effects of monetary shocks in sticky price models: a sufficient statistic approach." *The American Economic Review* 106 (10):2817–2851.

Alvarez, Fernando E. and Francesco Lippi. 2014. "Price setting with menu costs for multi product firms." *Econometrica* 82 (1):89–135.

———. 2019. "The Analytic Theory of a Monetary Shock." Working paper, EIEF.

Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello. 2016. "Monetary Shocks in Models with Inattentive Producers." *Review of Economic Studies* 83:421–459.

Baley, Isaac and Andres Blanco. 2019. "Firm Uncertainty Cycles and the Propagation of Nominal Shocks." *American Economic Journal: Macroeconomics* 11 (1):276–337.

Barro, Robert J. 1972. "A Theory of Monopolistic Price Adjustment." *Review of Economic Studies* 39 (1):17–26.

Berger, David and Joseph Vavra. 2018. "Dynamics of the U.S. price distribution." *European Economic Review* 103:60 – 82.

Caballero, Ricardo J. and Eduardo M. R. A. Engel. 1993a. "Microeconomic Adjustment Hazards and Aggregate Dynamics." *The Quarterly Journal of Economics* 108 (2):359–383.

———. 1999. "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S, s) Approach." *Econometrica* 67 (4):783–826.

Caballero, Ricardo J. and Eduardo M.R.A. Engel. 1993b. "Heterogeneity and output fluctuations in a dynamic menu-cost economy." *The Review of Economic Studies* 60 (1):95.

———. 2007. "Price stickiness in Ss models: New interpretations of old results." *Journal of Monetary Economics* 54 (Supplement):100–121.

Calvo, Guillermo A. 1983. "Staggered prices in a utility-maximizing framework." *Journal of Monetary Economics* 12 (3):383–398.

Caplin, Andrew S and Daniel F Spulber. 1987. "Menu Costs and the Neutrality of Money." *The Quarterly Journal of Economics* 102 (4):703–25.

Carvalho, Carlos. 2006. "Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks." *The B.E. Journal of Macroeconomics* 6 (3):1–58.

Carvalho, Carlos and Oleksiy Kryvtsov. 2018. "Price Selection." Staff Working Papers 18-44, Bank of Canada.

Casella, George and Roger L Berger. 2002. *Statistical inference*, vol. 2. Duxbury Pacific Grove, CA.

Cavallo, Alberto. 2015. "Scraped Data and Sticky Prices." NBER Working Papers 21490, National Bureau of Economic Research, Inc.

Cavallo, Alberto and Roberto Rigobon. 2016. "The Billion Prices Project: Using online prices for measurement and research ." *Journal of Economic Perspectives* 30 (2):151–178.

Costain, James and Anton Nakov. 2011a. "Distributional dynamics under smoothly state-dependent pricing." *Journal of Monetary Economics* 58 (6):646 – 665.

———. 2011b. "Price Adjustments in a General Model of State–Dependent Pricing." *Journal of Money, Credit and Banking* 43:385–406.

Dixit, Avinash. 1991. "Analytical Approximations in Models of Hysteresis." *Review of Economic Studies* 58 (1):141–51.

Dotsey, Michael, Robert G. King, and Alexander L. Wolman. 1999. "State-Dependent Pricing And The General Equilibrium Dynamics Of Money And Output." *The Quarterly Journal of Economics* 114 (2):655–690.

Dotsey, Michael and Alexander L. Wolman. 2020. "Investigating nonneutrality in a state-dependent pricing model with firm-level productivity shocks." *International Economic Review* 61 (1):159–188.

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo. 2011. "Reference Prices, Costs, and Nominal Rigidities." *American Economic Review* 101 (1):234–62.

Gautier, Erwan and Ronan Le Saout. 2015. "The Dynamics of Gasoline Prices: Evidence from Daily French Micro Data." *Journal of Money, Credit and Banking* 47 (6):1063–1089.

Golosov, Mikhail and Robert E. Jr. Lucas. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115:171–199.

Klenow, J. Peter and Benjamin Malin. 2010. "Microeconomic Evidence on Price-Setting." *Handbook of monetary economics* 3:231–284.

Kotlarski, Ignacy. 1967. "On characterizing the gamma and the normal distribution." *Pacific J. Math.* 20 (1):69–76.

Lian, Hairong, Peiguang Wang, and Weigao Ge. 2009. "Unbounded upper and lower solutions method for Sturm?Liouville boundary value problem on infinite intervals." *Nonlinear Analysis: Theory, Methods & Applications* 70 (7):2627–2633.

Nakamura, Emi and Jon Steinsson. 2010. "Monetary Non-Neutrality in a Multisector Menu Cost Model." *The Quarterly Journal of Economics* 125 (3):961–1013.

Petrella, Ivan, Emiliano Santoro, and Lasse de la Porte Simonsen. 2018. "Time-varying Price Flexibility and Inflation Dynamics." Discussion Paper 13027, CEPR.

Reis, Ricardo. 2006. "Inattentive producers." *Review of Economic Studies* 73 (3):793–821.

Sheremirov, Viacheslav. 2019. "Price dispersion and inflation: New facts and theoretical implications." *Journal of Monetary Economics* .

Woodford, Michael. 2009. "Information-Constrained State-Dependent Pricing." *Journal of Monetary Economics* 56:s100–s124.