

# The Macroeconomics of Sticky Prices with Generalized Hazard Functions\*

Fernando Alvarez

University of Chicago and NBER

Francesco Lippi

LUISS University and EIEF

Aleksei Oskolkov

University of Chicago

June 22, 2021

## Abstract

We give a full 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 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. Second, we derive a sufficient statistic for the aggregate effect of a monetary shock: given an *arbitrary* generalized hazard function, the cumulative impulse response of output to a once-and-for-all monetary shock is proportional to the ratio of the kurtosis of the steady-state distribution of price changes over the frequency of price adjustment. We prove that Calvo's model yields the upper bound and Golosov and Lucas's model the lower bound on this measure within the class of random menu cost models.

*Key Words:* Random Menu cost, Impulse response functions, Monetary Shocks, Generalized Hazard Function, Flexibility Index, sufficient statistic

*JEL Classification Numbers:* E3, E5

---

\*We are grateful to Andrey Alexandrov, Ricardo Caballero, Alberto Cavallo, Luca Dedola, Eduardo Engel, for detailed suggestions and discussion of our paper. We also benefited from the comments of Isaac Bailey, Andres Blanco, Klaus Adam, Erwan Gautier, Andrea Ferrara, Anton Nakov, Ricardo Reis, Raphael Schoenle, Rob Shimer, Nathan Zorzi, our discussant Andre Silva and participants to the 2019 ECB-PRISMA Conference in Lisbon, the 2019 workshop *Recent advances in macroeconomics* EIEF (Rome), the Macro-International Workshop at the University of Chicago, NYU, the Paris School of economics, Bocconi University, the Bank of Finland, the University of Michigan, the University of Wisconsin, and the CEPR Joint Conference on Monetary Policy Tools, the Bank of Norway Workshop, the Hydra Conference in Athens, and the students in the 2019 Monetary Economics at The University of Chicago and EIEF and LUISS for their comments. An early version circulated with the title "The Analytics of Monetary Shocks with Generalized Hazard Functions". This material is based, in part, on work supported by the National Science Foundation under grant number SES-1559459.

# 1 Introduction and Summary of Results

Firms' price setting decisions are central to the analysis of nominal shocks. Alternative modeling of such decisions leads to substantively different implications for the relevance, or otherwise the irrelevance, of price stickiness. Understanding the implications of different price setting mechanisms is thus important, but sharp analytical results are difficult to obtain since the analysis involves the dynamic behavior of a cross section of heterogeneous firms. Seminal contributions to this problem were given by [Caplin and Spulber \(1987\)](#); [Caplin and Leahy \(1991\)](#) for models where firms' *only* motive for price changes are *aggregate* nominal shocks. The growing micro evidence spawned by the work of [Bils and Klenow \(2004\)](#) has given rise to a new generation of models where price setting decisions are dominated by firm-level idiosyncratic shocks. In this paper we present sharp analytic results that characterize the propagation of monetary shocks in these new models.

We employ a setup that describes the firm's price setting decision through a "generalized hazard function", a function that relates the firm's price adjustment probability to its own state. Such a function provides a tractable description of the firm's behavior allowing us to study a vast variety of price setting models within a single framework. The notion of a generalized hazard function, and its derivation from first principles, were developed in seminal papers by [Caballero and Engel \(1999, 2007\)](#) and [Dotsey, King, and Wolman \(1999\)](#), and later revisited using information theoretical foundations by [Woodford \(2009\)](#) and [Costain and Nakov \(2011b\)](#). Compared to the workhorse [Calvo \(1983\)](#) model, where the adjustment probability is constant, a generalized hazard function  $\Lambda(x)$  allows it to depend on the state  $x$ , the firm's desired adjustment, such as the markup deviation from the desired level. Such state dependence is appealing theoretically, see e.g. [Barro \(1972\)](#); [Sheshinski and Weiss \(1977\)](#); [Dixit \(1991\)](#); [Golosov and Lucas \(2007\)](#), and has been found to be relevant empirically, see e.g. [Fougere, Le Bihan, and Sevestre \(2007\)](#); [Dias, Marques, and Santos Silva \(2007\)](#); [Eichenbaum, Jaimovich, and Rebelo \(2011\)](#); [Gautier and Saout \(2015\)](#).<sup>1</sup>

The setup nests a broad class of sticky price models where the firm's pricing decisions is represented by a symmetric generalized hazard function  $\Lambda(x)$ . The symmetry arises since we

---

<sup>1</sup>Several authors have employed the generalized hazard function in applications and empirical work. 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 \(forthcoming\)](#).

focus on economies where the idiosyncratic state is driftless, an accurate benchmark for low-inflation economies.<sup>2</sup> A large number of models are nested by this framework, including the “polar” versions: the canonical Calvo model with a constant hazard  $\Lambda(x) = \lambda$  and  $x$  is unbounded, and the [Goloso and Lucas \(2007\)](#) model with  $x$  bounded by the adjustment threshold  $\pm X$  and the hazard equals zero on  $|x| \in (0, X)$  with an “infinite hazard” at the adjustment thresholds.<sup>3</sup> Intermediate cases cover the so called Calvo-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 two main analytical results that give a thorough understanding of the workings of sticky price models, their mapping to the data, and the propagation of monetary shocks. First we show how to identify *all primitives* of the model from readily available data. 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 alternative 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  $\psi$  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 a similar result for a foundation where the firm selects the “probability” of adjustment  $\lambda$  in every period, subject to a cost  $c(\lambda)$ .<sup>4</sup> We show that every non-decreasing generalized hazard rate  $\Lambda(x)$  can be rationalized by a convex cost function  $c(\lambda)$ .

---

<sup>2</sup>See [Section 6](#) for an extension to the case of non-negligible inflation. 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.

<sup>3</sup>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 3](#) we provide a rigorous approximation of the behaviour at barrier as a very high hazard rate to justify this analogy.

<sup>4</sup>This is a simplified version of [Woodford \(2009\)](#), where the cost is modeled in a rational inattention framework. 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.

We complete the mapping between theory and data by showing how to use the (observed) distribution of price changes  $Q(\Delta p)$  and frequency of price changes  $N_a$  to fully identify the generalized hazard function  $\Lambda(x)$ .<sup>5</sup> 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 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. The reason is that, besides  $q(-x)N_a = \Lambda(x)f(x)$ , the functions  $\Lambda(x)$  and  $f(x)$  must jointly obey a Kolmogorov forward equation, a statistical relation that naturally links the existing distribution of price gaps to the probability of changing the gap. We derive the expression to retrieve  $f(x)$  and  $\Lambda(x)$  from  $q(\Delta p)$  and  $N_a$  in closed form, effectively solving two equations for two unknowns.<sup>6</sup>

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. To illustrate our procedure we use publicly available scraped-price data by [Cavallo \(2015\)](#) for the US, estimate the underlying distribution of price changes, measure its kurtosis, and recover  $\Lambda(x)$ ,  $f(x)$ , and  $G(\psi)$ . Interestingly, accounting for measurement error and aggregation, and correcting for unobserved heterogeneity, we find that the data are roughly consistent with a quadratic hazard function, which implies a value of kurtosis around 2, much smaller than those estimated in the previous literature. The latter result will matter for the propagation of nominal shocks. 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 find it to be about 6%, i.e.

---

<sup>5</sup>Such data have been heavily used to discipline sticky price models over the past two decades, see e.g. [Bils and Klenow \(2004\)](#); [Klenow and Malin \(2010\)](#); [Cavallo and Rigobon \(2016\)](#).

<sup>6</sup>The recovery of the function  $f(x)$  from observables relates to the work by [Baley and Blanco \(forthcoming\)](#) who obtain all the moments of  $f(x)$  even in the presence of drift and asymmetries in a simple setup a la [Nakamura and Steinsson \(2010\)](#).

about 94% of price changes show some state dependence. While we focus on the case of zero drift in the state, which is the one applicable to low-inflation countries, all the main results of this part hold in the presence of an arbitrary drift, i.e. in the presence of non-zero steady state inflation, as shown in [Section 6](#). Besides steady state inflation in price setting models, non-zero drift is relevant in other problems, such as depreciation rates on physical assets.

The second main result of the paper gives a sufficient statistic for the aggregate effect of a monetary shock. 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. [Baley and Blanco \(forthcoming\)](#) extended this result to the [Nakamura and Steinsson \(2010\)](#) setup. [Alvarez, Lippi, and Paciello \(2016\)](#) showed that the kurtosis result holds in a large class of rational inattention models, as proposed by [Reis \(2006\)](#), which are purely time dependent. This paper embeds, and provides a substantive generalization, of the previous results: we establish that the kurtosis result holds for *any* symmetric  $\Lambda(x)$  function, allowing for both finite and infinite boundaries of the inaction region. 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. This new result 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.<sup>7</sup>

Our analysis shows 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 \(1983\)](#)

---

<sup>7</sup>One reason the “kurtosis result” fails in their numerical analysis is likely a fraction of firms with flexible prices. Heterogeneity across firm types makes it essential to properly aggregate. Failing to do so will obfuscate the result, which holds for each firm’s type. We analyze the case with heterogenous firms in [equation \(49\)](#) and discussion around it.

case. The smallest, equal to one, corresponds to the pure menu cost model of Golosov and Lucas (2007). This result is interesting because non-decreasing hazard rates describe the entire set of models with random menu costs and information gathering, and thus it establishes Calvo as an upper bound within this broad class. For a Kurtosis higher than six one would need to come up with an economic foundation for a (locally) *decreasing* hazard function.

**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 the 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  $\Lambda$ . We show that in both models, given  $\Lambda$ , one can recover the primitive cost, either  $G$  or  $c$  (Theorem 1). 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  $\Lambda$  starting from the observed distribution of the size of price changes (Theorem 2) and presents an application using Cavallo's (2015) data set. Section 5 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 (Theorem 3). In Section 6 we extend the random menu cost model to accommodate inflation and derive the inverse mapping for that case (Theorems 4 and 5). Section 7 discusses scope and limitations of the results, and the relation with the recent literature. Section 8 concludes discussing extensions and future work. The proofs of the theorems are in Appendix A, the proofs of the remaining propositions are in Appendix B. Several extensions are developed in the appendices.

## 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.<sup>8</sup>

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. 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. If prices are not changed, the price gap  $x$  evolves as a standard Brownian Motion with zero drift and variance  $\sigma^2$ . The lack of drift indicates that the economy under consideration has no inflation.<sup>9</sup>

## 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 to eliminate a long tail of delayed adjustments that seems counterfactual. The generalized model allows the firm to draw a fixed menu cost  $\psi$  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

---

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

<sup>9</sup>See The Online appendix B in [Alvarez and Lippi \(2014\)](#) for a microfoundation of this model. [Nakamura et al. \(2018\)](#) and [Alvarez, Beraja, Gonzalez-Rozada, and Neumeyer \(2019\)](#) focus on a models with zero inflation and argue that they provide 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 for simple sS models. Moreover, these papers provides evidence consistent with this behaviour using variation on inflation rates close to zero for US and Argentina.

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:

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

Two points are worth making. First, given the symmetry of  $Bx^2$ , the value function is *symmetric around*  $x = 0$ .<sup>10</sup> A proof can be constructed by a simple guess and verify argument. Second, if  $\Psi = \infty$ , 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  $\min_z v(z)$  is the value right after adjustment, and given the symmetry of the return function, we have  $v(0) = \min_z v(z)$ . Thus we can simply write that for all  $x$

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

For the case where  $\Psi < \infty$  we can use that the optimal decision rule has a threshold  $X < \infty$  such

---

<sup>10</sup>This also relies on the the assumption that there is no drift in the uncontrolled process for  $x$ . The case with a non-zero drift is treated in [Section 6](#).



that if  $|x| \geq X$  the firm pays the fixed cost  $\Psi$ . Thus we can write

$$rv(x) = \begin{cases} Bx^2 + \frac{\sigma^2}{2}v''(x) + \kappa \int_0^{v(x)-v(0)} \psi dG(\psi) + \kappa [v(0) - v(x)] G(v(x) - v(0)), & |x| \leq X \\ r(v(0) + \Psi), & |x| > X \end{cases} \quad (1)$$

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

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

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 L](#) considers the case of a discrete distribution  $G$ , where  $\psi$  takes finitely many values.<sup>11</sup> In either case we have the following smooth pasting and optimal return point conditions:

$$v'(-X) = v'(X) = v'(0) = 0$$

We are now ready to define the *generalized hazard rate*,  $\Lambda : (-X, X) \rightarrow \mathbb{R}_+$ , 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  $G$ :

$$\Lambda(x) = \kappa G(v(x) - v(0)) \text{ for all } x \in (-X, X) \quad . \quad (2)$$

---

<sup>11</sup>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  $\Lambda$  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  $\kappa$ . While the function  $\Lambda$  is not defined at  $x = \pm X$ , we abuse notation and let  $\Lambda(X) = \lim_{x \rightarrow X} \Lambda(x) = \kappa G(\Psi) = \kappa$ .

### 2.1.1 Rationalizing a given generalized hazard $\Lambda$

We next show that any increasing, differentiable, symmetric and bounded hazard rate  $\Lambda$  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  $\Lambda$ , 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  $\Lambda$  and the values of three parameters:  $r$ ,  $B$ , and  $\sigma^2$ . Three remarks are in order. First, the values of the fixed costs  $\psi$  are measured relative to  $B$ , and thus the optimal decision rules depend only on the distribution of  $\psi/B$ . Second, we show that  $\sigma^2$ , while in principle unobservable, is encoded in the frequency and variance of price changes. Thus, once  $\Lambda$  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 L](#) considers discrete distributions of costs which imply an hazard  $\Lambda$  that is a step function. In this case we can recover  $G$  starting from  $\Lambda$  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

$$rv(x) = Bx^2 + \frac{\sigma^2}{2}v''(x) + \int_0^{v(x)-v(0)} \kappa[\psi + v(0) - v(x)]g(\psi)d\psi + \kappa[v(0) - v(x)]G(0) \quad (3)$$

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

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  $\Lambda$  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  $\Lambda$  describing the probability per unit of time of a price adjustment if the price gap is  $|x| < X$ . We have the following:

**LEMMA 1.** Assume  $\Lambda : [-X, X]$  be a positive function. Define  $u$  as the solution of the linear ordinary differential equation

$$[r + \Lambda(x)] u(x) = 2Bx + \frac{\sigma^2}{2} u''(x) \text{ for } x \in [0, X] \quad (4)$$

with boundary conditions  $u(0) = u(X) = 0$ . The solution for  $u$  exists and it is unique. If  $\Lambda$  is the GHF corresponding to the value function  $v$ , then

$$v(x) = u'(0) \frac{\sigma^2}{2r} + \int_0^x u(z) dz \text{ for } x \in [0, X] . \quad (5)$$

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:

**THEOREM 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 function  $\Lambda(\cdot) : (-X, X) \rightarrow \mathbb{R}_+$  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  $G(\cdot) : [0, \Psi] \rightarrow [0, 1]$  with a density  $g(\cdot)$ , continuous on  $(0, \Psi)$ , and possibly a mass point  $G(0) > 0$ , that uniquely rationalizes  $\Lambda$  as the generalized hazard function of the value function that solves

equation (3). Using the auxiliary function  $u$  in Lemma 1 and  $U(x) = \int_0^x u(z)dz$  for  $x \in (0, X)$ ,

$$\kappa = \lim_{x \uparrow X} \Lambda(x) \quad , \quad \Psi = U(X) \quad , \quad G(0) = \frac{\Lambda(0)}{\kappa} \quad (6)$$

$$g(U(x)) = \frac{\Lambda'(x)}{u(x)\kappa} \text{ for all } x \in (0, X) \text{ with } \psi = U(x) \quad (7)$$

The theorem allows us to retrieve the primitives of a fully specified price setting problem starting from any given non-decreasing hazard function  $\Lambda$ . 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  $\Lambda'(\cdot) > 0$ . See Section 6 for an identification result that extends Theorem 1 to the case of drift on the state, i.e. to non-zero steady state inflation.

**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  $|x| \in [0, X]$ .<sup>12</sup> We can solve for the auxiliary function  $u(x)$  using Lemma 1. This yields a polynomial:

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

satisfying the ODE in equation (4) and the boundary conditions  $u(0) = u(X) = 0$ . Straightforward application of the method of undetermined coefficients gives the recursive relation

$$a_3 = \frac{(r + \Lambda_0)a_1 - 2B}{3\sigma^2} \quad , \quad a_{2i+1} = \frac{(r + \Lambda_0)a_{2i-1} + \Lambda_2 a_{2i-3}}{\sigma^2 i(2i + 1)} \quad \text{for } i = 2, 3, \dots$$

All coefficients are determined as a function of  $a_1$ , which is pinned down by the boundary condition  $u(X) = 0$ . Application of Theorem 1 gives  $U(x) = \sum_{i=1}^{\infty} \frac{a_{2i-1}}{2i} x^{2i}$ , the value function  $v(x) =$

---

<sup>12</sup>If  $\Lambda$  is symmetric and smooth, it 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 H we show that if  $\Lambda$  does not admit a quadratic approximation around  $x = 0$ , the underlying density  $g$  exhibits non-generic behavior.

$a_1 \frac{\sigma^2}{2r} + U(x)$ , the arrival rate  $\kappa = \Lambda_0 + \Lambda_2 X^2$ , the distribution function  $G(\psi)$  with

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

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

$$g(U(x)) = \frac{2\Lambda_2 x}{\kappa u(x)} \quad (8)$$

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  $\ell$ . 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  $\Lambda$ . As in the previous case, we also allow the firm to pay a deterministic menu cost  $\Psi$  to change its price with certainty. This  $\Psi$  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  $\Lambda$  can be rationalized by some increasing and convex cost function  $c$ . One difference between the two setups is that the resulting  $\Lambda$  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:

$$rv(x) = \min \left\{ Bx^2 + \frac{\sigma^2}{2} v''(x) + \min_{\ell \geq 0} \{ \ell (v(0) - v(x)) + c(\ell) \}, r(\Psi + v(0)) \right\} \quad (9)$$

We assume that the cost function  $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is increasing and convex in  $\ell$ , and that  $c(\ell) \rightarrow \infty$  as  $\ell \rightarrow \infty$ . 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  $\Lambda$ . Allowing for flat segments in  $c$  implies that the minimizer  $\ell^*$  may be an interval for some  $x$ , which we can represent with a discontinuity in  $\Lambda$  at that value of  $x$ . We can now state a result that echoes the one in [Theorem 1](#):

**PROPOSITION 1.** Fix a discount rate, curvature, variance, and a value of the threshold  $(r, B, \sigma^2, X)$ , all positive. Let  $\Lambda(\cdot) : (-X, X) \rightarrow \mathbb{R}_+$ . 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  $c(\cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and a real number  $\Psi$  that rationalize the postulated decision rule as in [equation \(9\)](#). Moreover, the marginal cost  $c'(\cdot)$  can be uniquely constructed by solving a second order linear ordinary differential equation.

[Appendix G](#) 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  $\Lambda$  is a more fundamental object. Furthermore, as explained above, the model of this section allows a slightly larger set of generalized hazard functions  $\Lambda$ .

### 3 Steady State observable statistics

In this section we show how to use the hazard function  $\Lambda$  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  $\Psi$  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  $\Lambda$  that summarizes the firm's optimal decisions:

**ASSUMPTION 1.** Let  $\Lambda : (-X, X) \rightarrow \mathbb{R}_+$ , 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  $\Lambda$  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  $f(\cdot) : (-X, X) \rightarrow \mathbb{R}_+$ . 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:

$$f(x)\Lambda(x) = \frac{\sigma^2}{2}f''(x) \text{ for all } x \in [0, X), x \neq 0 \text{ and } x \notin \mathbb{J} \quad (10)$$

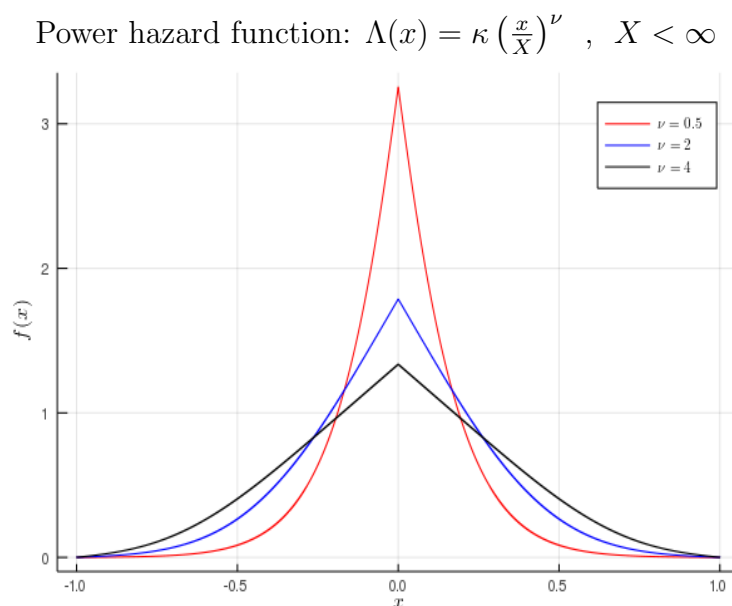
$$\text{with boundary conditions: } \frac{1}{2} = \int_0^X f(x)dx \quad \text{and} \quad \lim_{x \rightarrow X} f(x) = 0. \quad (11)$$

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 \(10\)-\(11\)](#) 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 \(2018\)](#).

**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 because 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

$$N_a = 2 \left[ \int_0^X f(x)\Lambda(x)dx - \frac{\sigma^2}{2}f'(X) \right]. \quad (12)$$

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



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:<sup>13</sup>

$$N_a \text{Var}(\Delta p) = \sigma^2 \tag{13}$$

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 \(12\)](#) to replace the Kolmogorov forward equation for  $f$ , and integrate by parts to obtain that:

$$s \equiv \frac{\int_{-X}^X \Lambda(x) f(x) dx}{N_a} = 1 - \frac{\sigma^2 |f'(X)|}{N_a} = 1 - \frac{|f'(X)|}{|f'(0)|} \text{ since } N_a = \sigma^2 |f'(0)| \tag{14}$$

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  $x = 0$ .

---

<sup>13</sup>The key assumption for this result to hold is that the price gap is closed upon adjustment. See [Appendix A](#), proof of Theorem 3, for a detailed argument.



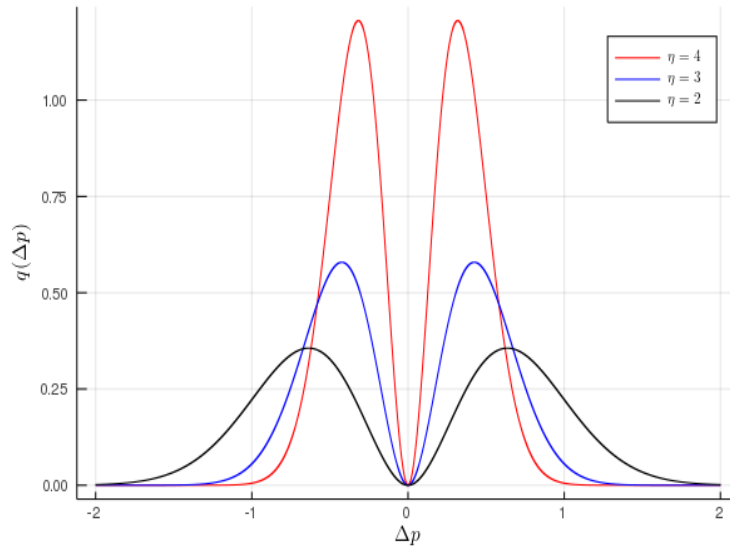
**Distribution of price changes.** Recall that upon any price change the firm “closes” its gap  $x$ , i.e. the size of the adjustment is  $\Delta p = -x$ . If  $X < \infty$  then the distribution of price changes has a mass point at  $\Delta p = -X$ . The mass of such price changes is equal to  $\frac{\sigma^2}{2}|f'(X)|$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

$$\Delta p = \begin{cases} -x & \text{w/ density } q(-x) \equiv \frac{\Lambda(x)f(x)}{N_a} \text{ for } x \in (0, X) \\ -X & \text{w/ probability } \frac{\sigma^2|f'(X)|}{N_a} \end{cases} \quad (15)$$

Note that  $1 - s$ , as defined in [equation \(14\)](#), 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 \(15\)](#), where the density of price gaps  $f$  solves the Kolmogorov forward equation

equation (10). 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 (14). The parameter  $\eta$  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 (15):

$$Var(\Delta p) = \frac{2 \left[ \int_0^X x^2 \Lambda(x) f(x) dx - X^2 \frac{\sigma^2}{2} f'(X) \right]}{N_a} \quad (16)$$

$$Kurt(\Delta p) = \frac{2 \left[ \int_0^X x^4 \Lambda(x) f(x) dx - X^4 \frac{\sigma^2}{2} f'(X) \right]}{N_a} \frac{1}{[Var(\Delta p)]^2} \quad (17)$$

**Standardization.** It is useful to rescale the firm's decision rule to isolate the role of the shape of  $\Lambda$  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  $\Lambda : (-X, X) \rightarrow \mathbb{R}_+$  satisfying Assumption 1. Given the triplet  $\{X, \Lambda, \sigma^2\}$ , we can compute the corresponding density of price changes  $q(\cdot) : (-X, X) \rightarrow \mathbb{R}$ , 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 2.** 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  $\tilde{X} = bX$ ,  $\tilde{\Lambda}(z) = \Lambda(z/b)$  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  $z \in (-Xb, Xb)$ .

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,  $\tilde{Var}(\Delta p) = 1$ . 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  $\Lambda$ , 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  $\Lambda$  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 is easy to see that  $\{\hat{X}, \hat{\Lambda}, \hat{\sigma}^2\}$  has  $\hat{N}_a = N_a/k$  and  $\hat{Q}(x) = Q(x)$  for all  $x$ .<sup>14</sup>

**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  $\Lambda$  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 3.** Let  $X < \infty$  and let  $\Lambda : [0, X) \rightarrow \mathbb{R}_+$  be a continuous generalized hazard function, where  $f : [0, X] \rightarrow \mathbb{R}_+$  is its corresponding invariant density, assumed to be symmetric. Let  $\Lambda_k : [0, \infty) \rightarrow \mathbb{R}_+$  be defined as  $\Lambda_k(x) = \Lambda(x)$  if  $x < X$  and  $\Lambda_k(x) = k$  otherwise. Let also  $f_k : [0, \infty) \rightarrow \mathbb{R}_+$  be the invariant density associated with  $\Lambda_k$ , also assumed symmetric for negative  $x$ 's. Then  $f_k$  converges uniformly to  $f$  in  $[0, X]$  as  $k \rightarrow \infty$ .

## 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  $\Lambda$  from the observable distribution of price changes. These two objects then allow us to recover the underlying distribution  $G$  of adjustment costs  $\psi$  in a random menu cost model of [Section 2.1](#). We apply the algorithm to data taken from [Cavallo \(2015\)](#), fitting the distribution of

---

<sup>14</sup>In [Appendix I](#) we consider an alternative normalization, suitable for the case where  $X < \infty$ .

price changes  $Q$  and recovering  $f$ ,  $\Lambda$ , and  $G$ . For future reference, we pay particular attention to estimating the kurtosis of the distribution of price changes.

To do this, we first characterize the restrictions that an increasing hazard function  $\Lambda$  imposes on  $Q$  and establish a mapping from the observables (price changes) to the distribution of price gaps  $f$  and adjustment hazard  $\Lambda$ . 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 2](#). 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  $\Lambda$ , and from them obtain the distribution of random menu cost  $G$  using the characterization in [Theorem 1](#). In [Section 6](#) we show that the identification result extends to the case of drift in the state, i.e. to non-zero steady state inflation.

**Identification of  $f$  and  $\Lambda$ .** 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 4.** Let  $Q$  be the CDF of price changes corresponding to a generalized hazard function  $\Lambda$  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 procedure essentially amounts to solving a system of two equations in two unknowns:  $f$  and  $\Lambda$ . The first equation is the Kolmogorov forward [equation \(10\)](#) that shows how the generalized hazard  $\Lambda$  determines the distribution of price gaps. We first integrate it twice and replace  $\sigma^2$  as in [equation \(13\)](#). Once we have  $f$ , it is straightforward to get  $\Lambda$  using the second equation,  $f(x)\Lambda(x) = q(x)N_a$ . This one is simple accounting: the number

of firms changing prices by  $x$  is the number  $f(x)$  of firms with this gap times the probability  $\Lambda(x)$  that they change the price.

**THEOREM 2.** Let  $\Delta p$  be price changes, and let  $Q$  and  $q$  be the CDF and corresponding density of price changes corresponding to a generalized hazard function  $\Lambda$  satisfying [Assumption 1](#). Let  $N_a$  be the frequency of price changes. The density for the invariant distribution  $f(x)$  is given by

$$f(x) = \frac{2}{\text{Var}(\Delta p)} \left[ \int_x^X (1 - Q(z)) dz \right] \text{ for all } x \in (0, X) \quad (18)$$

and  $f(-x) = f(x)$ , where  $\text{Var}(\Delta p)$  is the variance of the price changes computed using  $Q$ . The generalized adjustment hazard  $\Lambda(x)$  is given by

$$\Lambda(x) = \frac{N_a \text{Var}(\Delta p)}{2} \frac{q(x)}{\int_x^X (1 - Q(z)) dz} \text{ for all } x \in [0, X) \quad (19)$$

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

Recall that the function  $\Lambda$  implied by the models of [Section 2.1](#) and [Section 2.2](#) is increasing in  $x \in (0, X)$ . If  $\Lambda$  is increasing in  $(0, X)$ , the right hand side of [equation \(19\)](#) must be increasing. At any  $x$  where  $\Lambda$  is differentiable,

$$\frac{\Lambda'(x)}{\Lambda(x)} = \frac{q'(x)}{q(x)} + \frac{1 - Q(x)}{\int_x^X (1 - Q(z)) dz} \geq 0 \text{ for all } x \in (0, X), x \notin \mathbb{J} \quad (20)$$

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 \(19\)](#) must be bounded. If  $\Lambda$  is increasing, this is equivalent to

$$\lim_{x \rightarrow X} \frac{q(x)}{\int_x^X (1 - Q(z)) dz} = \lim_{x \rightarrow X} \frac{q'(x)}{-(1 - Q(x))} \leq C \quad (21)$$

for some constant  $C$ . Moreover, if  $X < \infty$ , then  $\lim_{x \rightarrow X} (1 - Q(x)) > 0$  and hence  $\lim_{x \rightarrow X} q'(x)$  must be finite. If  $X = \infty$ , then  $\lim_{x \rightarrow X} \frac{q''(x)}{q(x)} \leq C$  by L'Hopital rule. Note that if  $q$  has exponential tails, [equation \(21\)](#) is satisfied even if  $X = \infty$ . Moreover, since the model of [Section 2.2](#) does *not* imply a bounded  $\Lambda$  it does not require [equation \(21\)](#). In [Section 6](#) we show that the result of

**Theorem 2** extends to the case where the state has a constant arbitrary non-zero drift, and provide closed form expressions for  $f$  and  $\Lambda$ .

**A simple measure of state dependence.** The expression for  $\Lambda$  in **Theorem 2** 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”:

$$\mathcal{C} \equiv \frac{\Lambda(0)}{N_a} \quad (22)$$

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<sup>+</sup> model. In that special case, the function  $\Lambda$  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 **Theorem 2** 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 (22)**, is given by

$$\mathcal{C} = \frac{Var(\Delta p)}{2 \mathbb{E}[|\Delta p|]} q(0). \quad (23)$$

Moreover, using **equation (12)** for  $N_a$ ,  $\mathcal{C} \leq 1$  if  $\Lambda$  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 (23)** 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  $\mathbb{E}[|\Delta p|]$ .<sup>15</sup>

## 4.1 Estimating $\Lambda$ and $Q$ with Unobserved Heterogeneity.

Armed with [Theorem 2](#) we use the data to recover  $f$ ,  $\Lambda$ , and the model primitives like the distribution of menu cost. However, the estimation of  $q$  in many datasets requires dealing with a simple yet pervasive form of unobserved heterogeneity. The model presented above assumes all observations are generated by the same model, and in particular that all goods share the same “scale” parameter  $\sigma$ . This assumption is often violated in the data where similar goods display different variance in the size of the price changes. The reason to account for unobserved heterogeneity is that a mixture of distributions with identical kurtosis but different variances will give rise to a higher kurtosis.<sup>16</sup> For instance, two categories of goods with identical frequency and kurtosis, and hence with an identical propagation of monetary shocks, might give rise to a larger estimate of the kurtosis if they feature different variance of shocks, thus biasing the predictions of the theory.

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 2](#) discusses exactly this type of transformation that changes the scale without affecting shape. The setup is similar to a random effect model, 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:

$$\Delta p_{it} = b_i \Delta \tilde{p}_t \text{ for } i \in I \text{ and } t \in T(i) \tag{24}$$

Here  $b_i$  corresponds to the scaling factor  $b$  in [Proposition 2](#). The six identification assumptions 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 4](#), for all  $t \in \bigcup_{i \in I} T(i)$

---

<sup>15</sup>Clearly, this statistic also holds in the case of non-zero drift, since  $\mathcal{C} \equiv \Lambda(x^*)/N_a = q(0)/f(x^*)$ , where the expression for  $f(x^*)$  in the presence of drift is given in [Section 6](#).

<sup>16</sup>See [Proposition 11](#) in [Appendix C](#) for the formal treatment of this result.

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 4](#) 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 have the following result:

**PROPOSITION 5.** Consider two pairs of integer numbers  $(j, k)$  and  $(j', k')$  such that  $j + k = j' + k'$ . Under the assumptions stated above we have:

$$\frac{\mathbb{E}[(\Delta\tilde{p}_t)^j]\mathbb{E}[(\Delta\tilde{p}_t)^k]}{\mathbb{E}[(\Delta\tilde{p}_t)^{j'}]\mathbb{E}[(\Delta\tilde{p}_t)^{k'}]} = \frac{\mathbb{E}[(\Delta p_{it})^j(\Delta p_{is})^k]}{\mathbb{E}[(\Delta p_{it})^{j'}(\Delta p_{is})^{k'}]} \quad (25)$$

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  $\Delta\tilde{p}_t$ :

$$\mathbb{E}[(\Delta\tilde{p}_t)^{2k+2}] = \mathbb{E}[(\Delta\tilde{p}_t)^{2k}] \cdot \frac{\mathbb{E}[(\Delta p_{it})^{2k+2}]}{\mathbb{E}[(\Delta p_{it})^{2k}(\Delta p_{is})^2]} \text{ for all } k \geq 0 \quad (26)$$

which only uses [equation \(25\)](#) 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 \(26\)](#), thus obtaining a *non-parametric identification* of the density  $q$ .<sup>17</sup> Second, for future reference we display an expression for the Kurtosis of  $\Delta\tilde{p}$ :

$$Kurt(\Delta\tilde{p}_t) = \frac{\mathbb{E}[(\Delta p_{it})^4]}{\mathbb{E}[(\Delta p_{it})^2(\Delta p_{is})^2]} = \frac{Kurt(\Delta p_{it})}{1 + \text{corr}(\Delta p_{it}^2, \Delta p_{is}^2)CV(\Delta p_{it}^2)CV(\Delta p_{is}^2)} \text{ for } t \neq s \quad (27)$$

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

---

<sup>17</sup>This fully characterizes the distribution of  $\Delta\tilde{p}_t$ , since its odd moments are equal to zero due to symmetry.



computed by pooling different goods without accounting for heterogeneity. Whenever the squares of the 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.

**Cavallo’s scraped data.** We use the scraped price data from the Billion Prices Project by Cavallo (2015) to reduce the measurement error present in other data sets for example due to time aggregation using the average revenue.<sup>18</sup> The time span of the 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.<sup>19</sup>

Table 1 presents the summary statistics of 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 heterogeneity (u). To implement the latter procedure, we use equation (27). This equation is a particular case of equation (25) 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

$$\frac{1}{\#I} \sum_{i \in I} \frac{1}{\#T(i)(\#T(i) - 1)} \sum_{t, s \in T(i), t \neq s} |\Delta p_{it}|^j |\Delta p_{is}|^k \quad (28)$$

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 C we tabulate implied correlations recovered from equation (27). These turn out to be

<sup>18</sup>Link: <http://www.thebillionpricesproject.com/datasets/>, using store number 1.

<sup>19</sup>We remove 87 (larger than 150 log points) out of 326,570 price changes for products with at least three spells.

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	3,437	74,464	0.002	0.341	3.4 (0.16)	1.7 (0.07)	0.08	0.07
119	3,225	56,527	0.002	0.33	3.8 (0.09)	2.0 (0.05)	0.09	0.06
1212	2,551	30,343	-0.001	0.25	3.5 (0.27)	2.1 (0.16)	0.04	0.04
122	1,401	27,321	0.002	0.34	3.0 (0.09)	1.7 (0.05)	0.12	0.09
118	1,388	30,111	0.003	0.31	3.6 (0.24)	2.0 (0.12)	0.08	0.08
117	1,154	20,995	0.007	0.31	3.5 (0.14)	2.0 (0.05)	0.07	0.06
561	1,032	17,724	0.002	0.26	3.3 (0.22)	1.8 (0.13)	0.03	0.03

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”. Bootstrapped standard errors in parenthesis.

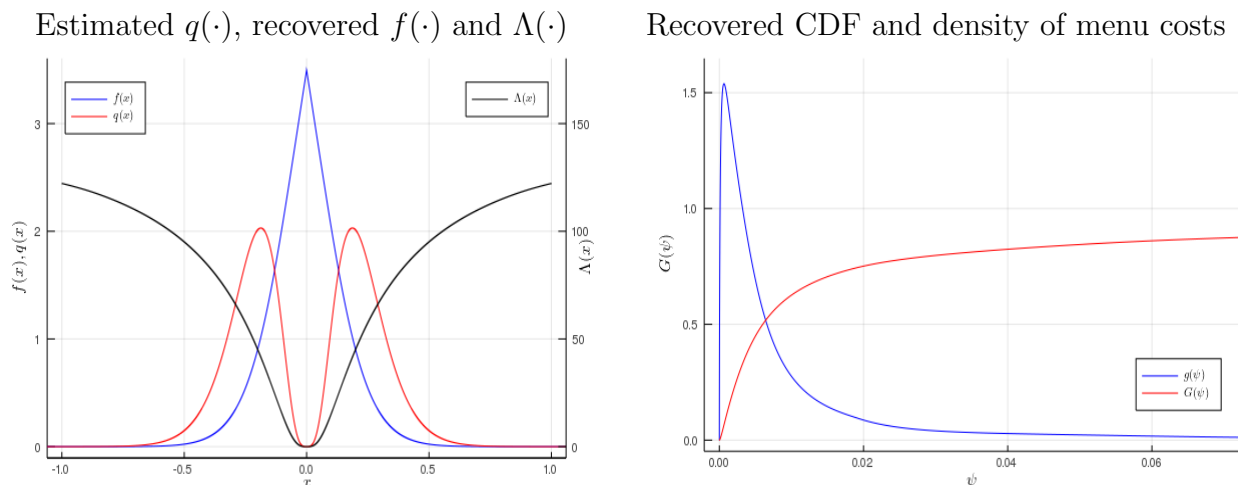
in the range between 0.29 and 0.45.

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 large number of moments, substantially decreasing precision. Instead, we estimate the Kurtosis and use the unit variance restriction to fit the scale and size of the Gamma distribution. The fitted density  $q$  is presented on the left panel of [Figure 3](#) together with the underlying generalized hazard function  $\Lambda$  and the density of price gaps  $f$ .<sup>20</sup> Finally, the right panel contains the distribution of menu cost recovered from the resulting hazard  $\Lambda$ . 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. The Gamma distribution is convenient, since by [Theorem 2](#)

<sup>20</sup>In [Appendix C](#) we detail our algorithm and show extended results from fitting a mixture of two Gamma distributions.

it implies a bounded  $\Lambda$ .<sup>21</sup>

Figure 3: Estimated distribution and implied cost functions (w. unobserved heterogeneity)



**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  $\Lambda$  (including the value of  $X$ , the barrier) and  $\sigma^2$ . Recall from [Theorem 2](#) 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 \(23\)](#). 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 estimated density  $q$  at  $\Delta p = 0$ .

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 \(23\)](#) produces an upward bias in the estimate. We derive an unbiased estimator, the result being analogous to the one in [Proposition 5](#). We express this estimator as a function of the pooled estimator  $\mathcal{C}_{pooled}$  and a correction due to the unobserved heterogeneity:

<sup>21</sup>In [Appendix C](#) we show the procedure to recover the cost function  $c$  corresponding to the model in [Section 2.2](#). This model allows for an unbounded  $\Lambda$ , 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.

PROPOSITION 6. Under the assumptions 1-5 stated above,

$$\mathcal{C} = \mathcal{C}_{pooled} \left( 1 + \frac{Cov(b_i^{-1}, b_i^2)}{\mathbb{E}[b_i^{-1}]\mathbb{E}[b_i^2]} \right) < \mathcal{C}_{pooled} \quad (29)$$

where the two components are given by

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

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  $(\Lambda, X, \sigma^2)$ , then there is no variation in  $b_i$ , and  $\mathcal{C} = \mathcal{C}_{pooled}$ .

## 5 A Sufficient Statistic for Small 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 by 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

$$m(x) = -\mathbb{E} \left[ \int_0^\tau x(t) dt \mid x(0) = x \right] \quad (30)$$

where  $\tau$  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 on the right hand side of [equation \(30\)](#) is with respect to the process for  $x$ , a jump-diffusion with jump intensity  $\Lambda(x)$ , diffusion variance  $\sigma^2$ , and zero drift. The function  $m : [-X, X] \rightarrow \mathbb{R}$  is once continuously differentiable, antisymmetric around  $x = 0$ , and satisfies:

$$m(x)\Lambda(x) = -x + \frac{\sigma^2}{2}m''(x) \text{ for all } x \text{ at which } \Lambda \text{ is continuous} \quad (31)$$

$$0 = m(X) \text{ if } X < \infty \text{ and } \lim_{x \rightarrow \infty} \frac{|m(x)|}{x} \leq \frac{1}{\inf_y \Lambda(y)} \text{ if } X = \infty . \quad (32)$$

Now we can define the cumulative impulse response to a monetary shock of size  $\delta$  as

$$\mathcal{M}(\delta) = \int_{-X}^X m(x - \delta) f(x) dx . \quad (33)$$

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 2](#), 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

$$\tilde{\mathcal{M}}(\delta) = \int_{-\tilde{X}}^{\tilde{X}-\delta} \tilde{m}(x) \tilde{f}(x + \delta) dx \quad (34)$$

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 7.** 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

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

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 \(34\)](#), and evaluating at  $\delta = 0$ . Summarizing, [Proposition 7](#) states 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  $\delta \geq 2\tilde{X}$ , 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.<sup>22</sup> The marginal version of this cumulative impulse response is

$$\mathcal{M}'(0) = - \int_{-X}^X m'(x) f(x) dx \quad (35)$$

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.

**THEOREM 3.** 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:

$$\mathcal{M}(\delta) = \frac{Kurt(\Delta p)}{6N_a} \delta + o(\delta^2) \quad (36)$$

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

Our results from [Theorem 1](#) show that only weakly increasing  $\Lambda$  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. [Theorem 3](#) holds for such functions too. It makes no reference to the micro-foundations behind  $\Lambda$  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.

---

<sup>22</sup>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.

**Kurtosis.** The next proposition discusses the properties of generalized hazard function that determine the Kurtosis of price changes. We concentrate on a case where we hold the adjustment frequency constant. This procedure allows us to isolate the effect of a change in  $\Lambda$  on selection from its effect on the frequency (recall that that fixing the frequency is equivalent to fixing the time units). Moreover, with the frequency fixed, the kurtosis of price changes directly maps into the approximate cumulative impulse response.

**PROPOSITION 8.** Fix  $N_a$  and consider two hazard functions  $\Lambda_1(x)$  and  $\Lambda_2(x)$  with the respective boundaries  $X_1$  and  $X_2$  with  $0 < X_2 \leq X_1 \leq \infty$ . Let  $\Lambda_1(0) > \Lambda_2(0)$  and let  $\Lambda_1(x) - \Lambda_2(x)$  change the 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  $\Lambda_1$  generates more adjustment for smaller  $x$ , and  $\Lambda_2$  generates more for larger ones. Selection is therefore more pronounced with  $\Lambda_2$ , 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<sup>+</sup> 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 **Theorem 3**, it also means a constant  $\Lambda$  maximizes  $\mathcal{M}'(0)$  for a fixed  $N_a$ . This highlights the role of selection. A strictly increasing rate of adjustment  $\Lambda$  implies positive selection, so the firms with larger deviations are more likely to adjust. When  $\Lambda$  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 8](#) 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  $\Lambda$  (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  $\Lambda$  as

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

To understand what it means for two functions to have the same curvature, take some arbitrary  $\Lambda$  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:

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

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  $\Lambda_1$  and  $\Lambda_2$  have the same curvature as  $\Lambda$ ). When  $\Lambda_1$  and  $\Lambda_2$  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.

## 5.1 Illustration with a power hazard function

In this section we describe the case where the generalized hazard is a power function with the power parameter  $\nu$ . 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 hazards functions with  $\nu = 2$ .

We use this example to illustrate how the parameters affecting the shape of  $\Lambda$  determine the Kurtosis of price changes. To do this, we first show that, with  $\nu$  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 stronger 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  $\nu$ , which governs the shape of  $\Lambda$ .

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

$$\rho z^\nu \hat{f}(z) = \hat{f}''(z) \text{ with } \hat{f}(1) = 0 \text{ and } \int_0^1 \hat{f}(z) dz = \frac{1}{2} \quad (37)$$

where  $\rho \equiv 2\kappa X^2/\sigma^2$ . The solution to [equation \(37\)](#) is given by

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

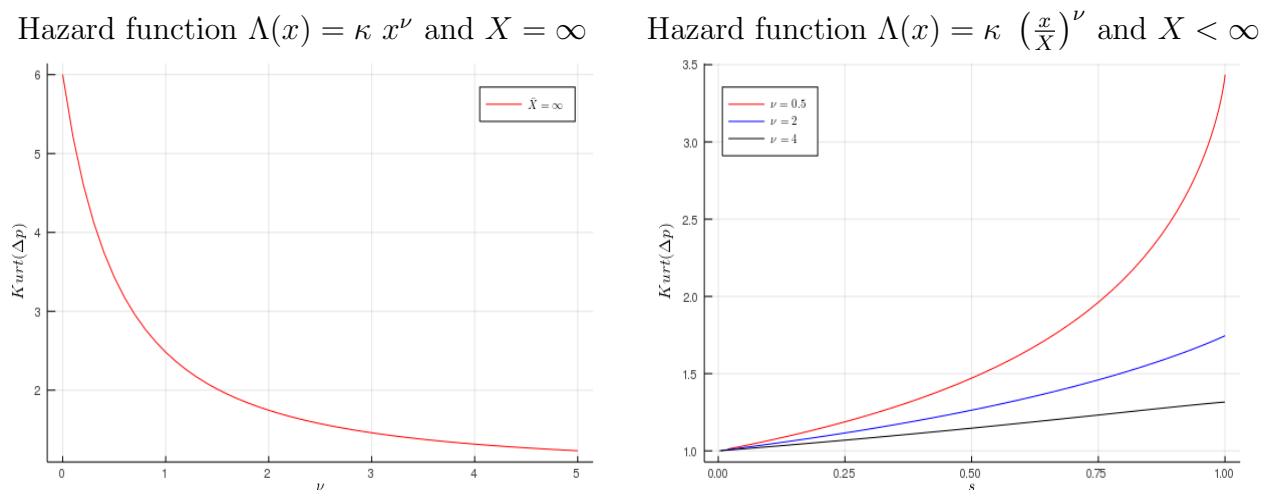
This is a combination of modified Bessel functions of the first and second kind  $\mathcal{I}_{\frac{1}{\nu+2}}$  and  $\mathcal{K}_{\frac{1}{\nu+2}}$  of order  $\frac{1}{\nu+2}$ , where the constants  $c_1, c_2$  are chosen to satisfy the two boundary conditions described in [equation \(37\)](#). The form  $\hat{f}$  depends on the parameters  $(\rho, \nu)$  that capture adjustment coming from random menu costs. Note that if  $X = \infty$  then  $c_1 = 0$ .

It follows that if two models have the same  $(\nu, \rho)$ , 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:

**PROPOSITION 9.** 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  $\nu$ , the function  $\hat{S}(\cdot, \nu)$  is increasing in  $\rho$ , and  $\hat{K}(\cdot, \nu)$  is decreasing in  $\rho$ .

Using this proposition we can fix  $s$ , say to  $s = 1$ , or  $X = \infty$ , making  $Kurt(\Delta p)$  a function of  $\nu$  only. Alternatively, fixing  $\nu$  makes  $Kurt(\Delta p)$  a function of the fraction of price changes strictly between barriers,  $s$ .

Figure 4: Kurtosis of power hazard function as  $\nu$  and  $s$  vary



The left panel of **Figure 4** displays the value of  $Kurt(\Delta p)$  for the case where  $X = \infty$  as a function of  $\nu$ . Note that Kurtosis goes from 6, corresponding to  $\nu = 0$ , or pure Calvo, to a value of 1, corresponding to  $\nu \rightarrow \infty$  which approximates Golosov and Lucas. Note that for  $\nu = 2$ , the quadratic case,  $Kurt(\Delta p) \approx 1.75$ .

The right panel of **Figure 4** 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 this relationship for three values of  $\nu$ . Fixing  $\nu$ , we vary  $\rho$  to obtain a larger share of price changes strictly inside the barriers  $s$ . This illustrates 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 lines show that a higher curvature  $\nu$  corresponds to a lower  $Kurt(\Delta p)$ .

## 6 Extending the setup to a non-negligible inflation

In this section we introduce a constant drift  $\mu$  into the price gap faced by the firm, so that in the absence of control the price gap evolves as  $dx = \mu dt + \sigma dW$ , where  $\mu$  represents (the negative of) the value of the steady state inflation. In this case the solution of the firm problem, i.e. its value function, the adjustment thresholds, and the generalized hazard  $\Lambda$  are no longer symmetric around  $x = 0$ . Below we analyze the effect of the non-zero drift  $\mu$  on the recovery of the distribution of  $G$  using  $\Lambda$ , and also on the recovery of  $\Lambda$  using the distribution of price changes  $q$  and  $Q$ .

**The firm's problem.** The value function for the problem with drift  $\mu$  is:

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

The presence of the drift  $\mu$  has several implications for the decision rules. Denote the optimal return point by  $x^*$ , so that  $x^* = \arg \min_x v(x)$ . It is no longer the case that  $x^* = 0$ , meaning that  $x^*$  does not coincide with the minimum of the flow cost function. Furthermore,  $v$  and  $\Lambda$  are no longer symmetric in  $x$  around  $x = x^*$ , and in general we need to distinguish the upper and lower bound of the range of inaction where it is not optimal to pay the fixed cost  $\Psi$ . We label the upper and lower bound of such region by  $\bar{x}$  and  $\underline{x}$  respectively. We no longer have  $X = \bar{x} = -\underline{x}$ , and in general  $\bar{x} - x^* \neq x^* - \underline{x}$ . The generalized hazard function  $\Lambda : (\underline{x}, \bar{x}) \rightarrow \mathbb{R}_+$  satisfies, as with no drift,  $\Lambda(x) = \kappa G(v(x) - v(x^*))$  in its domain. One can easily show that the value function  $v(\cdot)$  and the generalized hazard function  $\Lambda(\cdot)$  are both increasing in  $x$  for  $x \in (x^*, \bar{x})$  and decreasing in  $x$  for  $x \in (\underline{x}, x^*)$ , with  $v$  strictly so in each of the intervals.

**Rationalizing  $\Lambda$  in the presence of a drift.** We analyze the generalization of [Theorem 1](#) for  $\mu \neq 0$ . In this case it is not enough to have three arbitrary numbers  $\underline{x} < x^* < \bar{x}$  a bounded function  $\Lambda : [\underline{x}, \bar{x}] \rightarrow \mathbb{R}_+$  increasing in  $[x^*, \bar{x}]$  and decreasing in  $[\underline{x}, x^*]$  to recover  $v$  and  $G$ . Instead, we can construct a unique candidate value function, and then check the key monotonicity property for this candidate. If the property holds, then we can rationalize  $\Lambda$ , the associated thresholds  $\underline{x} < x^* < \bar{x}$ ,

and obtain the unique distribution  $G$  for the random menu cost.

We first present the generalization of [Lemma 1](#) to the presence of a drift. The main difference with the driftless case is that now  $u$  is defined in the entire domain of  $\Lambda$  instead of  $[0, \bar{x}]$ .

**LEMMA 2.** Take a positive function  $\Lambda : [\underline{x}, \bar{x}] \rightarrow \mathbb{R}_+$ . Let the function  $u$  solve the linear ordinary differential equation

$$[r + \Lambda(x)] u(x) = 2Bx + \mu u'(x) + \frac{\sigma^2}{2} u''(x) \text{ for } x \in [\underline{x}, \bar{x}] \quad (38)$$

with boundary conditions  $u(\underline{x}) = u(\bar{x}) = 0$ . The solution  $u$  exists and it is unique. If  $\Lambda$  is the GHF for the problem with value function  $v$  we have

$$v(x) = \frac{B}{r} (x^*)^2 + \frac{\sigma^2}{2r} u'(x^*) + \int_{x^*}^x u(z) dz \text{ for } x \in [\underline{x}, \bar{x}] \quad (39)$$

Now we are ready to present the natural extension of [Theorem 1](#). The main substantive difference is that it does not suffice that  $\Lambda$  be symmetric when  $\mu \neq 0$ . Indeed the “test” to see if  $\Lambda$  can be rationalized or not is whether the solution to [equation \(38\)](#) has the correct monotonicity properties, given by [equation \(40\)](#), which is set in terms of  $U$ , the integral of  $u$ .

**THEOREM 4.** Fix the discount rate  $r > 0$ , the curvature of the profit function  $B > 0$ , the volatility of shocks  $\sigma > 0$ , and the three thresholds  $\underline{x} < x^* < \bar{x}$ , where either both  $\bar{x}$  and  $\underline{x}$  are finite, or  $\underline{x} = -\bar{x} = \infty$ . Consider a function  $\Lambda(\cdot) : (\underline{x}, \bar{x}) \rightarrow \mathbb{R}_+$  that is differentiable, bounded, increasing on  $(x^*, \bar{x})$ , decreasing on  $(\underline{x}, x^*)$ , and with  $\lim_{x \rightarrow \bar{x}} \Lambda(x) = \lim_{x \rightarrow \underline{x}} \Lambda(x)$ . Let  $u$  be the unique solution of [equation \(38\)](#) in the domain  $[\underline{x}, \bar{x}]$ , with boundary conditions  $u(\bar{x}) = u(\underline{x}) = 0$ . Define  $U(x) \equiv \int_{x^*}^x u(z) dz$  for all  $x \in [\underline{x}, \bar{x}]$ . There exist positive real numbers  $\{\kappa, \Psi\}$  and a cost distribution  $G(\cdot) : [0, \Psi] \rightarrow [0, 1]$  with a density  $g(\cdot)$ , continuous on  $(0, \Psi)$ , and possibly a mass point  $G(0) > 0$ , that rationalizes  $\Lambda$  with a value function that solves [equation \(3\)](#) if and only if

$$U(\bar{x}) = U(\underline{x}), u(x) \leq 0 \text{ for all } x \in (\underline{x}, x^*), \text{ and } u(x) \geq 0 \text{ for all } x \in (x^*, \bar{x}) \quad (40)$$

When these conditions are met we have:

$$\kappa = \lim_{x \uparrow \bar{x}} \Lambda(x) \quad , \quad \Psi = U(\bar{x}) \quad \text{and} \quad , \quad G(0) = \frac{\Lambda(0)}{\kappa} \quad (41)$$

$$g(U(x)) = \frac{\Lambda'(x)}{u(x)\kappa} \quad \text{for all } x \in (\underline{x}, \bar{x}) \quad (42)$$

The practical use of this result is that it allows to check if  $\{\underline{x}, x^*, \bar{x}, \Lambda\}$  constitute a valid optimal policy for the problem described by  $\{B, r, \mu, \sigma^2\}$  and *some* distribution of random fixed cost  $G$  with *some* fixed cost  $\Psi$ . To find out if they do, it is necessary and sufficient to solve the linear o.d.e. in [equation \(38\)](#) and check the monotonicity properties [equation \(40\)](#) from its solution.

## 6.1 From price changes to the GHF ( $\Lambda$ ) in the presence of a drift

In this section we extend [Theorem 2](#) recovering  $f$  and  $\Lambda$  from a distribution of price changes with density  $q$  and CDF  $Q$  for the case where the price gaps have a drift  $\mu$ . The logic is the same as before, but the expressions for the density of the invariant distribution  $f$  and the generalized hazard function  $\Lambda$  are more cumbersome than the in the case with zero drift.

The first step is to write the Kolmogorov Forward equation for the invariant distribution in the case where the price gap as a drift  $\mu$ , which becomes:

$$0 = -\mu f'(x) - \Lambda(x)f(x) + \frac{\sigma^2}{2} f''(x) \quad \text{for all } x \neq x^* \quad \text{and } x \in [\underline{x}, \bar{x}] \quad (43)$$

$$0 = f(\underline{x}) = f(\bar{x}) \quad \text{and} \quad \int_{\underline{x}}^{\bar{x}} f(x) dx = 1. \quad (44)$$

The distribution of price changes satisfies the straightforward generalization of the case with no drift, i.e. it has density  $q$  and two mass points at its extremes:

$$q(x^* - x) = \frac{f(x)\Lambda(x)}{N_a} \quad \text{for all } x \in [\underline{x}, \bar{x}]$$

$$dQ(x^* - \bar{x}) = -\frac{\sigma^2}{2} f'(\bar{x})/N_a \quad , \quad dQ(x^* - \underline{x}) = +\frac{\sigma^2}{2} f'(\underline{x})/N_a$$

where the number of price changes satisfies:

$$N_a = \int_{\underline{x}}^{\bar{x}} \Lambda(x)f(x)dx + \frac{\sigma^2}{2}f'(\underline{x}) - \frac{\sigma^2}{2}f'(\bar{x})$$

Suppose we have the number of price changes  $N_a$  as well as the distribution of price changes  $Q$ . We will first obtain the value of  $\mu$  and  $\sigma^2$ , and using them and  $Q$  we obtain explicit expressions for the invariant density  $f$  and for the generalized hazard function  $\Lambda$  with  $\{\underline{x}, x^*, \bar{x}\}$ .

Importantly, the thresholds  $\underline{x}$  and  $\bar{x}$  together with the resetting point  $x^*$  are not readily scanned from the data, because price gaps are not observable. Instead, the data directly show price changes, including the extreme values  $\bar{x} - x^*$  and  $x^* - \underline{x}$ . For this reason, we break the procedure that recovers  $\{f, \Lambda, \underline{x}, x^*, \bar{x}\}$  in two steps. First, we show how to recover the re-centered distribution of price gaps  $\tilde{f}(y) = f(y + x^*)/N_a \cdot \sigma^2/2$  supported on  $(\underline{x} - x^*, \bar{x} - x^*)$  and the re-centered hazard  $\tilde{\Lambda}(y)$  defined by  $\Lambda(y + x^*) = \tilde{\Lambda}(y)$  defined on the same interval. Second, we briefly discuss the recovery of  $x^*$  that leads to  $\{\underline{x}, \bar{x}, f, \Lambda\}$ .

**THEOREM 5.** Let  $Q$  be the CDF of the price changes, with density  $q$  in its interior, and let  $N_a$  be the frequency of price changes. Assume  $Q$  and  $N_a$  have been generated by the solution to a firm's problem for an uncontrolled price gap process with parameters  $\mu$  and  $\sigma^2$  with  $\mu \neq 0$  and  $\sigma^2 > 0$ . Let the thresholds  $\underline{x} < x^* < \bar{x}$ , the generalized hazard function  $\Lambda$ , and the distribution of price gaps  $f$  also be generated by the solution to this problem. Then,

$$\mu = -\mathbb{E}[\Delta p]N_a, \text{ and } \sigma^2 \text{ is the solution to } M_{\Delta p} \left( \frac{2\mu}{\sigma^2} \right) = 1 \quad (45)$$

where  $M_{\Delta p}(\varphi) \equiv \int \exp(\varphi z)dQ(z)$  is the moment generating function of  $\Delta p$ . Define  $\phi = 2\mu/\sigma^2$  and  $R(z_1, z_2) \equiv \int_{z_1}^{z_2} \exp(\phi z)dQ(z)$ . The recentered and rescaled distribution of price gaps defined by

$\tilde{f}(y) \equiv f(y + x^*)/N_a \cdot (\sigma^2/2)$  is

$$\tilde{f}(y) = \begin{cases} -\int_0^{\bar{x}-x^*} e^{\phi z} R(-z, 0) dz - \frac{\bar{a}}{\phi} [e^{\phi(\bar{x}-x^*)} - 1] & \text{if } y = 0 \\ \tilde{f}(0) + \int_y^0 e^{\phi z} R(0, -z) dz + \frac{\underline{a}}{\phi} [e^{\phi y} - 1] & \text{if } y \in [\underline{x} - x^*, 0) \\ \tilde{f}(0) + \int_0^y e^{\phi z} R(-z, 0) dz + \frac{\bar{a}}{\phi} [e^{\phi y} - 1] & \text{if } y \in (0, \bar{x} - x^*] \end{cases}$$

where the constants  $\bar{a} \equiv \tilde{f}'(0+) < 0$ , and  $\underline{a} \equiv \tilde{f}'(0-) > 0$  are

$$\bar{a} = \frac{\phi \left[ \int_{\underline{x}-x^*}^0 e^{\phi z} R(0, -z) dz - \int_0^{\bar{x}-x^*} e^{\phi z} R(-z, 0) dz \right] + e^{\phi(\underline{x}-x^*)} - 1}{e^{\phi(\bar{x}-x^*)} - e^{\phi(\underline{x}-x^*)}} \quad \text{and} \quad \underline{a} = 1 + \bar{a}$$

Finally,  $\tilde{\Lambda}(y) = q(-y)/\tilde{f}(y) \cdot (2/\sigma^2)$  for  $y \in (\underline{x} - x^*, \bar{x} - x^*)$  and  $y \neq 0$ .

Now armed with  $\Lambda(y + x^*) = \tilde{\Lambda}(y)$ , one can recover  $x^*$  and consequently  $\{\underline{x}, \bar{x}, f, \Lambda\}$ . One way to do it is to guess a value for  $x^*$  and use [equation \(38\)](#) to solve for the function  $u$  with boundary conditions  $u(\bar{x}) = u(\underline{x}) = 0$ :

$$[r + \Lambda(x - x^*)]u(x) = 2Bx + \mu u'(x) + \frac{\sigma^2}{2} u''(x) \quad (46)$$

The guess is correct if  $u(x^*) = 0$ . If  $u(x^*) \neq 0$ , the guess has to be updated.

Finally, to assess if the setup with  $\mu = 0$  is a good approximation for economies with low but non-zero inflation, we show the following:

**PROPOSITION 10.** Suppose the distribution of price changes  $Q(x; \mu)$  and adjustment frequency  $N_a(\mu)$  are generated by the solution to a firm's problem with thresholds  $\underline{x} < x^* < \bar{x}$  and generalized hazard function  $\Lambda : (\underline{x}, \bar{x}) \rightarrow \mathbb{R}_+$  for an uncontrolled price gap process with parameters  $\mu$  and  $\sigma^2$  with  $\mu \neq 0$  and  $\sigma^2 > 0$ . Suppose also they are analytical functions of  $\mu$ . Then,

$$Kurt(\Delta p; \mu) = Kurt(\Delta p; 0) + o(\mu^2) \quad (47)$$

$$N_a(\mu) = N_a(0) + o(\mu^2) \quad (48)$$



In words, the setup with no inflation provides approximations for  $Kurt(\Delta p)$  and  $N_a$  that work up to the second order. The odd powers in the expansion around zero inflation disappear because of the symmetry. The cumulative impulse response to a monetary shock in case  $\mu = 0$  is therefore a good approximation for economies with small inflation.

## 7 Scope, limitations and relation to the literature

This section discusses the scope and limitations of our results, and the implications for selected contributions that have analyzed related questions in models with sticky prices.

**Results with a non-negligible drift (inflation).** It was shown in [Section 6](#) that the possibility of establishing an inverse mapping is not affected by the presence of drift (Theorems 4 and 5 and the version with drift of Theorems 2 and 3). This is encouraging and suggests that the scheme might be fruitfully applied to similar problems with Ss behavior, such as the optimal capital investment problems studied by [Guiso, Lai, and Nirei \(2017\)](#); [Caballero and Engel \(1999\)](#); [Baley and Blanco \(forthcoming\)](#). Instead, the sufficient statistic result, as we review next, requires that inflation to be close to zero.

**Main assumptions and limitations of the sufficient statistic result.** Four assumptions are key for the sufficient statistic result of [Theorem 3](#). The first one is that the model has no inflation, so that several model objects display symmetry properties. While the assumption of zero inflation might seem restrictive, we showed in [Proposition 10](#) that it provides a good approximation to models where inflation is low. Empirical evidence consistent with the insensitivity of firms' decision with respect to inflation in a neighborhood of zero inflation, is presented in [Alvarez et al. \(2019\)](#) using data from Argentina and in [Nakamura et al. \(2018\)](#) using data from the US. In both cases the evidence for insensitivity covers inflation rates at least as large as 3 per cent per year, and even higher. The second key assumption for the result to hold is that upon adjustment the firm completely closes the price gap, i.e. that  $x$  is reset to zero. This assumption is violated in models

with strategic complementarities, or in models with price plans (as in [Eichenbaum, Jaimovich, and Rebelo \(2011\)](#)). In such cases [Theorem 3](#) is not an accurate summary of the impulse response and other methods must be used to approximate the cumulated output response (See [Alvarez and Lippi \(2019\)](#) and [Alexandrov \(2020\)](#) for some results on, respectively, price plans and large shocks).<sup>23</sup> A third assumption is that  $x$  follows a Brownian motion. This assumption allows us to use the Kolmogorov forward equation in the proof. In a model with leptokurtic shocks, such as [Midrigan \(2011\)](#), the proposition fails to hold. However, for moderate degrees of leptokurtic shocks, consistent with the data on the distribution of firms' nominal shocks, the proposition provides an informative benchmark (see Section 5 in [Alvarez, Le Bihan, and Lippi \(2016\)](#) and the numerical results in [Gautier and Le Bihan \(Forthcoming\)](#)). Finally, the fourth assumption is that the monetary shock is described as a small once and for all change in the common nominal cost. For instance, if the shock will be large enough, in a simple sS model a large fraction of firms will adjust on impact. Also, for large monetary shocks, the effect of steady state inflation may not be second order. See [Alexandrov \(2020\)](#) for a theoretical analysis, and [Alvarez and Neumeyer \(2020\)](#) for simulations and empirical evidence on the effect of non-negligible steady state inflation rates (25 per cent per year and larger) and/or large once and for all shocks.

**Aggregation across heterogenous firms.** We briefly discuss how [Theorem 3](#) 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, applying the Theorem's result to each group and aggregating, we obtain that the area under the IRF of *aggregate* output for a small monetary shock  $\delta$  is

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

---

<sup>23</sup>Nevertheless, the results in [Alvarez, Lippi, and Souganidis \(2021\)](#) show that with strategic complementarity the *ratio* of the CIR across models with and without strategic complementarities is the same, and hence equal to the one given by the ration of Kurtosis to Frequency.

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 11](#) in [Appendix C](#), this average is also different from the kurtosis of the pooled data. This result can be seen as an instance of Jensen’s inequality which is important when applying the theorem to an economy with heterogenous sectors. This applies even if all the groups have the same kurtosis.<sup>24</sup> However, if groups are heterogenous in duration (or expenditures), then the kurtosis 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 \(2020\)](#). The formula in [equation \(49\)](#) implies that the flexible price firms must be given a zero weight in the computation of  $\mathcal{M}$ , which is computed only on the mass of firms with sticky prices. Notice that this is very different from computing the  $\mathcal{M}$  as the ratio of the cross-sectional average kurtosis and the average frequency: the cumulated impulse response computed using “average data” would be zero, while the true one is not.

**Empirical tests of the sufficient static result.** The robustness of [Theorem 3](#), which holds in a broad variety of different models, has motivated us to pursue an empirical test of this theoretical prediction. Ongoing work in [Alvarez, Ferrara, Gautier, Le Bihan, and Lippi \(2021\)](#) uses data for a large number of firms, representative of the French economy, as well as the micro data to measure the cross sectional moments (frequency and kurtosis) for about 120 PPI industries and 220 CPI categories. We use a Factor Augmented VAR to measure the sectoral responses to a monetary shock, as summarized by the cumulative impulse response of sectoral prices (CIR), under three alternative identification schemes. The estimated CIR correlates with the kurtosis of price changes and the frequency of price adjustments consistently with the prediction of the theory (equal coefficients with opposite signs). The analysis also shows that other moments not suggested by the theory, such as the mean, standard deviation and skewness of the size-distribution of price

---

<sup>24</sup>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 \(2010\)](#).

changes, are not correlated with the CIR. A related analysis for the United States is presented in [Hong et al. \(2020\)](#), where the authors aim to unveil statistically significant correlations between sectoral producers price indices (24 month after the monetary shock) and a set of cross sectional moments of the distribution of price changes. In spite of the wording, which presents the empirical evidence as a rejection of the theory, we note that this evidence is not a test of the sufficient statistic proposition since the outcome variable is the *level response of prices* while theory concerns the *cumulated response of output*. To be clear, the dependent variable in the the regression in [Hong et al. \(2020\)](#) is *not* the one derived in the theory. Overall, while the data are potentially interesting, the bulk of that empirical analysis does not provide an explicit test of the theory.<sup>25</sup>

**Relation with Caballero Engel’s Flexibility Index.** [Caballero and Engel \(2007\)](#) introduced the notion of Flexibility Index ( $\mathcal{F}$ ), subsequently used in several studies such as [Berger and Vavra \(2018\)](#), as an inverse measure of monetary non-neutrality. In [Appendix D](#) we define  $\mathcal{F}$  in terms of the model, using [Caballero and Engel’s \(2007\)](#) formula, and study its relationship with the cumulated impulse response that was used in [Section 5](#) as a measure of monetary non-neutrality. We show that  $\mathcal{F}$  measures the slope of the impulse response of prices right after a small monetary shock. 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 \(33\)](#). We use this simple model to display several examples where  $\mathcal{F}$  is *not* an accurate summary of the effect on output, neither of its cumulative response, nor of its short term response.

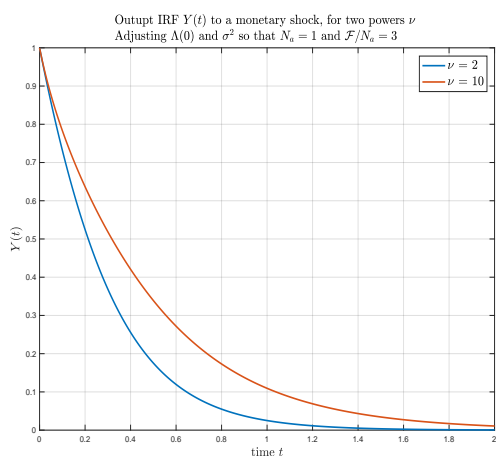
[Figure 5](#) displays several non-pathological examples derived for a model with a power hazard function  $\Lambda(x) = \Lambda(0) + \kappa x^\nu$  (see [Appendix D](#) for details) with different parameters  $\nu$  and  $\Lambda(0)$  chosen to produce a frequency  $N_a = 1$  and a constant flexibility index  $\mathcal{F} = 3$ . The left panel shows

---

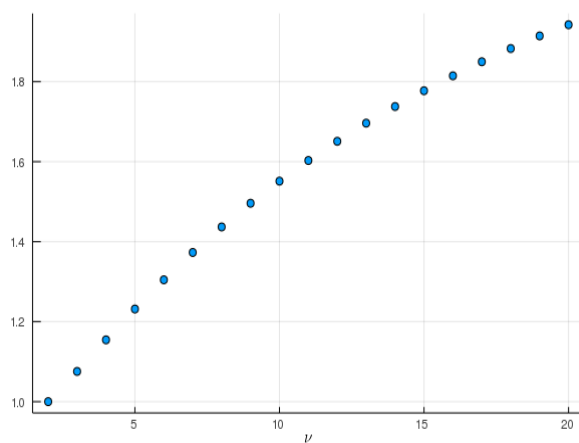
<sup>25</sup>In addition to the (wrong) units of the outcome variable, we note that several specifications used in that paper, such as those where when kurtosis is used as the only regressor, are inconsistent with the theory that prescribes that *both* kurtosis and frequency are part of the specification since they are in general not orthogonal (for instance both vary following a change in the distribution of adjustment costs). It is interesting that, in spite of all the wedges between the evidence and the theory, in the specifications where kurtosis and frequency are both entered as regressors the sign of the estimated coefficients are consistent with the sufficient statistic proposition and often statistically significant (see e.g. column 1 and 6 of Table 1 in their Section VIII and the columns of Table 12).

Figure 5: Limitations of the Flexibility Index (all cases have  $\mathcal{F} = 3$  and  $N_a = 1$ )

IRF for two cases with same  $\mathcal{F}$  and  $N_a$



Values of  $Kurt(\Delta p)$  relative to case with  $\nu = 2$



two cases with the same initial slope  $\mathcal{F}$  of the impulse response, producing a cumulative impulse response that differ markedly, even in the short run: for instance, the ratio of the two IRF evaluated at the expected average duration of prices i.e. at  $t = 1$  is higher than 4. The right panel shows several cases producing a constant flexibility index, also equal to 3, with substantially different values of the cumulated output response that doubles as  $\nu$  increases (plotted on the vertical axis relative to the case where  $\nu = 2$ ).

## 8 Concluding remarks and future research

This paper discussed the foundations of the generalized hazard function, a flexible modeling block used in several sticky-price setups, and provided a mapping to the primitives underlying the firm's optimization problem.

Our first contribution was to show how to identify the generalized hazard function from observable objects, such as the distribution of the size of price changes, by means of simple closed form equations. In [Appendix E](#) we complement these results by establishing that, under regularity conditions, the survival function  $S(t)$ , measuring the distribution of price *durations*, uniquely

identifies  $\Lambda(x)$  as well.<sup>26</sup> We illustrate this result with an application to US micro pricing data. Our empirical strategy accounts for unobserved heterogeneity across products, which we show to be important in estimating the Kurtosis of price changes.

The second main result is the analytic generalization of the “Kurtosis result” from [Alvarez, Le Bihan, and Lippi \(2016\)](#) to a considerably larger set of models. As described above, this result establishes a sufficient statistic to compute the cumulative output response to a small monetary shock: the ratio of the kurtosis of the steady-state distribution of price changes over the frequency of price adjustment. We also proved that the Calvo model yields the maximum amount of monetary non-neutrality within the class spanned by the random menu cost models, while the canonical menu cost developed by [Golosov and Lucas \(2007\)](#) yields the minimum effect.

Future developments concern both empirical applications as well as theoretical extensions. Ongoing research using sectoral PPI and CPI data for France, together with the micro moments underlying these sectors, reveals considerable empirical support of the informativeness of the sufficient statistic identified in [Theorem 3](#), see [Alvarez et al. \(2021\)](#). Also, our analytical setup can be extended to study multiproduct firms, as in [Midrigan \(2011\)](#); [Alvarez and Lippi \(2014\)](#), and to study the entire profile of the IRF (as opposed to the cumulative impulse response) using the methods explored in [Alvarez and Lippi \(2019\)](#).

## References

- Alexandrov, Andrey. 2020. “The Effects of Trend Inflation on Aggregate Dynamics and Monetary Stabilization.” Mimeo, University of Mannheim.
- 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, Andrea Ferrara, Erwan Gautier, Herve Le Bihan, and Francesco Lippi. 2021. “Empirical Investigation of a Sufficient Statistic for Monetary Policy Shocks.” Working paper 09, EIEF.
- Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis. 2021. “Price Setting with Strategic Complementarities as a Mean Field Game.” Tech. rep., University of Chicago, manuscript.

---

<sup>26</sup>Summarizing, the generalized hazard rate  $\Lambda$  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).

- Alvarez, Fernando and Andres Neumeier. 2020. “The Pass-through of Large Cost Shocks in an Inflationary Economy.” In *”Changing Inflation Dynamics, Evolving Monetary policy”*, edited by G. Castex, G. Gali and D. Saravia :7–48.
- 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. forthcoming. “Aggregate Dynamics in Lumpy Economies.” *Econometrica* .
- 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.
- Bils, Mark and Peter J. Klenow. 2004. “Some Evidence on the Importance of Sticky Prices.” *Journal of Political Economy* 112 (5):947–985.
- 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 and John Leahy. 1991. “State-Dependent Pricing and the Dynamics of Money and Output.” *The Quarterly Journal of Economics* 106 (3):683–708.
- 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.
- Dias, D.A., C. Robalo Marques, and J.M.C. Santos Silva. 2007. “Time- or state-dependent price setting rules? Evidence from micro data.” *European Economic Review* 51 (7):1589 – 1613.
- 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.
- Fougere, Denis, Herve Le Bihan, and Patrick Sevestre. 2007. “Heterogeneity in Consumer Price Stickiness.” *Journal of Business & Economic Statistics* 25 (3):247–264.
- Gautier, Erwan and Hervé Le Bihan. Forthcoming. “Shocks vs Menu Costs: Patterns of Price Rigidity in an Estimated Multi-Sector Menu-Cost Model.” *The Review of Economics and Statistics* – (–):–.
- 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.
- Guiso, Luigi, Chaoqun Lai, and Makoto Nirei. 2017. “An Empirical Study of Interaction-Based Aggregate Investment Fluctuations.” *The Japanese Economic Review* 68 (2):137–157.



- Hong, Gee Hee, Matthew Klepacz, Ernesto Pasten, and Raphael Schoenle. 2020. “The Real Effects of Monetary Shocks: Evidence from Micro Pricing Moments.” Working Papers Central Bank of Chile 875, Central Bank of Chile.
- 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.
- Midrigan, Virgiliu. 2011. “Menu Costs, Multi-Product Firms, and Aggregate Fluctuations.” *Econometrica*, 79 (4):1139–1180.
- Nakamura, Emi and Jon Steinsson. 2010. “Monetary Non-Neutrality in a Multisector Menu Cost Model.” *The Quarterly Journal of Economics* 125 (3):961–1013.
- Nakamura, Emi, Jon Steinsson, Patrick Sun, and Daniel Villar. 2018. “The Elusive Costs of Inflation: Price Dispersion during the U.S. Great Inflation\*.” *The Quarterly Journal of Economics* 133 (4):1933–1980.
- 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* .
- Sheshinski, Eytan and Yoram Weiss. 1977. “Inflation and Costs of Price Adjustment.” *Review of Economic Studies* 44 (2):287–303.
- Woodford, Michael. 2009. “Information-Constrained State-Dependent Pricing.” *Journal of Monetary Economics* 56:s100–s124.

## Appendix:

# The Macroeconomics of Sticky Prices with Generalized Hazard Functions

Fernando Alvarez, Francesco Lippi, and Aleksei Oskolkov

## A Proofs of the Theorems

**Proof.** (of [Lemma 1](#)). Define the function  $U(x) \equiv v(x) - v(0)$  and rewrite [equation \(3\)](#) as

$$rU(x) = Bx^2 + \frac{\sigma^2}{2}(U''(x) - v''(0)) - \kappa \int_0^{U(x)} G(\psi) d\psi \quad \text{for } x \in [0, X] \quad (50)$$

with boundary conditions  $U'(X) = 0$  and  $U(X) = \Psi$ . Note that by definition  $U(0) = 0$ . To obtain [equation \(50\)](#) we used integration by parts on the right hand side of [equation \(3\)](#):

$$\begin{aligned} \int_0^{U(x)} [\psi - U(x)] G'(\psi) d\psi &= G(\psi)\psi \Big|_0^{U(x)} - \int_0^{U(x)} G(\psi) d\psi - U(x) \int_0^{U(x)} G'(\psi) d\psi \\ &= G(\psi)\psi \Big|_0^{U(x)} - \int_0^{U(x)} G(\psi) d\psi - U(x) [G(U(x)) - G(0)] \\ &= - \int_0^{U(x)} G(\psi) d\psi + U(x)G(0) \end{aligned}$$

Next differentiate both sides of [equation \(50\)](#) with respect to  $x$  to obtain:

$$[r + \kappa G(U(x))] U'(x) = 2Bx + \frac{\sigma^2}{2} U'''(x) \quad \text{for } x \in [0, X] \quad (51)$$

with boundary conditions given by:  $U'(X) = 0$  and  $U'(0) = 0$ . The first boundary condition is smooth pasting. Note that if  $X = \infty$  we do not have smooth pasting, but since  $v$  is bounded above so is  $U$ , then it must be that  $\lim_{x \rightarrow \infty} U'(x) = 0$ , and hence the analogous boundary condition holds in the case where  $X$  is unbounded. The second boundary is implied by the symmetry and differentiability of  $v(\cdot)$ , and hence of  $U(\cdot)$ , around  $x = 0$ . Thus, solving for the value function in [equation \(3\)](#) is equivalent to solving for  $U(\cdot)$  in [equation \(51\)](#) with its corresponding boundary conditions.

Now define  $u(x) \equiv U'(x)$  and rewrite [equation \(51\)](#) using that  $\Lambda(x) = \kappa G(U(x))$ , by [equation \(2\)](#). This gives the o.d.e. in [equation \(4\)](#). The boundary conditions described above in terms of  $U'$  thus become  $u(X) = u(0) = 0$ .

**Uniqueness and invertibility.** Note that [equation \(4\)](#) is a linear second order ordinary differential equation of the Sturm-Liouville type with two Dirichlet boundary conditions, where we write:  $L(u)(x) \equiv [r + \Lambda(x)] u(x) - \frac{\sigma^2}{2} u''(x)$  and thus the equation above can be written as  $L(u)(x) = 2Bx$ .

The function  $\Lambda(\cdot)$  defining the operator  $L$  is continuous, so it has a unique solution  $u(\cdot)$ . To see this let  $L(u)(x) = 2Bx$  and let  $\{\theta_j, \varphi_j\}$  be the eigenvalues and orthonormal eigenfunctions of  $L$  satisfying the Dirichlet boundary conditions, i.e. solving  $L(\varphi_j) = \theta_j\varphi_j$  and with  $\varphi_j(0) = \varphi_j(X) = 0$ . By linearity we have  $L(\sum_j \alpha_j \varphi_j) = \sum_j \theta_j \alpha_j \varphi_j$  for any square integrable sequence  $\{\varphi_j(\cdot)\}$ . Then we can choose  $\{\alpha_j\}$  so that  $u(x) = \sum_j \theta_j \alpha_j \varphi_j(x)$ , with the equality in the  $L^2$  sense. In particular we can set  $\alpha_j = \langle \varphi_j, u \rangle / \theta_j$ . Again, the case of  $X = \infty$ , requires a slightly different argument for the existence of its solution. In particular, the existence of a solution is guaranteed by Theorem 3.1 in [Lian, Wang, and Ge \(2009\)](#). By the Maximum principle then  $u(x) > 0$  since  $2Bx > 0$  in  $(0, X)$ . Since  $u > 0$  then  $U$  is increasing and thus it is invertible.

**Value function.** We construct  $v(\cdot)$  as follows. Recall  $u = U'$  and  $U(0) = 0$ , we have

$$U(x) = \int_0^x u(z)dz \text{ for all } x \in [0, X] \quad \text{and} \quad \Psi = U(X) .$$

From the definition of  $U(x) = v(x) - v(0)$  and [equation \(3\)](#) we have

$$v''(0) = U''(0) = u'(0) \text{ and } rv(0) = v''(0) \frac{\sigma^2}{2} \text{ so } v(0) = u'(0) \frac{\sigma^2}{2r}$$

which gives [equation \(5\)](#) in the lemma. Note that  $v(\cdot)$  is increasing because  $u(x) > 0$  on  $(0, X)$  as established above.  $\square$

**Proof.** (of [Theorem 1](#)). We now construct the fixed cost  $\Psi$ , the Poisson arrival rate  $\kappa$ , the value of  $G(0)$  and the density  $G'(\cdot)$  that rationalize the generalized hazard rate  $\Lambda(\cdot)$  using the function  $u(\cdot)$ . We use [equation \(2\)](#),  $\Lambda(x) = \kappa G(U(x))$  for all  $x \in [0, X]$ , which evaluated at  $x = 0$  implies  $\Lambda(0) = \kappa G(0)$ . Denote by  $w(\cdot) \equiv U^{-1}(\cdot)$ , the inverse function of  $U(\cdot)$ , mapping  $[0, \Psi]$  onto  $[0, X]$ . Set  $\kappa$  to be  $\kappa = \Lambda(X)$  to ensure that  $G(\Psi) = 1$ . Differentiating the expression above with respect to  $x$ , we have  $G'(U(x))U'(x) = \frac{\Lambda'(x)}{\Lambda(X)}$  for all  $x \in (0, X)$  and thus

$$G'(\psi) = G'(U(w(\psi))) = \frac{\Lambda'(w(\psi))}{u(w(\psi))\Lambda(X)} = \frac{\Lambda'(w(\psi))}{u(w(\psi))\Lambda(X)} \text{ for all } \psi \in (0, \Psi)$$

which gives the density of  $G'$  in terms of the function  $u$  defined in [Lemma 1](#).  $\square$

**Proof.** (of [Theorem 2](#)) Without loss of generality, given the assumed symmetry, let  $q(\cdot)$  be the density of minus price changes, so that  $q(x)N_a = \Lambda(x)f(x)$ . Denote the minus price changes by  $\Delta p$ . We will use four equations for  $x > 0$ :

$$\begin{aligned} f''(x) &= \frac{2}{\sigma^2} q(x) N_a \\ f'(x) &= f'(X) - \int_x^X f''(t) dt \\ f(x) &= - \int_x^X f'(t) dt \\ \sigma^2 &= N_a \text{Var}(\Delta p) \end{aligned}$$

where we have used that  $f(X) = 0$ . Combining the first and the second equation we have ,

$$\begin{aligned} f'(x) &= f'(X) - \frac{2}{\sigma^2} N_a \int_x^X q(x) dx = f'(X) - \frac{2}{\sigma^2} N_a \left( 1 + f'(X) \frac{\sigma^2}{2N_a} - Q(x) \right) \\ &= \frac{2}{\sigma^2} N_a (Q(x) - 1) \end{aligned}$$

where we have used that  $\lim_{x \rightarrow X} Q(x) \rightarrow 1 + f'(X) \frac{\sigma^2}{2N_a}$  as  $x \rightarrow X$ . Integrating further,

$$f(x) = \frac{2}{\sigma^2} N_a \int_x^\infty (1 - Q(t)) dt$$

Now using the last equation,

$$f(x) = \frac{2}{\text{Var}(\Delta p)} \int_x^\infty (1 - Q(t)) dt$$

Using the identity  $q(x)N_a = \Lambda(x)\bar{p}(x)$  once again,

$$\Lambda(x) = \frac{N_a \text{Var}(\Delta p)}{2} \frac{q(x)}{\int_x^\infty (1 - Q(t)) dt}$$

Finally, we check whether  $\Lambda(X) = \kappa < \infty$ . If  $X < \infty$ , then using L'Hopital we get

$$\Lambda(X) = \frac{N_a \text{Var}(\Delta p)}{2} \frac{q'(X)}{-f'(X) \frac{\sigma^2}{2}} < \infty$$

If  $X = \infty$ , we apply L'Hopital rule twice, since  $q(x) \rightarrow 0$  and  $Q(x) \rightarrow 1$  as  $x \rightarrow \infty$ . We obtain:

$$\Lambda(X) = \frac{N_a \text{Var}(\Delta p)}{2} \lim_{x \rightarrow \infty} \frac{q''(x)}{q(x)}$$

which is finite given our assumption on the tail of  $q$ . This completes the proof.  $\square$

**Proof.** (of [Theorem 3](#)). First note that the identity in [equation \(13\)](#),  $N \cdot \text{Var} = \sigma^2$ , holds in the model. Let  $x(0) = 0$ . Consider the process  $z(t) \equiv x(t)^2 - \sigma^2 t$  for  $t \geq 0$ . Using Ito's lemma we can verify that the drift of  $x^2$  is  $\sigma^2$ , and hence  $z(t)$  is a Martingale. Let  $\tau$  be a stopping time, i.e. an instant where a price adjustment occurs (anywhere in the state space, including the boundaries), so that  $x$  is reset at  $x(0) = 0$ . By the optional sampling theorem  $z(\tau)$ , the process stopped at  $\tau$ , is also a martingale. Then  $\mathbb{E} [z(\tau) \mid x(0)] = \mathbb{E} [x(\tau)^2 \mid x(0)] - \sigma^2 \mathbb{E} [\tau \mid x(0)] = x(0) = 0$ . Since  $N = 1/\mathbb{E} [\tau \mid x(0)]$  and  $\text{Var} = \mathbb{E} [x(\tau)^2 \mid x(0)]$  we get the identity in [equation \(13\)](#).

For simplicity, we focus next on the case with unbounded support  $\bar{X} \rightarrow \infty$  (the logic for the case with bounded support is identical but the equations are slightly more cumbersome). Using the definition of the density of price changes in [equation \(15\)](#) we can rewrite the identity as

$$\int_{-\infty}^{\infty} x^2 \Lambda(x) f(x) dx = \sigma^2 \tag{52}$$

it is then straightforward to write the formula for kurtosis over  $6N_a$  as:

$$\frac{Kur}{6N_a} = \frac{\int_{-\infty}^{\infty} x^4 \Lambda(x) f(x) dx}{6 \left( \int_{-\infty}^{\infty} x^2 \Lambda(x) f(x) dx \right)^2} = \frac{\int_{-\infty}^{\infty} x^4 \Lambda(x) f(x) dx}{6\sigma^4}$$

where the last passage uses [equation \(52\)](#). Using the Kolmogorov forward equation,

$$\int_{-\infty}^{\infty} x^4 \Lambda(x) f(x) dx = \frac{\sigma^2}{2} \int_{-\infty}^{\infty} x^4 f''(x) dx$$

Integrating by parts twice gives

$$\int_{-\infty}^{\infty} x^4 \Lambda(x) f(x) dx = 6\sigma^2 \int_{-\infty}^{\infty} x^2 f(x) dx$$

This allows us to write

$$\frac{Kur}{6N_a} = \frac{\int_{-\infty}^{\infty} x^2 f(x) dx}{\sigma^2} \tag{53}$$

Recall that we have a system of two equations:

$$\Lambda(x)f(x) = \frac{\sigma^2}{2} f''(x) \quad , \quad \Lambda(x)m(x) = \frac{\sigma^2}{2} m''(x) - x$$

Eliminate  $\Lambda$  to get:

$$\frac{\sigma^2}{2} \frac{m(x)f''(x)}{f(x)} = -x + \frac{\sigma^2}{2} m''(x)$$

Multiply both sides by  $f(x)x$  and rearrange:

$$\frac{\sigma^2}{2} [m(x)f''(x) - m''(x)f(x)]x = -x^2 f(x)$$

Integrate both sides from 0 to  $\infty$ :

$$\frac{\sigma^2}{2} \int_0^{\infty} [m(x)f''(x) - m''(x)f(x)]x dx = - \int_0^{\infty} x^2 f(x) dx$$

Perform integration by parts in the left-hand side using the fact that  $[m(x)f'(x) - m'(x)f(x)]' = m(x)f''(x) - m''(x)f(x)$ :

$$\begin{aligned} \frac{\sigma^2}{2} \int_0^{\infty} [m(x)f''(x) - m''(x)f(x)]x dx &= \frac{\sigma^2}{2} \left( [m(x)f'(x) - m'(x)f(x)]x \Big|_0^{\infty} - \int_0^{\infty} [m(x)f'(x) - m'(x)f(x)]dx \right) \\ &= -\sigma^2 \int_0^{\infty} m(x)f'(x) dx \end{aligned}$$

where the last equality uses integration by parts again. We used  $\mathbb{E}[m(x)] < \infty$  and  $m(\cdot)$  being almost linear at infinity to justify setting  $f'(x)m(x)x$  and  $f(x)m'(x)x$  at infinity to 0. Hence, we have

$$\sigma^2 \int_0^\infty m(x)f'(x)dx = \int_0^\infty x^2 f(x)dx$$

Plugging this result in [equation \(53\)](#) we have

$$\frac{Kur}{6N_a} = \int_{-\infty}^\infty m(x)f'(x)dx$$

It follows from the definition of  $\mathcal{M}(\delta)$  in [equation \(33\)](#) that the right hand side is the first derivative of the CIR with respect to  $\delta$ , evaluated at  $\delta = 0$ , or  $\mathcal{M}'(\delta)$ . This completes the proof.  $\square$

**Proof.** (of [Theorem 4](#)) The proof follows the same steps used for [Theorem 1](#).

**Proof.** (of [Theorem 5](#)) The proof will proceed in three steps.

*Step 1.* As a first step we show how to recover the value  $\mu$  and  $\sigma^2$  based on the distribution of  $\Delta p$ . The first moment of  $\Delta p$ , together with the frequency  $N$  gives  $\mu$  from  $\mathbb{E}[\Delta p] = -\mu N_a$ . To see this, let  $\tau$  the stopping time at which prices are changed, so we have  $\Delta p = x^* - x(\tau)$  with  $x(0) = x^*$  and  $dx = \mu dt + \sigma dW$ . Let  $y(t) = x^* - x(t) + \mu t$ . Note that  $y$  is a Martingale, and thus  $\mathbb{E}[y(\tau)] = 0 = \mathbb{E}[x^* - x(\tau)] + \mu \mathbb{E}[\tau]$  or  $N_a \mathbb{E}[\Delta p] = -\mu$ . To obtain  $\sigma^2$  we prove that the moment generating function of the distribution of price changes must equal one when evaluated at  $2\mu/\sigma^2$ , i.e.:

$$M_{\Delta p} \left( \frac{2\mu}{\sigma^2} \right) = \int e^{\frac{2\mu}{\sigma^2} \Delta p} dQ(\Delta p) = 1$$

where  $M_{\Delta p}(\cdot)$  is the moment generating function of the distribution of prices. To see why this has to be the case, we first define  $F_n$  as:

$$F_n \equiv \frac{1}{N_a} \left[ \int_{\underline{x}}^{\bar{x}} \frac{\sigma^2}{2} f''(x)(x^* - x)^n dx - \frac{\sigma^2}{2} f'(x)(x^* - x)^n \Big|_{\underline{x}} \right]$$

Using the KFE equation, we get

$$\begin{aligned} F_n &= \frac{1}{N_a} \int_{\underline{x}}^{\bar{x}} [\Lambda(x)f(x) + \mu f'(x)] (x^* - x)^n dx - \frac{\sigma^2}{2N_a} f'(x)(x^* - x)^n \Big|_{\underline{x}} \\ &= \int_{\underline{x}}^{\bar{x}} (x^* - x)^n q(x^* - x) dx - \frac{\sigma^2}{2N_a} f'(x)(x^* - x)^n \Big|_{\underline{x}} + \frac{1}{N_a} \mu \int_{\underline{x}}^{\bar{x}} f'(x)(x^* - x)^n dx \end{aligned}$$

Using the definition of  $q$  and integrating by parts:

$$\begin{aligned} F_n &= \int_{\underline{x}}^{\bar{x}} (x^* - x)^n q(x^* - x) dx - \frac{\sigma^2}{2N_a} f'(x)(x^* - x)^n \Big|_{\underline{x}} \\ &\quad + \frac{2\mu}{\sigma^2} \frac{1}{n+1} \frac{1}{N_a} \left[ \int_{\underline{x}}^{\bar{x}} \frac{\sigma^2}{2} f''(x)(x^* - x)^{n+1} dx - \frac{\sigma^2}{2} f'(x)(x^* - x)^{n+1} \Big|_{\underline{x}} \right] \end{aligned}$$

This implies

$$F_n = \mathbb{E}[\Delta p^n] + \frac{2\mu}{\sigma^2} \frac{1}{n+1} F_{n+1}$$

with

$$\begin{aligned} F_1 &= \frac{1}{N_a} \left[ \int_{\underline{x}}^{\bar{x}} \frac{\sigma^2}{2} (x^* - x) f''(x) dx - \frac{\sigma^2}{2} f'(x)(x^* - x) \Big|_{\underline{x}}^{\bar{x}} \right] \\ &= \frac{1}{N_a} \left[ \int_{\underline{x}}^{\bar{x}} \frac{\sigma^2}{2} f'(x) dx + \frac{\sigma^2}{2} (x^* - x) f'(x) \Big|_{\underline{x}}^{\bar{x}} - \frac{\sigma^2}{2} f'(x)(x^* - x) \Big|_{\underline{x}}^{\bar{x}} \right] \\ &= \frac{\sigma^2}{2N_a} f(x) \Big|_{\underline{x}}^{\bar{x}} = 0 \end{aligned}$$

Iterating the recursive expression for  $F_n$  with this starting condition and assuming that  $\mu \neq 0$ ,

$$0 = \frac{2\mu}{\sigma^2} F_1 = \sum_{j=1}^{\infty} \left( \frac{2\mu}{\sigma^2} \right)^j \frac{1}{j!} \mathbb{E}[(\Delta p)^j]$$

Now the moment generating function can be written as

$$M_{\Delta p}(\varphi) \equiv \int e^{\varphi \Delta p} dQ(\Delta p) = 1 + \sum_{j=1}^{\infty} \frac{\varphi^j}{j!} \int (\Delta p)^j dQ(\Delta p)$$

Hence  $M_{\Delta p}(\frac{2\mu}{\sigma^2}) = 1$ .

*Step 2.* As an intermediate step we develop an alternative expression for  $N_a$ , which will be used below. By a mass preservation argument in the time dependent version of the Kolmogorov equation, continuity of  $f$  at  $x = x^*$ , and the boundary conditions at  $x = \bar{x}$  and  $x = \underline{x}$ ,

$$0 = \frac{\sigma^2}{2} [f'(\bar{x}) - f'(x_+^*) + f'(x_-^*) - f'(\underline{x})] - \int_{\underline{x}}^{\bar{x}} \Lambda(x) f(x) dx$$

Replacing the expression for  $N_a$  we obtain  $N_a = \frac{\sigma^2}{2} [f'(x_-^*) - f'(x_+^*)]$ .

*Step 3.* Now we turn to obtain the invariant distribution of price gaps  $f$ . Using the definition of the density of price changes into the Kolmogorov forward equation for any  $x \in (\underline{x}, \bar{x}) / \{x^*\}$ ,

$$f''(x) = \frac{2}{\sigma^2} [f'(x)\mu + \Lambda(x)f(x)] = \frac{2\mu}{\sigma^2} f'(x) + \frac{2N_a}{\sigma^2} q(x^* - x)$$

which is a non-homogenous first order ordinary differential equation with constant coefficient. Letting  $\bar{a} \equiv f'(x_+^*) < 0$ , and  $\underline{a} \equiv f'(x_-^*) > 0$ , we can solve the initial value problem for  $f'$  using the definition of the function  $R$  in the statement of the theorem:

$$\begin{aligned} f'(x; \underline{a}) &= e^{\frac{2\mu}{\sigma^2}(x-x^*)} \left[ \underline{a} - \frac{2N_a}{\sigma^2} R(0, x^* - x) \right] \text{ for } x \in (\underline{x}, x^*) \\ f'(x; \bar{a}) &= e^{\frac{2\mu}{\sigma^2}(x-x^*)} \left[ \bar{a} + \frac{2N_a}{\sigma^2} R(x^* - x, 0) \right] \text{ for } x \in [x^*, \bar{x}) \end{aligned}$$

which can be verified to solve the first order linear o.d.e. for  $f'$  in each of the segments. Hence,

$$f(x) = f(x^*) + \frac{2N_a}{\sigma^2} \int_x^{x^*} e^{\frac{2\mu}{\sigma^2}(z-x^*)} R(0, x^* - z) dz + \frac{a\sigma^2}{2\mu} \left[ e^{\frac{2\mu}{\sigma^2}(x-x^*)} - 1 \right] \text{ for } x \in (\underline{x}, x^*)$$

$$f(x) = f(x^*) + \frac{2N_a}{\sigma^2} \int_{x^*}^x e^{\frac{2\mu}{\sigma^2}(z-x^*)} R(x^* - z, 0) dz + \frac{\bar{a}\sigma^2}{2\mu} \left[ e^{\frac{2\mu}{\sigma^2}(x-x^*)} - 1 \right] \text{ for } x \in [x^*, \bar{x}]$$

We now derive two equations that  $\underline{a}$  and  $\bar{a}$  must satisfy. Imposing  $f(\bar{x}) = f(\underline{x}) = 0$  and that  $f$  is continuous at  $x = x^*$ , we get

$$f(x^*) = \int_{\underline{x}}^{x^*} f'(x; \underline{a}) dx = -\frac{2N_a}{\sigma^2} \int_{\underline{x}}^{x^*} e^{\frac{2\mu}{\sigma^2}(x-x^*)} R(0, x^* - x) dx + \frac{a\sigma^2}{2\mu} \left[ 1 - e^{\frac{2\mu}{\sigma^2}(x-x^*)} \right]$$

$$= -\int_{x^*}^{\bar{x}} f'(x; \bar{a}) dx = -\frac{2N_a}{\sigma^2} \int_{x^*}^{\bar{x}} e^{\frac{2\mu}{\sigma^2}(x-x^*)} R(x^* - x, 0) dx - \frac{\bar{a}\sigma^2}{2\mu} \left[ e^{\frac{2\mu}{\sigma^2}(\bar{x}-x^*)} - 1 \right]$$

Thus, the system of two linear independent equations in  $\underline{a}$  and  $\bar{a}$  is:

$$\frac{2N_a}{\sigma^2} \left[ \int_{\underline{x}}^{x^*} e^{\frac{2\mu}{\sigma^2}(x-x^*)} R(0, x^* - x) dx - \int_{x^*}^{\bar{x}} e^{\frac{2\mu}{\sigma^2}(x-x^*)} R(x^* - x, 0) dx \right]$$

$$= \underline{a} \left[ \frac{1 - e^{\frac{2\mu}{\sigma^2}(x-x^*)}}{2\mu/\sigma^2} \right] + \bar{a} \left[ \frac{e^{\frac{2\mu}{\sigma^2}(\bar{x}-x^*)} - 1}{2\mu/\sigma^2} \right]$$

$$\frac{2N_a}{\sigma^2} = \underline{a} - \bar{a}$$

To arrive at the expressions in the statement of the theorem, replace  $2\mu/\sigma^2$  with  $\phi$  and normalize  $\bar{a}$  and  $\underline{a}$  by  $2N_a/\sigma^2$ :

$$\left[ \int_{\underline{x}}^{x^*} e^{\phi(x-x^*)} R(0, x^* - x) dx - \int_{x^*}^{\bar{x}} e^{\phi(x-x^*)} R(x^* - x, 0) dx \right] = \underline{a} \left[ \frac{1 - e^{\phi(x-x^*)}}{\phi} \right] + \bar{a} \left[ \frac{e^{\phi(\bar{x}-x^*)} - 1}{\phi} \right]$$

$$1 = \underline{a} - \bar{a}$$

The expression for  $f(x^*)$  is now

$$f(x^*) = -\frac{2N_a}{\sigma^2} \left( \int_{x^*}^{\bar{x}} e^{\phi(x-x^*)} R(x^* - x, 0) dx - \frac{\bar{a}}{\phi} [e^{\phi(\bar{x}-x^*)} - 1] \right)$$

The expressions for  $f(x)$  on both sides of  $x^*$  are

$$f(x) = f(x^*) + \frac{2N_a}{\sigma^2} \left( \int_x^{x^*} e^{\phi(z-x^*)} R(0, x^* - z) dz + \frac{a}{\phi} [e^{\phi(x-x^*)} - 1] \right) \text{ for } x \in (\underline{x}, x^*)$$

$$f(x) = f(x^*) + \frac{2N_a}{\sigma^2} \left( \int_{x^*}^x e^{\phi(z-x^*)} R(x^* - z, 0) dz + \frac{\bar{a}}{\phi} [e^{\phi(x-x^*)} - 1] \right) \text{ for } x \in [x^*, \bar{x}]$$

It is now straightforward to change the variables from  $x$  to  $y = x - x^*$  and go from  $f(x)$  to  $\tilde{f}(y + x^*) = f(x)/(2N_a/\sigma^2)$ . This completes the proof.  $\square$



## B Proofs of the Propositions

**Proof.** (of Proposition 1) The details are in Section G. The proof follows the same steps as that of Theorem 1.

**Proof.** (of Proposition 2) To show this let the density of the invariant distribution be  $\tilde{f}(z) = f(z/b)/b$ . This function solves the KFE for  $\tilde{\Lambda}$  and  $\tilde{\sigma}^2$ . This can be verified using that  $f$  solves the KFE for  $\Lambda$  and  $\sigma^2$ . Since  $N_a = -\sigma^2 f'(0)$  and  $\tilde{N}_a = -\tilde{\sigma}^2 \tilde{f}'(0)$  then it implies that  $\tilde{N}_a = N_a$  for any  $b$ . Also we can see that  $\tilde{q}(z) = q(z/b)/b$ , by using  $q(x) = \Lambda(x)f(x)/N_a$  and  $\tilde{q}(z) = \tilde{\Lambda}(z)\tilde{f}(z)/\tilde{N}_a$  for all  $z \in (-Xb, Xb)$ . Using the formula for a change on variable, and the relationship between  $q$  and  $\tilde{q}$  and of  $\Lambda$  and  $\tilde{\Lambda}$  we get  $\int_{-X}^X \Lambda(x)f(x)dx = \int_{-X}^X \tilde{\Lambda}(z)\tilde{f}(z)dz$ , and thus  $\tilde{s} = s$ .  $\square$

**Proof.** (of Proposition 3) We start by describing the o.d.e and boundary that  $f$  and  $f_k$  satisfy. For  $f$  we have:

$$\begin{aligned}\Lambda(x)f(x) &= \frac{\sigma^2}{2}f''(x) \text{ for all } x \in (0, X) \\ f(X) &= 0 \\ 1/2 &= \int_0^X f(x)dx\end{aligned}$$

For  $f_k$  we have

$$\begin{aligned}\Lambda(x)f_k(x) &= \frac{\sigma^2}{2}f_k''(x) \text{ for all } x \in (0, X) \\ kf(x) &= \frac{\sigma^2}{2}f_k''(x) \text{ for all } x \in (X, \infty) \\ 1/2 &= \int_0^X f_k(x)dx + \int_X^\infty f_k(x)dx\end{aligned}$$

and that  $p_k$  has a continuous first derivative at  $x = X$ . We can then solve for  $f_k$  for  $x > X$ , obtaining  $f_k(x) = f_k(X)e^{-\eta(x-X)}$  for all  $x > X$ , where  $\eta = \sqrt{2k}/\sigma$ . Thus, using the required continuity we can write:

$$\begin{aligned}\Lambda(x)f_k(x) &= \frac{\sigma^2}{2}f_k''(x) \text{ for all } x \in (0, X) \\ f_k'(X) &= -\eta f_k(X) \\ 1/2 &= \int_0^X f_k(x)dx + f_k(X)/\eta\end{aligned}$$

Now consider the solutions of the homogenous second order o.d.e. given by  $\sigma^2/2f''(x) = \Lambda(x)f(x)$  for  $x \in [0, X]$ . Given the assumption that  $\Lambda$  is continuous, we know that the solution is given by linear combinations of two linearly independent functions  $g_1, g_2$  defined  $[0, X]$ . This functions depend on the interval  $(0, X)$ , the constant  $\sigma > 0$  only. Thus we can write the solution of each of the two o.d.e. above as:

$$\begin{aligned}f_k(x) &= a_k g_1(x) + b_k g_2(x) \\ f(x) &= a g_1(x) + b g_2(x)\end{aligned}$$

for all  $x \in [0, X]$ . The coefficients  $a_k, b_k, a, b$  can be chosen to satisfy the two boundary conditions written for  $f$  and  $f_k$ . We can use the homogeneity of the boundary conditions and preliminary set  $a_k = a = 1$ , drop the boundary conditions given by the integral equation for each system, use  $\bar{b}, b_k$  to solve the remaining boundary conditions at  $X$ , and then find  $a, a_k$  and rescale  $b, b_k$  to satisfy the two integral equations. To do so, let  $\hat{b} = b/a$  and  $\hat{b}_k = a_k/b_k$ . Thus we write the remaining boundary conditions:

$$\begin{aligned} f(X) = 0 & \text{ becomes } 0 = g_1(X) + \hat{b}g_2(X) \\ f'_k(X) = -\eta f_k(X) & \text{ becomes } g'_1(X) + \hat{b}_k g'_2(X) = -\eta \left[ g_1(X) + \hat{b}_k g_2(X) \right] \end{aligned}$$

equivalently we can write:

$$\hat{b} = -\frac{g_1(X)}{g_2(X)} \text{ and } \hat{b}_k = -\frac{\eta g_1(X) + g'_1(X)}{\eta g_2(X) + g'_2(X)}$$

Furthermore let  $I_i \equiv \int_0^X g_i(x)dx$  for  $i = 1, 2$  so that we can write the remaining boundary conditions as:

$$\begin{aligned} 1/2 = aI_1 + bI_2 & \implies a = \frac{1}{2(I_1 + I_2\hat{b})} \\ 1/2 = a_k I_1 + b_k I_2 + a_k \frac{g_1(X)}{\eta} + b_k \frac{g_2(X)}{\eta} & \implies a_k = \left( I_1 + \hat{b}_k I_2 + \frac{g_1(X)}{\eta} + \hat{b}_k \frac{g_2(X)}{\eta} \right) / 2 \end{aligned}$$

Note that, given the expression for  $\eta$ , taking  $k \rightarrow \infty$  it is equivalent to take  $\eta \rightarrow \infty$ . Then, using L'Hopital in the second equation we obtain that  $\hat{b}_k \rightarrow \hat{b}$ , which then implies that  $a_k \rightarrow a$  and finally  $b_k \rightarrow b$ . Now we can compare  $f_k$  and  $f$  to obtain:

$$\begin{aligned} |f_k(x) - f(x)| &= |(a_k - a)g_1(x) + (b_k - b)g_2(x)| \\ &\leq |a_k - a||g_1(x)| + |b_k - b||g_2(x)| \text{ for all } x \in [0, X] \end{aligned}$$

Since  $g_1$  and  $g_2$  are continuous in  $x$ , then they are bounded in  $[0, X]$ . Thus as  $k \rightarrow \infty$  we have that  $f_k$  converges uniformly to  $f$ .  $\square$

**Proof.** (of [Proposition 4](#)) Absolute continuity of  $Q(\cdot)$  follows from continuity of  $f(\cdot)$  on  $(-X, X)/\{0\}$  and boundedness of  $\Lambda(\cdot)$  on  $(-X, X)$ . Symmetry of  $q(\cdot)$  follows from both  $f(\cdot)$  and  $\Lambda(\cdot)$  being symmetric, and its continuity follows from the continuity of  $f(\cdot)$ .

That  $Q(\cdot)$  is fully identified by all its moment requires either  $X < \infty$  or the existence of its moment generating function in some neighborhood of zero when  $X = \infty$ . This is Theorem 2.3.11 in [Casella and Berger \(2002\)](#). Take the case  $X = \infty$ . We will show the existence of the moment generating function in a neighborhood of zero, which amounts to convergence of a series

$$\sum_{n=0}^{\infty} \frac{(ia)^n \mathbb{E}[x^n]}{n!}$$

for some  $a > 0$ . Due to symmetry, all odd moments are zero, so we will prove that the even moments grow no faster than the factorial.

Consider an even moment  $\mathbb{E}[x^{2k+2}]$ :

$$\mathbb{E}[x^{2k+2}] = \int_{-\infty}^{\infty} x^{2k+2} q(x) dx = \frac{2}{N_a} \int_0^{\infty} x^{2k+2} \Lambda(x) f(x) dx = \frac{\sigma^2}{N_a} \int_0^{\infty} x^{2k+2} f''(x) dx$$

This uses the definition of and symmetry  $q(\cdot)$  and the KFE. Integrate the right-hand side by parts twice:

$$\frac{\sigma^2}{N_a} \int_0^{\infty} x^{2k+2} f''(x) dx = \frac{\sigma^2(2k+2)(2k+1)}{N_a} \int_0^{\infty} x^{2k} f(x) dx$$

Here we used the fact that, due to **Assumption 1**,  $\Lambda(\cdot)$  is bounded away from zero for  $x > x^H$ , so the decay rate of  $q(\cdot)$  is no slower than exponential. This drives the intermediate terms from integration by parts to zero.

Now we will prove that

$$\int_0^{\infty} x^{2k} f(x) dx \leq \xi \int_0^{\infty} x^{2k} \Lambda(x) f(x) dx$$

for some number  $\xi$  that does not depend on  $k$ . Two cases are interesting. First is when there is a number  $\lambda_1 > 0$  such that  $\Lambda(x) > \lambda$  with probability one with respect to the measure defined by  $f(\cdot)$ . In this case,

$$\int_0^{\infty} x^{2k} f(x) dx \int_0^{\infty} \frac{1}{\Lambda(x)} x^{2k} \Lambda(x) f(x) dx < \frac{1}{\lambda} \int_0^{\infty} x^{2k} \Lambda(x) f(x) dx$$

and we are done. Now assume, on the contrary, for any positive number  $\lambda$  there is a positive measure (corresponding to  $f(\cdot)$ ) of  $x$  such that  $\Lambda(x) < \lambda$ . Recall that, by **Assumption 1**, there exist  $x^H > 0$  and  $\lambda > 0$  such that  $\Lambda(x) > \lambda$  for  $x > x^H$ . Then there exists a pair of numbers  $(\lambda_2, x_2)$  with and two sets  $A_1$  and  $A_2$  such that  $A_1 = \{x : \Lambda(x) < \lambda_2\}$ ,  $A_2 = [x_2, \infty)$ , the measures of  $A_1$  and  $A_2$  associated with  $f(\cdot)$  are equal to  $F > 0$ , and

$$\int_{A_1} \left( F\Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) f(x) dx = - \int_{A_2} \left( F\Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) f(x) dx \quad (54)$$

To see why these sets exist, take first  $x_2 = x^H$ . If there is no  $\lambda_2 < \lambda$  such that the measure of  $\{x : \Lambda(x) < \lambda_2\}$  is equal to  $[x_2, \infty)$ , increase  $x_2$  until there is. Since  $X = \infty$ , the measure of  $[x_2, \infty)$  decreases continuously as  $x_2$  increases, so for any  $\lambda_1 < \lambda$  the value of  $x_2 \geq x^H$  such that the measures of  $A_2$  and  $A_1$  are equal exists.

Now consider the difference

$$\begin{aligned}
& F \int_{A_1 \cup A_2} x^{2k} \Lambda(x) f(x) dx - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \int_{A_1 \cup A_2} x^{2k} f(x) dx \\
&= \int_{A_1 \cup A_2} \left( F \Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) x^{2k} f(x) dx \\
&= \int_{A_1} \left( F \Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) x^{2k} f(x) dx + \int_{A_2} \left( F \Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) x^{2k} f(x) dx
\end{aligned}$$

Consider the last line. We know from [equation \(54\)](#) that the expression in brackets under the first integral is negative, and that under the second integral is positive. This is because they are the sum to zero, and  $\Lambda(x)$  is greater on  $A_2$  than on  $A_1$ . We also know that  $x \leq x^H$  on  $A_1$  and  $x \geq x^H$  on  $A_2$ . Hence,

$$\begin{aligned}
& \int_{A_1} \left( F \Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) x^{2k} f(x) dx + \int_{A_2} \left( F \Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) x^{2k} f(x) dx \\
& \geq (x^H)^{2k} \left[ \int_{A_1} \left( F \Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) f(x) dx + \int_{A_2} \left( F \Lambda(x) - \int_{A_1 \cup A_2} \Lambda(x) f(x) dx \right) f(x) dx \right] \\
& = 0
\end{aligned} \tag{55}$$

This insures

$$\int_{A_1 \cup A_2} x^{2k} f(x) dx \leq \frac{F}{\int_{A_1 \cup A_2} \Lambda(x) f(x) dx} \int_{A_1 \cup A_2} x^{2k} \Lambda(x) f(x) dx = \xi_1 \int_{A_1 \cup A_2} x^{2k} \Lambda(x) f(x) dx$$

At the same time,

$$\int_{\mathbb{R}_+ / \{A_1 \cup A_2\}} x^{2k} f(x) dx \leq \frac{1}{\lambda_2} \int_{\mathbb{R}_+ / \{A_1 \cup A_2\}} x^{2k} \Lambda(x) f(x) dx = \xi_2 \int_{\mathbb{R}_+ / \{A_1 \cup A_2\}} x^{2k} \Lambda(x) f(x) dx$$

Hence,

$$\int_0^\infty x^{2k} f(x) dx \leq \max\{\xi_1, \xi_2\} \int_0^\infty x^{2k} \Lambda(x) f(x) dx = \max\{\xi_1, \xi_2\} \mathbb{E} [x^{2k}]$$

Pluggin this to what was obtained before,

$$\mathbb{E} [x^{2k+2}] \leq \frac{\sigma^2(2k+2)(2k+1) \max\{\xi_1, \xi_2\}}{N_a} \mathbb{E} [x^{2k}]$$

This implies that the series in question converges, and thus the moment generating function exists, at least in the circle of the radius  $\sqrt{N_a / (\sigma^2 \max\{\xi_1, \xi_2\})}$ .  $\square$

**Proof.** (of Proposition 5) Under the identification assumptions,

$$\begin{aligned} \frac{\mathbb{E}[(\Delta p_{it})^j (\Delta p_{is})^k]}{\mathbb{E}[(\Delta p_{it})^{j'} (\Delta p_{is})^{k'}]} &= \frac{\mathbb{E}[b_i^{j+k} (\Delta \tilde{p}_t)^j (\Delta \tilde{p}_s)^k]}{\mathbb{E}[b_i^{j'+k'} (\Delta \tilde{p}_t)^{j'} (\Delta \tilde{p}_s)^{k'}]} = \frac{\mathbb{E}[(b_i)^{j+k}] \mathbb{E}[(\Delta \tilde{p}_t)^j] \mathbb{E}[(\Delta \tilde{p}_s)^k]}{\mathbb{E}[(b_i)^{j'+k'}] \mathbb{E}[(\Delta \tilde{p}_t)^{j'}] \mathbb{E}[(\Delta \tilde{p}_s)^{k'}]} \\ &= \frac{\mathbb{E}[(\Delta \tilde{p}_t)^j] \mathbb{E}[(\Delta \tilde{p}_t)^k]}{\mathbb{E}[(\Delta \tilde{p}_t)^{j'}] \mathbb{E}[(\Delta \tilde{p}_t)^{k'}]} \end{aligned}$$

The first equality uses  $\Delta p_{it} = b_i \Delta \tilde{p}_t$ . The second one uses mutual independence of  $b_i$ ,  $\Delta \tilde{p}_t$ , and  $\Delta \tilde{p}_s$ . The last one uses the fact that  $\Delta \tilde{p}_t$  and  $\Delta \tilde{p}_s$  are identically distributed.  $\square$

**Proof.** (of Proposition 6) Start with  $Q(x)$ :

$$Q(x) = \mathbb{P}\{\Delta p_{it} \leq x\} = \int_0^\infty \mathbb{P}\left\{\Delta \tilde{p}_t \leq \frac{x}{b_i} \mid b_i\right\} dH(b_i) = \int_0^\infty \mathbb{P}\left\{\Delta \tilde{p}_t \leq \frac{x}{b_i}\right\} dH(b_i) \quad (56)$$

The last equality uses the mutual independence of  $\Delta \tilde{p}_t$  and  $b_i$ . Differentiate with respect to  $x$ :

$$q(x) = \partial_x \mathbb{P}\{\Delta p_{it} \leq x\} = \int_0^\infty \frac{1}{b_i} \partial_x \mathbb{P}\left\{\Delta \tilde{p}_t \leq \frac{x}{b_i}\right\} dH(b_i)$$

Evaluate at  $x = 0$ :

$$q(0) = \int_0^\infty \frac{1}{b_i} \tilde{q}(0) dH(b_i) = \mathbb{E}[b_i^{-1}] \tilde{q}(0) \quad (57)$$

Now turn to  $\mathcal{C}_{pooled}$ :

$$\mathcal{C}_{pooled} = \frac{q(0) \text{Var}(\Delta p_{it})}{2 \mathbb{E}[|\Delta p_{it}|]} = \frac{\tilde{q}(0) \mathbb{E}[b_i^{-1}] \mathbb{E}[b_i^2] \text{Var}(\Delta \tilde{p}_t)}{2 \mathbb{E}[b_i] \mathbb{E}[|\Delta \tilde{p}_t|]} = \mathcal{C} \frac{\mathbb{E}[b_i^{-1}] \mathbb{E}[b_i^2]}{\mathbb{E}[b_i]}$$

Hence,

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

That the correction multiplier is smaller than one follows from the correlation between  $1/b_i$  and  $b_i^2$  being negative. Next we find the expression for the correction as a function of the moments:

$$\frac{\mathbb{E}[|\Delta p_{it}|]}{\mathbb{E}[|\Delta p_{it}|^{-1}] \mathbb{E}[|\Delta p_{it}|]} = \frac{\mathbb{E}[b_i]}{\mathbb{E}[b_i^{-1}] \mathbb{E}[b_i^2]} \frac{\mathbb{E}[|\Delta \tilde{p}_t|]}{\mathbb{E}[|\Delta \tilde{p}_t|^{-1}] \mathbb{E}[|\Delta \tilde{p}_t|^2]} = \frac{\mathbb{E}[b_i]}{\mathbb{E}[b_i^{-1}] \mathbb{E}[b_i^2]} \frac{\mathbb{E}[|\Delta p_{it}|]}{\mathbb{E}[|\Delta p_{it}|^{-1}] \mathbb{E}[|\Delta p_{is}|^2]}$$

Hence,

$$\frac{\mathbb{E}[b_i]}{\mathbb{E}[b_i^{-1}] \mathbb{E}[b_i^2]} = \frac{\mathbb{E}[|\Delta p_{it}|^{-1}] \mathbb{E}[|\Delta p_{is}|^2]}{\mathbb{E}[|\Delta p_{it}|^{-1}] \mathbb{E}[|\Delta p_{it}|]}$$

This completes the proof.  $\square$

**Proof.** (of Lemma 3) Denote  $S^n(t) \equiv \frac{\partial^n}{\partial t^n} S(t)$ . We will derive the following recursion:

$$S^{(n)}(t) = \mathbb{E} \left[ F_n(x(t)) e^{-\int_0^t \Lambda(x(s)) ds} \mid x(0) = 0 \right] \text{ for all } t \geq 0 \text{ and all } n = 1, 2, \dots \quad (58)$$

for a sequence of functions  $F_n : \mathbb{R} \rightarrow \mathbb{R}$ . For  $n = 1$  it follows from differentiating equation (76) with respect to  $t$ :

$$S^{(1)}(t) = -\mathbb{E} \left[ \Lambda(x(t)) e^{-\int_0^t \Lambda(x(s)) ds} \mid x(0) = 0 \right] \quad (59)$$

thus  $F_1(x) = -\Lambda(x)$ . For the induction step, assume that equation (58) hold and we will differentiate it with respect to  $t$ . To do this, since  $F_n(x(t))$  is an Ito's process, and thus not differentiable with respect to time, we use Ito's lemma for the product of two Ito's process, namely  $F_n(x(t))$  and  $Z(t) \equiv e^{-\int_0^t \Lambda(x(s)) ds}$ , the second one being a degenerate one, since it has bounded variation. We then use that  $dF_n(x(t)) = \partial_{xx} F_n(x(t)) \frac{\sigma^2}{2} dt + \partial_x F_n(x(t)) \sigma dW$ , since  $x$  has no drift, and  $dZ(t) = -\Lambda(x(t)) Z(t) dt$ . Thus,

$$\begin{aligned} S^{(n+1)}(t) &\equiv \lim_{\Delta \downarrow 0} \frac{S^{(n)}(t + \Delta) - S^{(n)}(t)}{\Delta} \\ &= \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathbb{E} [F_n(x(t + \Delta)) Z(t + \Delta) - F_n(x(t)) Z(t) \mid x(0) = 0] \\ &= \mathbb{E} \left[ \left( \frac{\sigma^2}{2} \partial_{xx} F_n(x(t)) - \Lambda(x(t)) F_n(x(t)) \right) Z(t) \mid x(0) = 0 \right] \\ &= \mathbb{E} \left[ \left( \frac{\sigma^2}{2} \partial_{xx} F_n(x(t)) - \Lambda(x(t)) F_n(x(t)) \right) e^{-\int_0^t \Lambda(x(s)) ds} \mid x(0) = 0 \right] \end{aligned}$$

which give us a recursion for  $F_n$ :

$$F_{n+1}(x) = \frac{\sigma^2}{2} \partial_{xx} F_n(x) - \Lambda(x) F_n(x) \text{ for all } x \quad (60)$$

Finally, evaluating the  $n^{th}$  derivatives of  $S$  at  $t = 0$  we have:

$$S^{(n)}(0) = F_n(0) \text{ and all } n = 1, 2, \dots \quad (61)$$

This completes the proof.  $\square$

**Proof.** (of Proposition 7) In the text.  $\square$

**Proof.** (of Proposition 8) Let the price gap distributions that correspond to  $\Lambda_1$  and  $\Lambda_2$  be  $f_1$  and  $f_2$ . Recall that for a fixed  $N_a$  and  $\sigma^2$  we have  $f_1'(0) = f_2'(0)$  and it is sufficient to compare

$$\int_0^\infty f_1(x) x^2 dx \text{ against } \int_0^\infty f_2(x) x^2 dx$$

(1) We first claim that the graph of the function  $f_1(x) - f_2(x)$  cannot cross the  $x$ -axis from above. That is, there is no segment  $[a, b]$  such that  $f_1(x) - f_2(x) = 0$  on this segment,  $f_1(x) - f_2(x) > 0$  to the left of  $a$ , and  $f_1(x) - f_2(x) > 0$  to the right of  $b$ . Note that this nests the case when  $a = b$  and hence  $[a, b]$  is a single point. Suppose such a segment exists. Then one of the two statements is true: either  $\Lambda_1(x) \geq \Lambda_2(x)$  for all  $x \leq a$  or  $\Lambda_1(x) \leq \Lambda_2(x)$  for all  $x \geq b$ .

In the first case, the graph of  $f_1(x) - f_2(x)$  never crosses the  $x$ -axis again to the left of  $a$ . If it does cross it at some  $c < a$ , on  $(c, a)$  we have  $f_1(x) > f_2(x)$  and hence  $\Lambda_1(x)f_1(x) > \Lambda_2(x)f_2(x)$ , implying  $f_1''(x) > f_2''(x)$ . But this contradicts  $f_1'(c) - f_2'(c) \geq 0$  and  $f_1'(a) - f_2'(a) \leq 0$  holding simultaneously. Hence, for all  $x < a$  we have  $f_1(x) > f_2(x)$ , implying  $\Lambda_1(x)f_1(x) > \Lambda_2(x)f_2(x)$  and  $f_1''(x) > f_2''(x)$  on  $(0, a)$ . But since  $f_1'(a) \leq f_2'(a)$ , in this region we have  $f_1'(x) < f_2'(x)$ , which contradicts  $f_1'(0) = f_2'(0)$ .

In the second case, the graph of  $f_1(x) - f_2(x)$  never crosses the  $x$ -axis again to the right of  $b$ . If it does cross it at some  $d > b$ , on  $(b, d)$  we have  $f_1(x) < f_2(x)$  and hence  $\Lambda_1(x)f_1(x) < \Lambda_2(x)f_2(x)$ , implying  $f_1''(x) < f_2''(x)$ . But this contradicts  $f_1'(b) - f_2'(b) \leq 0$  and  $f_1'(d) - f_2'(d) \geq 0$  holding simultaneously. Hence the graph of  $f_1(x) - f_2(x)$  never crosses the  $x$ -axis again to the right of  $b$ , which already rules out  $X_1 > X_2$ . Moreover, if  $X_1 = X_2 \leq \infty$ , it must hold that  $f_1'(X_1) \geq f_2'(X_1)$ , which contradicted by  $f_1'(x) < f_2'(x)$  for  $x > b$ . The latter follows from  $f_1'(b) - f_2'(b) \leq 0$  and  $f_1''(x) < f_2''(x)$  for  $x > b$ .

(2) Since the graph of the function  $f_1(x) - f_2(x)$  cannot cross the  $x$ -axis from above, it can only cross the  $x$ -axis from below. We know that there must be at least one crossing, because  $f_1$  and  $f_2$  are continuous and both integrate to one. Hence, the function  $f_1(x) - f_2(x)$  is non-positive until some point and non-negative after some point until  $X_1$ . Moreover, there are segments of strict positivity and strict negativity. Hence,

$$\int_0^{X_1} (f_1(x) - f_2(x))x^2 dx > 0$$

This completes the proof.  $\square$

**Proof.** (of [Corollary 2](#)) Fix  $X$  and let  $\Lambda_1(x) \equiv \lambda_1$  on  $(0, X)$  correspond to the Calvo<sup>+</sup> model. The other hazard function,  $\Lambda_2$ , is at least somewhere strictly increasing. We claim it cannot be that  $\Lambda_2(x) \geq \lambda_1$  for all  $x$ . Assume toward a contradiction that this is the case.

Then it cannot be that the graph of  $f_2(x) - f_1(x)$  crosses the  $x$ -axis from below on  $(0, X)$ . If it does, there is a segment  $[a, b]$  such that  $f_2(x) - f_1(x)$  is positive to the right of  $b$ . But then the graph of  $f_2(x) - f_1(x)$  never crosses the  $x$ -axis on  $(b, X]$  again, because if it did cross it at some  $d > b$ , we would have  $\Lambda_2(x)f_2(x) > \Lambda_1(x)f_1(x)$  on  $(b, d)$ , implying  $f_2''(x) > f_1''(x)$  on  $(b, d)$ , which contradicts  $f_2'(b) \geq f_1'(b)$  and  $f_2'(d) \leq f_1'(d)$  holding simultaneously. But we know that  $f_1(X) = f_2(X) = 0$ , which yields a contradiction.

Neither can it be that the graph of  $f_2(x) - f_1(x)$  crosses the  $x$ -axis from above on  $(0, X)$ . If it does, there is a segment  $[a, b]$  such that  $f_2(x) - f_1(x)$  is positive to the left of  $a$ . But then the graph of  $f_2(x) - f_1(x)$  never crosses the  $x$ -axis on  $[0, a)$  again, because if it did cross it at some  $c < a$ , we would have  $\Lambda_2(x)f_2(x) > \Lambda_1(x)f_1(x)$  on  $(c, a)$ , implying  $f_2''(x) > f_1''(x)$  on  $(c, a)$ , which contradicts  $f_2'(a) \leq f_1'(a)$  and  $f_2'(c) \geq f_1'(c)$  holding simultaneously. Hence,  $\Lambda_2(x)f_2(x) > \Lambda_1(x)f_1(x)$  on  $(c, a)$ , implying  $f_2''(x) > f_1''(x)$  on  $(c, a)$ . But together with  $f_2'(a) \leq f_1'(a)$  this contradicts  $f_1'(0) = f_2'(0)$ .

Hence, the graph of  $f_2(x) - f_1(x)$  does not cross the  $x$ -axis from above or below on  $(0, X)$ . But  $\Lambda_2$  is not identically equal to  $\lambda_1$ , so  $f_2$  cannot coincide with  $f_1$  everywhere. This yields the contradiction. Now we know that  $\Lambda_2(x) < \lambda_1$  for some  $x$ . Since  $\Lambda_2$  is non-decreasing, the conditions of [Proposition 8](#) are satisfied, and  $\Lambda_1$  generates a higher kurtosis of price changes. This completes the proof.  $\square$

**Proof.** (of [Corollary 3](#)) Let  $X_1 > X_2$  and let  $\Lambda_1$  and  $\Lambda_2$  be constants  $\lambda_1$  and  $\lambda_2$  on their intervals. We claim that  $\lambda_1 > \lambda_2$ . Assume toward the contradiction  $\lambda_1 \leq \lambda_2$ . We know that the graph of the function  $f_1(x) - f_2(x)$  must cross the  $x$ -axis from below at some point, because  $f_1(X_2) > 0$ ,

$f_2(X_2) = 0$ , and both  $f_1$  and  $f_2$  integrate to one. Hence, there is a point  $a$  such that  $f_1(x) < f_2(x)$  to the left of  $a$ . Then the graph of  $f_1(x) - f_2(x)$  never crosses the  $x$ -axis on  $(0, a)$  again, since if it did there would be a point  $c < a$  such that on  $(c, a)$  we have  $f_1(x) < f_2(x)$  and hence  $\Lambda_1(x)f_1(x) < \Lambda_2(x)f_2(x)$ , implying  $f_1''(x) < f_2''(x)$  everywhere on  $(c, a)$ . The latter contradicts  $f_1'(a) \geq f_2'(a)$  and  $f_1'(c) \leq f_2'(c)$  holding simultaneously.

But that the graph of  $f_1(x) - f_2(x)$  never crosses the  $x$ -axis on  $(0, a)$  again means that  $f_1(x) < f_2(x)$  and hence  $\Lambda_1(x)f_1(x) < \Lambda_2(x)f_2(x)$ , implying  $f_1''(x) < f_2''(x)$  everywhere on  $(0, a)$ . Together with  $f_1'(a) \geq f_2'(a)$  this contradicts  $f_1'(0) = f_2'(0)$ . Hence,  $\lambda_1 > \lambda_2$ . The pair  $\Lambda_1$  and  $\Lambda_2$  thus qualify for the [Proposition 8](#), and  $\Lambda_1$  generates a higher kurtosis of price changes. Hence, within the space of constant hazard functions with barriers higher  $X$  generate higher Kurtoses. By [Proposition 3](#), the kurtosis for  $X = \infty$  is the limit of any sequence generated by constant hazard functions with  $X_k \rightarrow \infty$ . Without loss of generality, the sequence can be constructed as monotone, so the kurtosis for  $X = \infty$  is higher than any its element. But the kurtosis for an arbitrary  $\Lambda$  is majorized by that corresponding to a constant  $\tilde{\Lambda}$  with the same barrier. Hence, the kurtosis for a constant  $\Lambda$  and  $X = \infty$  is the highest possible one. This completes the proof.  $\square$

**Proof.** (of [Corollary 4](#)) If the two hazard functions have the same curvature  $k(x)$ , it means that

$$\Lambda_1(x) = \Lambda_1(0) + \Lambda_1'(0) \int_0^x e^{\int_0^z \frac{k(w)}{w} dw} dz$$

$$\Lambda_2(x) = \Lambda_2(0) + \Lambda_2'(0) \int_0^x e^{\int_0^z \frac{k(w)}{w} dw} dz$$

We have  $\mathcal{C}_1 > \mathcal{C}_2$  if and only if  $\Lambda_1(0) > \Lambda_2(0)$ . Using the same method as in the proof of [Corollary 2](#), we can show that, since the frequency of adjustment is the same, there exists a  $z < X$  such that  $\Lambda_1(z) < \Lambda_2(z)$ . Hence,  $\Lambda_1'(0) < \Lambda_2'(0)$ , and  $\Lambda_1(x) - \Lambda_2(x)$  is a decreasing function. The graphs of  $\Lambda_1(\cdot)$  and  $\Lambda_2(\cdot)$  thus only cross once, so they qualify for [Proposition 8](#), and  $Kurt_1(\Delta p) > Kurt_2(\Delta p)$ .  $\square$

**Proof.** (of [Proposition 9](#)) Fix  $\nu \geq 0$ . In [Lemma 4](#), we know that  $s$  increases in  $\rho$ , so it is sufficient to show that  $Kurt(\Delta p)$  also does. For this purpose, take some  $\rho_1 = 2\kappa_1 X_1^2 / \sigma_1^2$ . They generate  $f_1(\cdot)$  with

$$\kappa_1 \left( \frac{x}{X_1} \right)^\nu f_1(x) = \frac{\sigma_1^2}{2} f_1''(x)$$

Now we want to increase  $\rho_1$  to some  $\rho_2 > \rho_1$ . This can induce multiple  $f_2(\cdot)$ , since the distribution of price gaps also depends on  $X$  and  $\sigma^2$ . But the kurtosis of price changes only depends on  $\rho$ , so it suffices to show that one of the densities  $f_2(\cdot)$  corresponding to  $\rho_2$  generates a higher  $Kurt(\Delta p)$ . Let the new  $\rho_2$  and the density  $f_2(\cdot)$  be such that  $\sigma_1^2 = \sigma_2^2$  and  $f_1'(0) = f_2'(0)$ . To compare the Kurtosis in this case it is enough to evaluate the sign of

$$\int_0^{X_2} f_2(x)x^2 dx - \int_0^{X_1} f_1(x)x^2 dx = \int_0^{X_2} (f_2(x) - f_1(x))x^2 dx$$

First, from the proof of [Lemma 4](#) we know that  $\hat{p}_2'(0) < \hat{p}_1'(0)$ , which implies  $X_2 > X_1$  because  $f_2'(0) = f_1'(0)$ . This, in turn, implies that  $f_2(x) - f_1(x)$  is positive on  $(a, X_2)$  for some  $a < X_1$ .



Since  $f_1(\cdot)$  and  $f_2(\cdot)$  integrate to the same number over their supports, there must be a crossing  $b$ , to the left of which  $f_1(x) > f_2(x)$ . At this crossing,  $f_1'(b) \geq f_2'(b)$ . Now we will argue that there is no other crossing  $c < b$ .

Suppose, by way of contradiction, such a crossing exists. We have  $f_2'(c) \leq f_1'(c)$ . Subtract one Kolmogorov forward equations from the other:

$$x^\nu \left[ \frac{\kappa_2}{X_2^\nu} f_2(x) - \frac{\kappa_1}{X_1^\nu} f_1(x) \right] = \frac{\sigma^2}{2} [f_2(x) - f_1(x)]'' \quad (62)$$

Now there are two options:  $\kappa_2/X_2^\nu \geq \kappa_1/X_1^\nu$  or  $\kappa_2/X_2^\nu < \kappa_1/X_1^\nu$ . In the first case, since  $f_2'(c) \leq f_1'(c)$  and  $f_2(x) > f_1(x)$  to the left of  $c$ , from [equation \(62\)](#) we can conclude that  $f_2''(x) > f_1''(x)$  for  $x < c$ , and hence  $f_2'(x) - f_1'(x)$  only increases as  $x$  decreases. But this contradicts  $f_1'(0) = f_2'(0)$ . In the second case, since  $f_2'(c) \leq f_1'(c)$  and  $f_2(x) < f_1(x)$  to the right of  $c$ , from [equation \(62\)](#) we can conclude that  $f_2''(x) < f_1''(x)$  for  $x > c$ , and hence  $f_2'(x) - f_1'(x)$  only decreases as  $x$  decreases. But this contradicts  $f_2'(b) > f_1'(b)$ . Hence, there is no crossing to the left of  $b$ .

This means that  $f_2(x) - f_1(x)$  is negative on  $[0, b)$  and positive on  $(b, X_2)$ . Since it integrates to zero over this whole interval, its integral with any positive increasing function (such as  $x^2$ ) is positive. Hence, the kurtosis is higher for  $\rho_2 > \rho_1$ .  $\square$

**Proof.** (of [Lemma 4](#)) By the definition of  $\hat{f}(\cdot)$ , we have

$$\begin{aligned} \hat{f}(z) &= X f(zX) \\ \hat{f}'(z) &= X^2 f'(zX) \end{aligned}$$

The function  $\hat{f}(\cdot)$  itself is derived from

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

Computing the Kurtosis,

$$Kurt(\Delta p) = \frac{12N_a}{\sigma^2} \int_0^X f(x)x^2 dx = -12f'(0) \int_0^X f(x)x^2 dx = -12\hat{f}'(0) \int_0^1 \hat{f}(z)z^2 dz$$

Since  $\hat{f}(\cdot)$  is completely determined by  $\rho$  and  $\hat{\Lambda}(\cdot)$ , this quantity does not depend on other parameters. The share of adjustment between the boundaries is

$$s = 1 - \frac{f'(X)}{f'(0)} = 1 - \frac{\hat{f}'(1)}{\hat{f}'(0)} \quad (63)$$

It also only depends on  $\hat{\Lambda}(\cdot)$  and  $\rho$ . The frequency of price changes is given by

$$N_a = -\sigma^2 f'(0) = -\frac{\sigma^2}{X^2} \hat{f}'(0)$$

From this we have  $\hat{n}(\rho) = -\hat{f}'(0)$ , so  $\hat{n}(\rho)$  only depends on  $\hat{\Lambda}$  and  $\rho$ . In the case when  $\rho = 0$  the Kolmogorov forward equation is solved by a linear  $\hat{f}(\cdot)$ , and the slope is  $-1$  from the boundary

condition and the normalization. Hence,  $\hat{n}(0) = 1$ . Now take the other statistic:

$$\frac{\text{Kurt}(\Delta p)}{6N_a} = \frac{2X^2}{\sigma^2} \int_0^1 \hat{f}(z)z^2 dz = \frac{X^2}{6\sigma^2} \hat{m}(\rho)$$

Here the function  $\hat{m}(\rho)$  is twelve times the integral of  $\hat{f}(z)z^2$  which only depends on  $\hat{\Lambda}(\cdot)$  and  $\rho$ . In the case when  $\rho = 0$  we have  $\hat{f}(z) = 1 - z$  for  $z \in [0, 1]$  and hence  $\hat{m}(0) = 1$ .

Now fix the shape  $\hat{\Lambda}(\cdot)$ . Consider two different values of  $\rho$ :  $\rho_1 > \rho_2$ . They generate two distributions  $\hat{f}_1(\cdot)$  and  $\hat{f}_2(\cdot)$ . Taking the difference between the Kolmogorov forward equations that define them,

$$\hat{\Lambda}(z)(\rho_1 \hat{f}_1(z) - \rho_2 \hat{f}_2(z)) = (\hat{f}_1(z) - \hat{f}_2(z))''$$

It holds that  $\hat{f}_1(1) = \hat{f}_2(1)$ , so there must be another point  $y \in (0, 1)$  where  $\hat{f}_1(y) = \hat{f}_2(y)$ , because  $\hat{f}_1(\cdot)$  and  $\hat{f}_2(\cdot)$  integrate to the same number. Moreover, this point must be a crossing, meaning that  $\hat{f}_1(z) - \hat{f}_2(z)$  has different signs on to the left and to the right of it. Suppose  $\hat{f}_1(z) - \hat{f}_2(z)$  is positive to the right of  $y$ . This means  $\hat{f}'_1(y) - \hat{f}'_2(y) \geq 0$ . But then to the right of  $y$  it holds that  $\hat{f}'_1(z) - \hat{f}'_2(z) > 0$ , since the left-hand side of [equation \(63\)](#) is positive. Hence, the difference between  $\hat{f}_1(\cdot)$  and  $\hat{f}_2(\cdot)$  only increases to the right of  $y$ , and they cannot cross again at  $z = 1 > y$ . This is a contradiction. The crossing is therefore such that  $\hat{f}'_1(y) - \hat{f}'_2(y) \leq 0$ . But then to the left of  $y$  it holds that  $\hat{f}'_1(z) - \hat{f}'_2(z) < 0$ , since the right-hand side of [equation \(63\)](#) is positive in this region. The difference between  $\hat{f}_1(z)$  and  $\hat{f}_2(z)$  increases as  $z$  decreases, as does the difference between  $\hat{f}'_1(z)$  and  $\hat{f}'_2(z)$ . Hence, the crossing is unique and  $\hat{f}'_1(0) < \hat{f}'_2(0)$ . Moreover,  $\hat{f}_1(z) - \hat{f}_2(z) > 0$  for  $z \in [0, y)$  and  $\hat{f}_1(z) - \hat{f}_2(z) < 0$  for  $z \in (y, 1]$ . From the latter fact together with  $\hat{f}_1(1) = \hat{f}_2(1)$  it follows that  $\hat{f}'_1(1) > \hat{f}'_2(1)$ . To summarize:

- there is a unique  $y \in (0, 1)$  such that  $\hat{f}_1(z) - \hat{f}_2(z) > 0$  for  $z \in [0, y)$  and  $\hat{f}_1(z) - \hat{f}_2(z) < 0$  for  $z \in (y, 1]$ ;
- $\hat{f}'_1(0) < \hat{f}'_2(0)$
- $\hat{f}'_1(1) > \hat{f}'_2(1)$

From the first bulletpoint it follows that  $\hat{m}(\cdot)$  decreases in  $\rho$ . This is because  $\hat{f}_1(\cdot) - \hat{f}_2(\cdot)$  integrates to zero over  $(0, 1)$ . Since it is positive until some  $z$  and negative afterwards, its integral with increasing positive functions (such as  $z^2$ ) is always negative. From the second bulletpoint it follows that  $\hat{n}(\cdot)$  increases in  $\rho$ , because  $\hat{n}(\rho_i) = -\hat{f}'_i(0)$ . From the second and the third bulletpoints combined it follows that  $s$  increases in  $\rho$ , because  $\hat{f}'(1)$  and  $\hat{f}'(0)$  are both negative, so their ratio decreases with  $\rho$ . This completes the proof.  $\square$

**Proof.** (of [Proposition 10](#)) First, observe that if  $v(x; \mu)$  and  $\{\underline{x}(\mu), x^*(\mu), \bar{x}(\mu)\}$  represent a solution to the firm's problem with drift  $\mu$ , then  $\underline{x}(\mu) = -\bar{x}(-\mu)$ ,  $x^*(\mu) = -x^*(-\mu)$ ,  $\bar{x}(\mu) = -\underline{x}(-\mu)$ , and  $v(x; \mu) = v(-x; -\mu)$ . This can be verified directly by plugging. Hence,  $\Lambda(x; \mu) = \Lambda(-x; -\mu)$ , because  $\Lambda(x; \mu) = \kappa G(v(x; \mu) - v(x^*(\mu); \mu))$ .

Second, observe that if  $f(x; \mu)$  solves the Kolmogorov forward equation for  $\mu$  and  $\Lambda(x; \mu)$  then  $f(x; \mu) = f(-x; -\mu)$ . This can again be verified directly by plugging and using  $\Lambda(x; \mu) = \Lambda(-x; -\mu)$ . An implication of this symmetry is that  $f'(x; \mu) = -f'(-x; -\mu)$ . Hence, for the adjustment frequency we can write

$$N_a(\mu) = \int_{\underline{x}(\mu)}^{\bar{x}(\mu)} \Lambda(x; \mu) f(x; \mu) dx + \frac{\sigma^2}{2} f'(\underline{x}(\mu); \mu) - \frac{\sigma^2}{2} f'(\bar{x}(\mu); \mu) \quad (64)$$

$$= \int_{-\bar{x}(-\mu)}^{-\underline{x}(-\mu)} \Lambda(-x; -\mu) f(-x; -\mu) dx - \frac{\sigma^2}{2} f'(-\underline{x}(\mu); -\mu) + \frac{\sigma^2}{2} f'(-\bar{x}(\mu); -\mu) \quad (65)$$

$$= \int_{\underline{x}(-\mu)}^{\bar{x}(-\mu)} \Lambda(x; -\mu) f(x; -\mu) dx + \frac{\sigma^2}{2} f'(\underline{x}(-\mu); -\mu) - \frac{\sigma^2}{2} f'(\bar{x}(-\mu); -\mu) = N_a(-\mu) \quad (66)$$

In a similar vein, using  $q(x) = \Lambda(x)f(x)/N_a$  and hence  $q(x; \mu) = q(-x; -\mu)$ , we can write for any even moment of  $Q(\cdot)$

$$\mathbb{E} [\Delta p^{2k}] (\mu) = \int_{\underline{x}(\mu)}^{\bar{x}(\mu)} x^{2k} q(x; \mu) dx = \int_{-\bar{x}(-\mu)}^{-\underline{x}(-\mu)} x^{2k} q(-x; -\mu) dx = \int_{\underline{x}(-\mu)}^{\bar{x}(-\mu)} x^{2k} q(x; -\mu) dx = \mathbb{E} [\Delta p^{2k}] (-\mu)$$

This holds for the fourth moment and variance, so it holds for Kurtosis as well. Hence, both  $Kurt(\Delta p)$  and  $N_a$  are symmetric in  $\mu$ . They are also analytical functions of  $\mu$  and can be written as

$$Kurt(\Delta p) = \sum_{i=0}^{\infty} a_i \mu^i \quad (67)$$

$$N_a = \sum_{i=0}^{\infty} b_i \mu^i \quad (68)$$

The odd terms in these infinite sums must be zero, meaning

$$Kurt(\Delta p) = a_0 + o(\mu^2) \quad (69)$$

$$N_a = b_0 + o(\mu^2) \quad (70)$$

This completes the proof.  $\square$

**Proof.** (of [Proposition 11](#)) The proof is in the statement of the proposition.

**Proof.** (of [Proposition 12](#)) Differentiating  $\Omega$

$$\Omega'(\delta) = X f(-X + \delta) + \int_{-X}^{-X+\delta} f(x) dx$$

taking  $\delta \rightarrow 0$ , since the invariant distribution satisfies  $f(-X) = 0$ , we have  $\Omega'(0) = 0$ .

Now we seek to characterize  $\lim_{t \downarrow 0} \omega_\delta(t; \delta)$ . We will show that  $\lim_{t \downarrow 0} \omega_\delta(t; 0) = \infty$  if  $X < \infty$ .

For this case we replace the initial condition by  $f(x + \delta)$  by  $f(x) + f'(x)\delta$  where  $f$  is the density of the invariant distribution. We can omitt the contribution from the term  $f(x)$ , since it is equal to zero by virtue of being the invariant distribution.

The KFE gives the following properties:

1. For all  $t > 0$ , since  $-X$  is an exit point,  $f(-X, t) = 0$ .
2. For all  $t > 0$ , there exists  $\underline{x}(t) > -X$ , so that  $f(x, t) < f(x, 0) = f'(x)\delta > 0$  for all  $x \in [-X, \underline{x}(t)]$ . This follows because  $f(x, t)$  is differentiable in  $x$  and  $f(-X, t) = 0$ .
3. For all  $x \in (-X, 0)$  we have:  $f(x, t) \rightarrow f(x, 0)$  as  $t \downarrow 0$ . This follows since  $f(x, t)$  is differentiable in time  $t$  for all  $x$ .

From these properties we obtain that  $f'(-X, t) \rightarrow \infty$  as  $t \downarrow 0$ . Hence,  $\omega_\delta(0, 0) = \infty$ .  $\square$

**Proof.** (of [Proposition 13](#)) The frequency of adjustment is given by

$$\begin{aligned} N_a &= \int_{-\infty}^{\infty} f(x)(\Lambda(0) + \kappa x^\nu) dx = \int_{-\infty}^{\infty} \tilde{f}(z) \left( \Lambda(0) + \kappa \left( \frac{z}{\eta} \right)^\nu \right) dz \\ &= \frac{\kappa}{\eta^\nu} \int_{-\infty}^{\infty} p(z)(\alpha + z^\nu) dz = \frac{\kappa}{\eta^\nu} \tilde{N}(\nu, \alpha) = \frac{\beta^2 \eta^2}{2} \tilde{N}(\nu, \alpha) \end{aligned}$$

The flexibility index is

$$\begin{aligned} \mathcal{F} &= - \int_{-\infty}^{\infty} x(\Lambda(0) + \kappa x^\nu) f'(x) dx = - \int_{-\infty}^{\infty} z \left( \Lambda(0) + \kappa \left( \frac{z}{\eta} \right)^\nu \right) p'(z) dz \\ &= - \frac{\kappa}{\eta^\nu} \int_{-\infty}^{\infty} z(\alpha + z^\nu) \tilde{f}'(z) dz = \frac{\kappa}{\eta^\nu} \left( \int_{-\infty}^{\infty} \tilde{f}(z)(\alpha + z^\nu) dz + \nu \int_{-\infty}^{\infty} p(z) z^\nu dz \right) \\ &= \frac{\kappa}{\eta^\nu} \left( \tilde{N}(\nu, \alpha)(1 + \nu) - \nu \alpha \right) = \frac{\beta^2 \eta^2}{2} \left( \tilde{N}(\nu, \alpha)(1 + \nu) - \nu \alpha \right) \end{aligned}$$

The distribution of price changes is given by

$$q(x) = \frac{f(x)(\Lambda(0) + \kappa x^\nu)}{N_a} = \frac{\eta(\eta x)(\alpha + (\eta x)^\nu)}{\tilde{N}(\nu, \alpha)}$$

To compute the kurtosis, we need the fourth moment and the variance:

$$\begin{aligned} \mathbb{E}[\Delta p^4] &= \int_{-\infty}^{\infty} x^4 q(x) dx = \frac{1}{\eta^4 N(\nu, \alpha)} \int_{-\infty}^{\infty} z^4 p(z)(\alpha + z^\nu) dz \\ \mathbb{E}[\Delta p^2] &= \int_{-\infty}^{\infty} x^2 q(x) dx = \frac{1}{\eta^2 N(\nu, \alpha)} \int_{-\infty}^{\infty} z^2 p(z)(\alpha + z^\nu) dz \end{aligned}$$

These expressions imply that  $\mathbb{E}[\Delta p^4]/\mathbb{E}[\Delta p^2]^2$  only depends on  $(\nu, \alpha)$ .  $\square$

**Proof.** (of [Proposition 14](#)) Let  $f_1(x)$  and  $f_2(x)$  be the price gap distributions generated by  $\Lambda_1(x)$  and  $\Lambda_2(x)$ . Assume without loss that  $\kappa_1 < \kappa_2$ . We will first prove that  $\Lambda_1(0) > \Lambda_2(0)$  whenever  $N_a$  is the same in the two models. That  $Kurt_1(\Delta p) > Kurt_2(\Delta p)$  will then follow from [Proposition 8](#). Finally, we will show that  $\mathcal{F}_1 < \mathcal{F}_2$ .

(1) Suppose by contradiction that  $\Lambda_1(0) \leq \Lambda_2(0)$ . Then,  $\Lambda_1(x) < \Lambda_2(x)$  for all  $x > 0$ . Since  $N_a$  and  $\sigma^2$  are the same in the two models, we know that  $f'_1(0) = f'_2(0)$ .

Suppose there is a point  $a > 0$  at which the graph of  $f_1(x)$  crosses that of  $f_2(x)$  from below. That is,  $f_1(a) = f_2(a)$  and  $f_1(x) < f_2(x)$  to the left of  $a$ . Then the graphs of  $f_1(x)$  and  $f_2(x)$  never

cross again to the left of  $a$ . If they did cross at some point  $b < a$ , we would have  $f_1'(a) \geq f_2'(a)$  and  $f_1'(b) \leq f_2'(b)$ , so that  $f_1'(a) - f_1'(b) \geq f_2'(a) - f_2'(b)$ , but this is impossible, since  $f_1(x) < f_2(x)$  and  $\Lambda_1(x) < \Lambda_2(x)$  on  $(a, b)$ , while  $\sigma^2 f_i''(x)/2 = \Lambda_i(x) f_i(x)$  for  $i \in \{1, 2\}$ . Hence,  $f_1(x) < f_2(x)$  for all  $x < a$ , which contradicts  $f_1'(0) = f_2'(0)$  for the same reason.

Suppose there is a point  $c > 0$  at which the graph of  $f_1(x)$  crosses that of  $f_2(x)$  from above. That is,  $f_1(c) = f_2(c)$  and  $f_1(x) < f_2(x)$  to the right of  $c$ . Then the graphs of  $f_1(x)$  and  $f_2(x)$  never cross again to the right of  $c$ . If they did cross at some point  $d > c$ , we would have  $f_1'(d) \geq f_2'(d)$  and  $f_1'(c) \leq f_2'(c)$ , so that  $f_1'(d) - f_1'(c) \geq f_2'(d) - f_2'(c)$ , but this is impossible, since  $f_1(x) < f_2(x)$  and  $\Lambda_1(x) < \Lambda_2(x)$  on  $(c, d)$ , while  $\sigma^2 f_i''(x)/2 = \Lambda_i(x) f_i(x)$  for  $i \in \{1, 2\}$ . Hence,  $f_1(x) < f_2(x)$  for all  $x > c$ , which contradicts  $f_1'(x) - f_2'(x) \rightarrow 0$  as  $x \rightarrow \infty$  for the same reason.

By what was said above, the graphs of  $f_1(x)$  and  $f_2(x)$  cannot cross, but they must, since these functions integrate to the same number and have the same limit at infinity. Hence,  $\Lambda_1(0) \leq \Lambda_2(0)$  is impossible when  $\sigma^2$  and  $N_a$  are the same in the two models.

(2) Now since  $\kappa_1 < \kappa_2$  and  $\Lambda_1(0) > \Lambda_2(0)$ , the two generalized hazard functions  $\Lambda_1(x)$  and  $\Lambda_2(x)$  satisfy the conditions of [Proposition 8](#). From this it follows that  $Kurt_1(\Delta p) > Kurt_2(\Delta p)$ .

(3) The flexibility index for the power-plus case is given by

$$\mathcal{F} = \int_{-\infty}^{\infty} f(x)(\Lambda(x) + \Lambda'(x)x)dx = (1 + \nu)N_a - \nu\Lambda(0)$$

Since the two models deliver the same  $N_a$  and  $\nu$  is fixed, the one with a greater intercept has a smaller  $\mathcal{F}$ . This completes the proof.  $\square$

**Proof.** (of [Proposition 15](#)). We will make two observations, one about  $\Lambda$  and one about  $F_n$  required to establish the two main results of the proposition. Then we will use [Lemma 3](#) finish the proof.

The first observation is that the symmetry of  $\Lambda$  around  $x = 0$  implies that all the odd numbered derivatives evaluated at  $x = 0$  of  $\Lambda$  are equal to zero.

The second observation is a property of the function  $F_n(x)$  generated by the recursion in [equation \(60\)](#), which can be written as:

$$F_n(x) = \tilde{F}_n(x) - \left(\frac{\sigma^2}{2}\right)^{n-1} \frac{\partial^{2n-2}\Lambda(x)}{\partial x^{2n-2}}$$

where  $\tilde{F}_n(\cdot)$  depends only on the level of  $\Lambda(\cdot)$  and at most the first  $2n - 1$  derivatives of  $\Lambda(\cdot)$ , evaluated at  $x$ . This property can be established by induction. It is true for  $F_1(x) = -\Lambda(x)$  for  $n = 1$ . Now assume it holds for  $n$ , and we will show that it holds  $n + 1$ . To do so we compute  $F_{n+1}$  according to the recursion. On this computation, the first term is the product of  $\sigma^2/2$  times the sum of the second derivative of  $\tilde{F}_n(x)$  with respect to  $x$  and of the second derivative of  $-(\sigma^2/2)^{n-1} \partial^{2n-2}\Lambda(x)/\partial x^{2n-2}$  with respect to  $x$ . The remaining term,  $-\Lambda(x)F_n(x)$ , involves no derivatives. This finishes the induction step, and thus established the desired result for  $F_n$ .

1. If we know the function  $\Lambda(x)$ , then we can recursively compute  $F_n(x)$  from [equation \(60\)](#). Evaluating this expressions at  $x = 0$  and using [equation \(61\)](#) we obtain all the derivatives of  $S$  evaluated at  $t = 0$ . In particular, these expressions only use the level and the even derivatives of  $\Lambda$  evaluated at  $x = 0$ . If  $S$  is analytical, the expansion of  $S$  at  $t = 0$  gives the values everywhere.

2. If we know the function  $S$ , we can take all its derivatives at  $t = 0$ , and by [equation \(61\)](#) we know all the values of  $F_n(0)$  for  $n \geq 1$ . Next we argue that the recursion in [equation \(60\)](#) evaluated at  $x = 0$ , will give us all the even order derivatives of  $\Lambda$  evaluated at  $x = 0$ . Since  $\Lambda$  is symmetric, all the derivatives of odd order, evaluated at  $x = 0$ , so we are only interested in the even derivatives at  $x = 0$ . Next we argue that, algorithmically, we can recursively recover the derivatives up to order  $2n - 2$  with  $\{F_n(0)\}$  for  $j = 1, \dots, 2n - 2$ . First we note that  $\Lambda(0)$  and  $\Lambda''(0)$  are given by  $F_1(0)$  and  $F_2(0)$ . Now assume we know all the derivatives up to order  $2n - 2$ . Then, given the value of  $\partial^{n+1}S(0)/\partial t^{n+1} = F_{n+1}(0)$ , the known values for  $\Lambda(0)$ ,  $F_n(0)$ , and  $\sigma^2$ , using the recursion we obtain the implied value for  $\partial_{xx}F_n(0)$ . Using that  $F_n$  depend at most on  $2n - 2$  derivatives of  $\Lambda$ , as well as the particular expression derived above, we obtain the value of  $\partial^{2n}\Lambda(0)/\partial t^{2n}$ . This completes the induction step, and hence establishes the desired property, and hence the level and all the derivatives of  $\Lambda$  at  $x = -0$  have been recovered. Finally, since  $\Lambda$  is assumed to be analytical, an expansion around  $x = 0$  gives its value at any other  $x$ .

This completes the proof.  $\square$

## C Estimation and measurement issues

In this appendix we present our estimation algorithm and some additional results. The next proposition shows that if we have a sample with mixed  $N$  different type of products all with the same kurtosis but with different variance, then the kurtosis of the price changes of such a mixture is higher than the kurtosis for each of them.

**PROPOSITION 11.** Assume that  $\Delta p$  is a mixture of  $N$  distributions, with weights  $\{\omega_j\}_{j=1}^N$ . Assume that for each distribution  $j$ , price changes have the same kurtosis  $K$ , but they may have different variance  $V_j$ . Then

$$Kurt(\Delta p) = \frac{\sum_j \omega_j K V_j^2}{\left[\sum_j \omega_j V_j\right]^2} = K \frac{\sum_j \omega_j V_j^2}{\left[\sum_j \omega_j V_j\right]^2} = K \frac{\sum_j J(V_j)\omega_j}{J\left(\sum_j V_j\omega_j\right)} \geq K$$

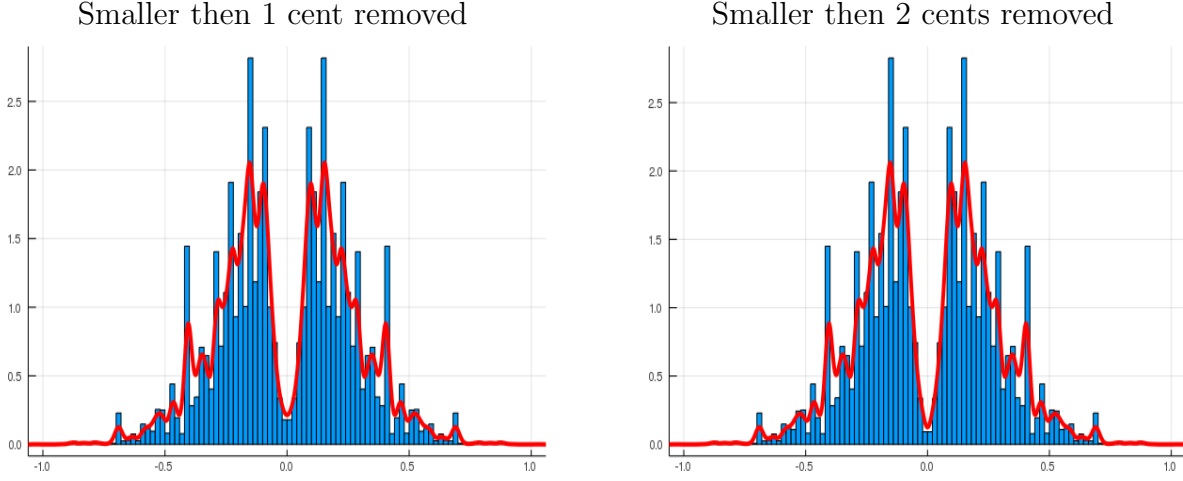
with strict inequality if the distribution of  $\{V_j\}_{j=1}^N$  is not degenerate, since  $J(V) = V^2$  is a strictly convex function.

The proof is contained in the statement of the proposition. Next, we plot the symmetrized histograms with fitted densities for two data cleaning procedures: the one that eliminates price changes smaller than 2 cents in absolute value, and the one eliminating those smaller than 1 cent in absolute value. The distributions are very close, with immaterial differences in the bars around zero.

We use the method of moments to estimate the mixture of two Gamma distributions with the parameters  $\omega$  (the weight),  $(\alpha_1, \beta_1)$  and  $(\alpha_2, \beta_2)$ . The moments of  $|\Delta \tilde{p}_t|$  we use are denoted by  $\gamma_{j,k}$ :

$$\gamma_{j,k} = \frac{\mathbb{E}[|\Delta \tilde{p}_t|^{j+k}]}{\mathbb{E}[|\Delta \tilde{p}_t|^j]\mathbb{E}[|\Delta \tilde{p}_t|^k]}$$

Figure 6: Distribution of price changes in a narrow category



Pooling all products for category 561 “Non-durable household goods”

For a mixture of two Gamma distributions with the weight  $\xi$  on the first one,

$$\gamma_{j,k} = \frac{\left[ \beta_2^{j+k} \omega \frac{\Gamma(\alpha_1 + j + k)}{\Gamma(\alpha_1)} + \beta_1^{j+k} (1 - \omega) \frac{\Gamma(\alpha_2 + j + k)}{\Gamma(\alpha_2)} \right]}{\left[ \beta_2^j \omega \frac{\Gamma(\alpha_1 + j)}{\Gamma(\alpha_1)} + \beta_1^j (1 - \omega) \frac{\Gamma(\alpha_2 + j)}{\Gamma(\alpha_2)} \right]} \left[ \beta_2^k \omega \frac{\Gamma(\alpha_1 + k)}{\Gamma(\alpha_1)} + \beta_1^k (1 - \omega) \frac{\Gamma(\alpha_2 + k)}{\Gamma(\alpha_2)} \right] \quad (71)$$

Using these moments allows us to recover  $\omega$ ,  $\alpha_1$ ,  $\alpha_2$ , and the ratio  $\beta_1/\beta_2$ . The exact values of  $\beta_1$  and  $\beta_2$  are pinned down by the normalization  $\mathbb{E}[|\Delta \tilde{p}_t|] = 1$ . To estimate  $\gamma_{j,k}$ , we rely on [Proposition 5](#):

$$\frac{\mathbb{E}[|\Delta \tilde{p}_t|^{j+k}]}{\mathbb{E}[|\Delta \tilde{p}_t|^j] \mathbb{E}[|\Delta \tilde{p}_t|^k]} = \frac{\mathbb{E}[|\Delta p_{it}|^{j+k}]}{\mathbb{E}[|\Delta p_{it}|^j] \mathbb{E}[|\Delta p_{is}|^k]}$$

For all seven product categories, we get four moments ( $\hat{\gamma}_{1,1}$ ,  $\hat{\gamma}_{2,1}$ ,  $\hat{\gamma}_{3,1}$ , and  $\hat{\gamma}_{3,2}$ ) from the data and solve the system of four analogs of [equation \(71\)](#). We minimize the sum of deviations squared with equal weights. The results are presented in [Table 2](#).

Category	$\hat{\gamma}_{11}$	$\hat{\gamma}_{21}$	$\hat{\gamma}_{31}$	$\hat{\gamma}_{32}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1/\hat{\beta}_2$	$\hat{\omega}$	$\hat{\alpha}_{22}$
111	1.248	1.406	1.507	1.787	2.099	12.190	228.677	0.161	4.248
119	1.282	1.507	1.702	2.381	1.058	6.012	91.439	0.109	3.747
1212	1.242	1.476	1.786	2.9230	0.599	3.873	73.414	0.000	4.151
122	1.243	1.397	1.508	1.903	1.848	9.779	173.048	0.131	4.460
118	1.289	1.539	1.777	2.552	3.123	9.836	0.628	0.580	3.610
117	1.281	1.511	1.721	2.484	0.967	5.442	84.154	0.089	3.801
561	1.216	1.394	1.586	2.271	0.998	5.783	103.470	0.031	4.782

Table 2: Moments taken from the data and the estimated parameters

Specializing to the case with a single Gamma distribution  $\omega = 1$  allows us to recover the expressions for  $\alpha$  in closed form. Consider  $\gamma_{j,1}$  for some  $j$ :

$$\gamma_{j,1} = \frac{\Gamma(\alpha + j + 1)\Gamma(\alpha)}{\Gamma(\alpha + j)\Gamma(\alpha + 1)} = 1 + \frac{j}{\alpha}$$

Hence,

$$\alpha = \frac{j}{\gamma_{j,1} - 1} \quad (72)$$

Since we attach particular importance to the kurtosis, we would also like to use  $\gamma_{2,2}$ :

$$\gamma_{j,2} = \frac{\Gamma(\alpha + j + 2)\Gamma(\alpha)}{\Gamma(\alpha + j)\Gamma(\alpha + 2)} = \frac{(\alpha + j + 1)(\alpha + j)}{(\alpha + 1)\alpha} = \left(1 + \frac{j + 1}{\alpha}\right) \frac{\gamma_{j,1}}{\gamma_{1,1}}$$

This leads to

$$\alpha = \frac{(j + 1)\gamma_{j,2}}{\gamma_{j,2}\gamma_{1,1} - \gamma_{j,1}} \quad (73)$$

Notice that  $\beta$ , the scale of the distribution, drops out from these expressions, because  $\gamma_{j,k}$  are dimensionless moments. We use a linear combinations of expressions in [equation \(72\)](#) and [equation \(73\)](#) with  $\hat{\gamma}_{j,1}$  for  $j \in \{1, 2\}$  and  $\hat{\gamma}_{22}$  as estimators of  $\alpha$ . Consistency requires the weights of the combinations to sum to one, and we make them inversely proportional to the bootstrapped variance of the estimators of summands. The estimates are presented in the last column of [Table 2](#): the estimate  $\hat{\alpha}_{22}$  is constructed from  $\hat{\gamma}_{11}$ ,  $\hat{\gamma}_{21}$ , and  $\hat{\gamma}_{22}$ .

In [Table 3](#) we present some additional statistics. First, we tabulate skewness of the distribution of price changes to show that the distributions are close to symmetric. Then, we contrast the estimates of the Kurtosis with the full sample and with the first two price changes only. The difference between them is suggestive of a strong correlation between consecutive price changes (squared), and of a weaker correlation between distant price changes. As can be seen from [equation \(27\)](#), how much the underlying Kurtosis is different from that of the pooled distribution (without accounting for product heterogeneity) increases with this correlation. The implied correlation and the coefficient of variation (present in [equation \(27\)](#) as well) are tabulated in the remaining two columns.

Now we present the estimation procedure to recover the flow cost function from [Section 2.2](#). The model in this section permits  $\Lambda$  to be unbounded. We take advantage of that and work with a power hazard  $\Lambda(x) = \kappa x^\nu$ . This form of  $\Lambda$  gives rise to a specific functional form of  $Q$ . We compute the moments of  $Q$  as functions of  $(\kappa, \nu)$  and then estimate them using the method of moments.

Suppose  $\Lambda(x) = \kappa x^\nu$ . Denote  $\rho = 2\kappa/\sigma^2$ . The corresponding density of price gaps has to obey a Kolmogorov forward equation that has the form

$$\rho x^\nu f(x) = f''(x)$$



Category	Skewness	Kurtosis	Kurtosis ( $t = 1, 2$ )	Implied Correlation	$CV(\Delta\tilde{p}_t)$
111	-0.121	1.656 (0.065)	1.426 (0.071)	0.440	1.555
119	0.011	1.955 (0.050)	1.288 (0.042)	0.339	1.683
1212	-0.020	2.051 (0.162)	1.710 (0.186)	0.284	1.589
122	-0.025	1.677 (0.051)	1.189 (0.019)	0.390	1.398
118	-0.012	2.044 (0.118)	1.663 (0.150)	0.295	1.620
117	-0.004	1.989 (0.047)	1.422 (0.089)	0.303	1.577
561	-0.006	1.778 (0.133)	1.403 (0.066)	0.374	1.524

Table 3: Additional statistics

With  $X = \infty$ , the solution is

$$f(x) = \frac{x^{1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)}{2\int_0^\infty x^{1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)dx}$$

The distribution of price changes is then given by

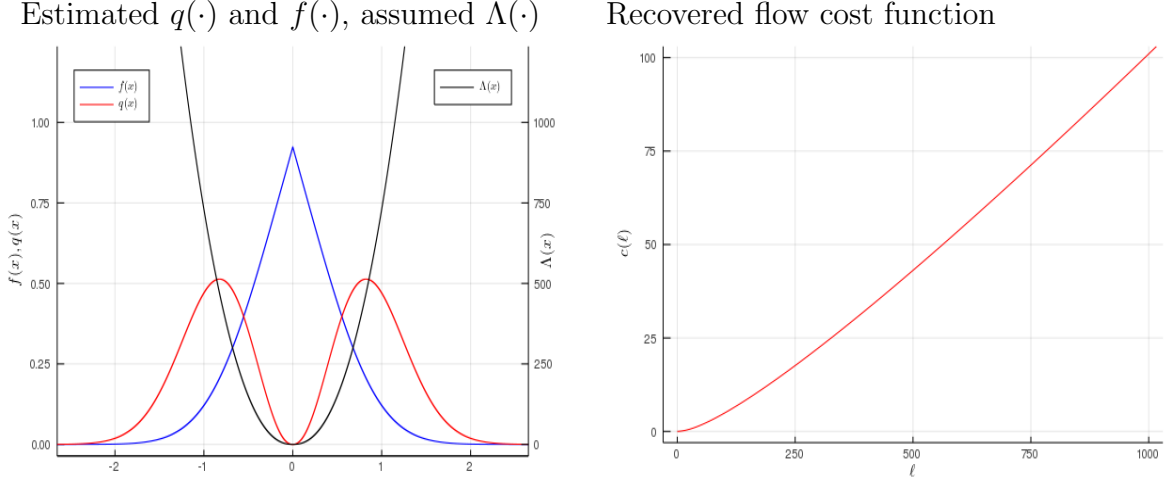
$$q(-x) = \frac{\kappa x^\nu f(x)}{N_a} = \frac{\kappa x^{\nu+1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)}{2N_a\int_0^\infty x^{1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)dx}$$

Since  $\mathbb{V}[\Delta\tilde{p}_t] = 1$ , we have  $\sigma^2 = N_a$ , so

$$q(-x) = \frac{\kappa x^\nu f(x)}{N_a} = \frac{\rho x^{\nu+1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)}{4\int_0^\infty x^{1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)dx}$$

This has to be a probability distribution, so it integrates two one. We also have the moment

Figure 7: Estimated distribution of price changes and implied cost function



condition  $\mathbb{E}[(\Delta\tilde{p}_t)^4] = Kurt(\Delta\tilde{p}_t)$ . Writing the two restrictions in a convenient form,

$$\int_0^{\infty} (\rho x^{\nu} - 2)x^{1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)dx = 0$$

$$\int_0^{\infty} (\rho x^{\nu+4} - 2Kurt(\Delta\tilde{p}_t))x^{1/2}\mathcal{K}_{1/(\nu+2)}\left(\frac{2\sqrt{\rho}}{\nu+2}x^{(\nu+2)/2}\right)dx = 0$$

From these two relations we can get  $(\hat{\rho}, \hat{\nu})$ . After that, using  $\sigma^2 = N_a$ , we can recover  $\hat{\kappa}$ :

$$\hat{\kappa} = \frac{\hat{\rho}N_a}{2}$$

The system of two restrictions can be solved exactly, and the model is just identified. The results for the category 561 ("non-durable household goods") are presented on [Figure 7](#). The estimated parameters are  $\hat{\nu} = 2.285$  and  $\hat{\kappa} = 30.747$ , corresponding to the Kurtosis 1.64, slightly below the quadratic case.

## D Flexibility Index: scope and limitations

The impulse response function (IRF) of the aggregate price level after a shock  $\delta$  can be written as

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

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

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

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:

$$\partial_t f(x, t) = -\Lambda(x) f(x, t) + \frac{\sigma^2}{2} \partial_{xx} f(x, t) \text{ for all } x \in [-X, X] \text{ and } t \geq 0, \quad (74)$$

$$f(X, t) = f(-X, t) = 0 \text{ for all } t > 0, \text{ and } f(x, 0) = f_0(x) \text{ for all } x \in [-X, X] \quad (75)$$

The initial jump is given by

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

The initial distribution  $f_0$  that we consider is a uniform shift by  $\delta$  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  $\delta$ . 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 \(74\)](#) and [equation \(75\)](#). 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 12.** Let  $\Omega$  and  $\omega$  be the jump and flow values of the IRF of prices at  $t = 0$ . Let  $X < \infty$ , let  $\Lambda$  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  $\Lambda$ .

Because of this result we will move to analyze the flexibility index for models with  $X = \infty$ , where it is finite. We will do so for a family of hazard functions which is a slight generalization of the one treated in [Section 5.1](#).

## D.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  $\Lambda(0)$ ,  $\kappa$ , and  $\nu$ , the other parameter of the model is  $\sigma^2$ . We introduce the parameter  $\eta$  and let  $\alpha$  be the adjusted intercept:

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

The quadratic case is  $\nu = 2$  and  $\alpha = 0$ . This adjusted intercept measures the relative magnitude of  $\Lambda(0)$  and the slope  $\kappa$ , 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  $\alpha$ .

**PROPOSITION 13.** Fix  $\sigma^2$  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

$$N_a = \frac{\eta^2 \sigma^2}{2} \tilde{N}(\nu, \alpha) \quad , \quad \frac{Kurt(\Delta p)}{6N_a} = \frac{1}{\eta^2 \sigma^2} \tilde{K}(\nu, \alpha) \quad , \quad \mathcal{F} = \frac{\eta^2 \sigma^2}{2} (\tilde{N}(\nu, \alpha)(1 + \nu) - \nu \alpha)$$

where  $\tilde{N}(\nu, \alpha)$  and  $\tilde{K}(\nu, \alpha)$  only depend on  $\nu$  and  $\alpha$ ;  $\tilde{N}(0, \alpha) \equiv 1 + \alpha$ , 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:

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

If two models have the same  $(\nu, \alpha)$ , the density of price changes in one is a rescaling of that in the other. This implies that kurtosis (and other scale-free statistics) is the same. If  $\eta$  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 a 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 14.** Assume that  $\Lambda$  is given by a power-plus function. Fix  $(\nu, \sigma^2)$  and take two different power plus generalized hazard functions  $\Lambda_1$  and  $\Lambda_2$ . 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 the right panel of [Figure 5](#) we display several economies with the same adjustment frequency  $N_a$ , and with the same Flexibility Index  $\mathcal{F}$ , that feature 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  $\nu$  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  $\nu$  increases from 2 to 20, as plotted in the figure. The slope of the impulse response at  $t = 0$  does not capture the area under it in a reliable way.

In the left panel of [Figure 5](#) we take two examples, 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 in the right panel of the figure. 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.

## E 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  $\Lambda$ . When  $X = \infty$ , the Survival function is given by

$$S(t) = \mathbb{E} \left[ e^{-\int_0^t \Lambda(x(s)) ds} \mid x(0) = 0 \right] \text{ for all } t \geq 0 \quad (76)$$

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  $\sigma^2$ . 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:

$$S(t) = \frac{e^{-t\Lambda(0)}}{\left( \cosh \left( t\sqrt{2\kappa\sigma^2} \right) \right)^{\frac{1}{2}}} \text{ and } h(t) = \Lambda(0) + \sqrt{\kappa \frac{\sigma^2}{2}} \tanh \left( t\sqrt{2\kappa\sigma^2} \right) \text{ for all } t \geq 0 \quad (77)$$

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 3.** Fix a value of  $\sigma^2 > 0$ , and assume that  $X = \infty$ . Assume that  $S$  is related to  $\Lambda$  by equation (76). The derivatives of the Survival function  $S$  a time  $t = 0$  and the derivatives of  $\Lambda$  at  $x = 0$  are related by the recursively generated functions  $\{F_n\}$  as follows:

$$\begin{aligned} \left. \frac{\partial^n S(t)}{\partial t^n} \right|_{t=0} &= F_n(0) \text{ and all } n = 1, 2, \dots \text{ where } F_n(\cdot) \text{ are given by} \\ F_{n+1}(x) &= \frac{\sigma^2}{2} \frac{\partial^2 F_n(x)}{\partial x^2} - \Lambda(x) F_n(x) \text{ for all } x \in \mathbb{R} \text{ and } n = 1, 2, \dots \text{ and} \\ F_1(x) &= -\Lambda(x) \text{ for all } x \in \mathbb{R} \end{aligned}$$

**Lemma 3** is the base of an algorithm to compute the derivatives of  $S$  at  $t = 0$  given  $\Lambda$  and the derivatives of  $\Lambda$  at  $x = 0$  given  $S$ . Using this lemma, we obtain the main result of this section:

**PROPOSITION 15.** Assume that  $\sigma^2 > 0$ ,  $X = \infty$ , and  $\Lambda$  satisfies **Assumption 1**. Let  $S$  be the Survival function of  $\Lambda$ , as in **equation (76)**. If the generalized hazard function  $\Lambda$  is analytical, then the Survival function  $S$  uniquely identifies  $\Lambda$ . Likewise, if the Survival function  $S$  is analytical, then the generalized hazard function  $\Lambda$  uniquely identifies  $S$ .

As remarked before, **Lemma 3** gives an algorithm to recursively compute an expansion of  $S$  based on the derivatives of  $\Lambda$ , or an expansion of  $\Lambda$  based on the derivatives of  $S$ . An implication of **Lemma 3** and **Proposition 15** 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.  $x = 0$ :

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

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

These formulas give a simple connection between the local behavior of  $\Lambda$  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 **Theorem 2**, in this case we use much more restrictive conditions for  $\Lambda$ , and obtain a more cumbersome representation — an infinite expansion instead of a closed-form expression involving an integral. In spite of this **Theorem 2** and **Proposition 15** have the same flavor: they show that if  $\Lambda$  is analytical and  $X = \infty$ , then  $\Lambda$  can be fully identified either using the information contained in the Survival function, i.e duration on price changes, and  $\sigma^2$ , which can be recovered from  $N_a$  and the variance of price changes with **equation (13)**. 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.

## Online Appendix:

# The Macroeconomics of Sticky Prices with Generalized Hazard Functions

Fernando Alvarez, Francesco Lippi, and Aleksei Oskolkov

## G Solution for the firm's alternative setup of **Section 2.2**

The first order condition for choice of  $\ell$  in [equation \(9\)](#) are:

$$c'_-(\ell^*(x)) \leq v(x) - v(0) \leq c'_+(\ell^*(x)) \text{ for all } x$$

where  $\ell^*(x)$  denotes the optimal policy, and where  $c'_-(\cdot)$  and  $c'_+(\cdot)$  denote the right and left derivatives of  $c$ . As in the previous case, we have that if  $\Psi < \infty$  there is a barrier  $X < \infty$  for which:  $v(X) = v(0) + \Psi$  and  $v'(X) = 0$ . Finally, by the same reasons as before, we have symmetry, i.e.  $v(x) = v(-x)$ , and  $\ell^*(x) = \ell^*(-x)$ . As before we can summarize the decision rule of the firm for  $x \in (-X, X)$  with a generalized hazard function:

$$\Lambda(x) = \ell^*(x) \text{ for all } x$$

To simplify the discussion, next we describe the case of a cost  $c$  that is continuously differentiable and strictly convex, where we simply have:

$$c'(\ell^*(x)) = v(x) - v(0) \text{ and } \Lambda(x) = (c')^{-1}(v(x) - v(0)) \text{ for all } x$$

We note that since  $v(x)$  is strictly increasing in  $x$  for  $x \in (0, X)$ , and  $c(\ell)$  is convex, then  $\Lambda(x)$  must also be increasing in  $x$  for  $x \in (0, X)$ .

Replacing  $\ell^*$  into the HBJ equation we obtain:

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

Let us assume that the cost function  $c$  has a continuous derivative. Defining, as before  $U(x) = v(x) - v(0)$ , with  $u = U' = v'$ , we can differentiate the HBJ equation in  $x \in (0, X)$ , and use the envelope to obtain:

$$[r + \Lambda(x)]u(x) = 2Bx + \frac{\sigma^2}{2}u''(x).$$

Using the boundaries  $u(0) = u(X) = 0$ , and the logic used in the proof of [Theorem 1](#) it is then straightforward to solve for  $\Psi = \int_0^X u(z)dz$ . The marginal cost of switching intensity is recovered using  $\Lambda(x) = (c')^{-1}(U(x))$ . The cost function itself, just as the value function, is only determined up to an additive constant, which is straightforward to verify from [equation \(9\)](#).

## H Properties of Distribution of Menu Cost

In this appendix we note that the posited behavior of  $\Lambda$  in a neighbourhood of  $x = 0$  or  $x = |X|$  determines whether the underlying density  $g$  is bounded. It is shown in [equation \(2\)](#) that the hazard function inherits the shape of the value function because of the underlying optimization: when the firm draws a fixed cost, what matters is how the value of the draw compares to the gains from adjustment. Taking a first order derivative of [equation \(2\)](#) gives

$$\Lambda'(x) = \kappa g(v(x) - v(0)) v'(x) \quad (78)$$

A bounded density  $g$  would make  $\Lambda'(x)$  have zero limits at  $x = 0$  and  $x = X$  because of the smooth-pasting conditions on  $v(x)$  at these points. Thus, if the hazard function of the inverse problem (the one that solves for  $g$  given  $\Lambda$ ) is not flat at 0 or  $\Psi$ , then the density  $g$  must be diverging. We formalize this observation next:

**COROLLARY 5.** Let  $\varepsilon > 0$  and suppose  $\Lambda'(x)$  is bounded away from zero for  $x \in (0, \varepsilon)$ . Then  $g(\psi)$  is unbounded on any  $(0, \psi)$ . Likewise, if  $\Lambda'(x)$  is bounded away from zero for  $x \in (X - \varepsilon, X)$  then  $g(\psi)$  is unbounded on any  $(\psi, \Psi)$ .<sup>27</sup>

We can also characterize the behavior of the density  $g$  around  $\psi = 0$  for different forms of  $\Lambda$  around  $x = 0$ . Take the limiting elasticity of the hazard

$$\nu = \lim_{x \downarrow 0} \frac{x\Lambda'(x)}{\Lambda(x) - \Lambda(0)}$$

If  $\Lambda$  is symmetric and smooth, it admits a quadratic approximation close to zero, and  $\nu = 2$ . Interestingly, deviations from  $\nu = 2$  imply irregular behavior of  $g$ . [Theorem 1](#) states that

$$g(x) = \frac{\Lambda'(x)}{\kappa u(x)}$$

But  $u(x)$  converges to zero as  $x \rightarrow 0$ , so the limit is tricky. To resolve the indeterminacy, notice that  $u(x)$  goes to zero linearly, since  $u''(0) = 0$  (immediate from the [equation \(4\)](#) defining  $u(x)$  in [Lemma 1](#)). Thus whether the limit is (i) zero, (ii) positive and finite, or (iii) infinite, depends respectively on whether  $\Lambda'(x)$  goes to zero (i) faster than a linear rate ( $\nu > 2$ ), (ii) at a linear rate ( $\nu = 2$ ), (iii) slower than a linear rate ( $\nu < 2$ ). We can formalize this:

**COROLLARY 6.** Suppose that  $\Lambda'(x)$  and  $g(\psi)$  both have (possibly infinite) right limits at zero. Then  $\lim_{\psi \downarrow 0} g(\psi) = \infty$  for  $\nu < 2$ ,  $0 < \lim_{\psi \downarrow 0} g(\psi) < \infty$  for  $\nu = 2$ , and  $\lim_{\psi \downarrow 0} g(\psi) = 0$  for  $\nu > 2$ .

This corollary states that a quadratic hazard function implies a density of  $\psi$  that is positive and finite around  $\psi = 0$ . If the leading term in  $\Lambda(x)$  is higher than quadratic ( $\nu > 2$ ) then the density must be zero, meaning that  $G$  is flat close to  $\psi = 0$ . A hazard function with a leading term  $\nu < 2$  implies a distribution of  $\psi$  with density that is diverging around  $\psi = 0$ .

---

<sup>27</sup>Since  $\Lambda(x)$  is symmetric, to be smooth at zero it has to have  $\Lambda'(0) = 0$ . The proof is done by standard analysis.



# I Alternative Normalization

We consider an alternative normalization to one used in [Proposition 2](#). This normalization requires that  $X < \infty$ . For a triplet  $\{\sigma^2, X, \Lambda\}$  we can define a new problem represented by pair  $\{\rho, \hat{\Lambda}\}$  where  $\hat{\Lambda} : (-1, 1) \rightarrow \mathbb{R}_+$  and where  $\rho$  is a scalar defined as follows:

$$\hat{\Lambda}(z) = \frac{\Lambda(zX)}{\kappa} \text{ for all } z \in [-1, 1] \text{ and } \rho = \frac{2\kappa X^2}{\sigma^2} \quad (79)$$

Note that this is the normalization used in [Proposition 2](#) with  $b = 1/X$ . This is a slight generalization of [Proposition 2](#), in that it allows to have some comparative static with respect to  $\kappa$ .

Given the triplet  $\{\sigma^2, X, \Lambda\}$  we can solve for  $f$  as indicated in [equation \(10\)](#). And given the pair  $\{\rho, \hat{\Lambda}\}$  we can solve for the probability density  $\hat{f}$ , using a change of variables:

$$\hat{f}(z) \equiv f(zX) X \text{ for all } z \in [-1, 1]$$

We note that  $\hat{f}$  satisfies the

$$\hat{\Lambda}(z) \rho \hat{f}(z) = \hat{f}''(z) \text{ for all } z \in [-1, 1] \text{ and } z \notin \mathbb{Z} \quad (80)$$

where  $z \in \mathbb{Z}$  if  $z = x/X$  and  $x \in \mathbb{J}$ . Moreover, the density  $\hat{f}$  must satisfy

$$\hat{f}(1) = \hat{f}(-1) = 0 \text{ and } \int_{-1}^1 \hat{f}(z) dz = 1$$

**LEMMA 4.** Consider two triplets  $\{\sigma, X, \Lambda\}$  such that both generate the function  $\hat{\Lambda}(\cdot)$  and the parameter  $\rho$  by using [equation \(79\)](#). The two triplets have the same Kurtosis of price changes  $Kurt(\Delta p)$  and the same share of adjustment in the interior  $s$ . Furthermore,

$$\begin{aligned} N_a &= \frac{\sigma^2}{X^2} \hat{n}(\rho) \\ \frac{Kurt(\Delta p)}{6N_a} &= \frac{X^2}{\sigma^2} \frac{\hat{m}(\rho)}{6} \\ s &= \hat{s}(\rho) \end{aligned}$$

where  $\hat{n}(\rho)$ ,  $\hat{m}(\rho)$  and  $\hat{s}(\rho)$  only depend on  $\hat{\Lambda}(\cdot)$  and  $\rho$ . Moreover,  $\hat{n}(\cdot)$  is increasing in  $\rho$ ,  $\hat{m}(\cdot)$  is decreasing in  $\rho$ ,  $\hat{s}(\cdot)$  is increasing in  $\rho$ , and  $\hat{n}(0) = \hat{m}(0) = \hat{s}(0) = 1$ .

## J Functional forms of $\langle f(x), m(x), \mathcal{T}(x) \rangle$ for integer $\nu$

The invariant density  $f$  has to be symmetric around  $x = 0$ , and has to satisfy:

$$\Lambda(x)f(x) = \frac{\sigma^2}{2}f''(x) \text{ for all } x \in [0, X], \quad (81)$$

$$\frac{1}{2} = \int_0^X f(x)dx \text{ and } f(X) = 0. \quad (82)$$

The contribution of an individual firm to the IRF is antisymmetric around  $x = 0$  and satisfies the following:

$$\Lambda(x)m(x) = -x + \frac{\sigma^2}{2}m''(x) \text{ for all } x \in [0, X], \quad (83)$$

$$m(0) = m(X) = 0. \quad (84)$$

Fianlly,  $\mathcal{T}(x)$  is symmetric around  $x = 0$  and satisfies

$$\Lambda(x)\mathcal{T}(x) = 1 + \frac{\sigma^2}{2}\mathcal{T}(x) \text{ for all } x \in [0, X],$$

$$\mathcal{T}(X) = 0 \text{ and } \mathcal{T}'(0) = 0.$$

The latter equality is a consequence of  $\mathcal{T}(\cdot)$  being continuously differentiable ay zero and antisymmetric.

Denote  $y = \sigma^2/2a$ . We will assume that the functions of interest are analytical, so we can write them as:

$$f(x) = \sum_{k=0}^{\infty} \alpha_k x^k \text{ for } x \in [0, X]$$

$$m(x) = \sum_{k=0}^{\infty} \beta_k x^k \text{ for } x \in [0, X]$$

$$\mathcal{T}(x) = \sum_{k=0}^{\infty} \gamma_k x^k \text{ for } x \in [0, X]$$

so that, in particular,  $\gamma_0 = \mathcal{T}(0)$ . Inserting these expressions into the equations above and using the functional form for  $\Lambda(\cdot)$ , we obtain:

$$a \sum_{k=0}^{\infty} \alpha_k x^{k+\nu} = \frac{\sigma^2}{2} \sum_{k=2}^{\infty} \alpha_k k(k-1)x^{k-2} \quad \text{for } x \in [0, X]$$

$$a \sum_{k=0}^{\infty} \beta_k x^{k+\nu} = \frac{\sigma^2}{2} \sum_{k=2}^{\infty} \beta_k k(k-1)x^{k-2} - x \quad \text{for } x \in [0, X]$$

$$a \sum_{k=0}^{\infty} \gamma_k x^{k+\nu} = \frac{\sigma^2}{2} \sum_{k=2}^{\infty} \gamma_k k(k-1)x^{k-2} + 1 \quad \text{for } x \in [0, X]$$

Matching the coefficient of each of the powers of  $x$  we have

$$\begin{aligned}\alpha_k &= y(k + \nu + 2)(k + \nu + 1)\alpha_{k+\nu+2} \text{ for } k \geq 0 \\ \beta_k &= y(k + \nu + 2)(k + \nu + 1)\beta_{k+\nu+2} \text{ for } k \geq 0 \\ \gamma_k &= y(k + \nu + 2)(k + \nu + 1)\gamma_{k+\nu+2} \text{ for } k \geq 0\end{aligned}$$

The symmetry and smoothness properties also lead to

$$\beta_0 = \beta_2 = \gamma_1 = 0 \tag{85}$$

Relabelling the coefficients, we can write the sums as

$$\begin{aligned}f(x) &= \alpha_0 \left( 1 + \sum_{j=1}^{\infty} \xi_{p,j} y^{-j} x^{j(\nu+2)} \right) + \alpha_1 x \left( 1 + \sum_{j=1}^{\infty} \eta_{p,j} y^{-j} x^{j(\nu+2)} \right) \\ m(x) &= \beta_1 x \left( 1 + \sum_{j=1}^{\infty} \xi_{m,j} y^{-j} x^{j(\nu+2)} \right) + \beta_3 x^3 \left( 1 + \sum_{j=1}^{\infty} \eta_{m,j} y^{-j} x^{j(\nu+2)} \right) \\ \mathcal{T}(x) &= \gamma_0 \left( 1 + \sum_{j=1}^{\infty} \xi_{t,j} y^{-j} x^{j(\nu+2)} \right) + \gamma_2 x^2 \left( 1 + \sum_{j=1}^{\infty} \eta_{t,j} y^{-j} x^{j(\nu+2)} \right)\end{aligned}$$

Here the coefficients  $\xi_{\cdot,j}$  and  $\eta_{\cdot,j}$  are given by

$$\begin{aligned}\xi_{p,j} &= \prod_{i=1}^j \frac{1}{i(\nu+2)(i(\nu+2)-1)} & \eta_{p,j} &= \prod_{i=1}^j \frac{1}{i(\nu+2)(i(\nu+2)+1)} \\ \xi_{m,j} &= \prod_{i=1}^j \frac{1}{i(\nu+2)(i(\nu+2)+1)} & \eta_{m,j} &= \prod_{i=1}^j \frac{1}{(i(\nu+2)+2)(i(\nu+2)+3)} \\ \xi_{t,j} &= \prod_{i=1}^j \frac{1}{i(\nu+2)(i(\nu+2)-1)} & \eta_{t,j} &= \prod_{i=1}^j \frac{1}{(i(\nu+2)+1)(i(\nu+2)+2)}\end{aligned}$$

Now define the following parameter:

$$Z = \frac{X^{\nu+2}}{y} = 2aX^\nu \frac{X^2}{\sigma^2} = 2\kappa\mathcal{T}_0$$

It will be useful in pinning down the coefficients. Here  $\tilde{\Lambda}$  is the left limit of the hazard rate when  $x$  approaches  $X$ , and  $\mathcal{T}_0$  is the expected time to adjustment when  $a = 0$ .

Consider first  $f(\cdot)$ . The boundary condition is

$$\begin{aligned}0 = f(X) &= \alpha_0 \left( 1 + \sum_{j=1}^{\infty} \xi_{p,j} y^{-j} X^{j(\nu+2)} \right) + \alpha_1 x \left( 1 + \sum_{j=1}^{\infty} \eta_{p,j} y^{-j} X^{j(\nu+2)} \right) \\ &= \alpha_0 \left( 1 + \sum_{j=1}^{\infty} \xi_{p,j} Z^j \right) + \alpha_1 X \left( 1 + \sum_{j=1}^{\infty} \eta_{p,j} Z^j \right)\end{aligned}$$

Define additionally  $\xi_{\cdot,0} = \eta_{\cdot,0} = 1$ . The condition that  $f(\cdot)$  is a density states

$$\begin{aligned} \frac{1}{2} &= \int_0^X f(x)dx = \alpha_0 X \left( 1 + \sum_{j=1}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+1} \right) + \alpha_1 X^2 \left( \frac{1}{2} + \sum_{j=1}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+2} \right) \\ &= \alpha_0 X \sum_{j=0}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+1} + \alpha_1 X^2 \sum_{j=0}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+2} \end{aligned}$$

This leads to

$$\begin{aligned} \alpha_1 &= \frac{1}{2X^2} \frac{\left( \sum_{j=0}^{\infty} \xi_{p,j} Z^j \right)}{\sum_{j=0}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+2} \left( \sum_{j=0}^{\infty} \xi_{p,j} Z^j \right) - \sum_{j=0}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+1} \left( \sum_{j=0}^{\infty} \eta_{p,j} Z^j \right)} \\ &= \frac{1}{2X^2} \hat{\alpha}_1(\nu, Z) \end{aligned}$$

$$\begin{aligned} \alpha_0 &= -\frac{1}{2X} \frac{\left( \sum_{j=0}^{\infty} \eta_{p,j} Z^j \right)}{\sum_{j=0}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+2} \left( \sum_{j=0}^{\infty} \xi_{p,j} Z^j \right) - \sum_{j=0}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+1} \left( \sum_{j=0}^{\infty} \eta_{p,j} Z^j \right)} \\ &= \frac{1}{2X} \hat{\alpha}_0(\nu, Z) \end{aligned}$$

Now observe that the integral of  $f(x)x^2$  is in fact proportional to  $X^2$  for a fixed  $Z$ :

$$\begin{aligned} \int_0^X f(x)x^2 dx &= \alpha_0 X^3 \sum_{j=0}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+3} + \alpha_1 X^4 \sum_{j=0}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+4} \\ &= \frac{X^2}{2} \left[ \hat{\alpha}_0(\nu, Z) \sum_{j=0}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+3} + \hat{\alpha}_1(\nu, Z) \sum_{j=0}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+4} \right] \end{aligned}$$

To determine  $m(\cdot)$  and  $\mathcal{T}(\cdot)$ , it is useful to consider separately the cases  $\nu \geq 1$  and  $\nu = 0$ . Start with  $\nu \geq 1$ . In this case, in addition to [equation \(85\)](#), we know that

$$3\sigma^2\beta_3 = 1 \text{ and } \sigma^2\gamma_2 = -1$$

The boundary conditions are  $m(X) = \mathcal{T}(X) = 0$ , so

$$\begin{aligned} -\beta_1 &= \frac{X^2}{3\sigma^2} \left( \frac{1 + \sum_{j=1}^{\infty} \eta_{m,j} Z^j}{1 + \sum_{j=1}^{\infty} \xi_{m,j} Z^j} \right) \\ \gamma_0 &= \frac{X^2}{\sigma^2} \left( \frac{1 + \sum_{j=1}^{\infty} \eta_{t,j} Z^j}{1 + \sum_{j=1}^{\infty} \xi_{t,j} Z^j} \right) \end{aligned}$$

The functional forms are then

$$m(x) = -\frac{xX^2}{3\sigma^2} \left( \frac{1 + \sum_{j=1}^{\infty} \eta_{m,j} Z^j}{1 + \sum_{j=1}^{\infty} \xi_{m,j} Z^j} \right) \sum_{j=0}^{\infty} \xi_{m,j} y^{-j} x^{j(\nu+2)} + \frac{x^3}{3\sigma^2} \sum_{j=0}^{\infty} \eta_{m,j} y^{-j} x^{j(\nu+2)}$$

$$\mathcal{T}(x) = \frac{X^2}{\sigma^2} \left( \frac{1 + \sum_{j=1}^{\infty} \eta_{t,j} Z^j}{1 + \sum_{j=1}^{\infty} \xi_{t,j} Z^j} \right) \sum_{j=0}^{\infty} \xi_{t,j} y^{-j} x^{j(\nu+2)} - \frac{x^2}{\sigma^2} \sum_{j=0}^{\infty} \eta_{t,j} y^{-j} x^{j(\nu+2)}$$

Observe that for  $\mathcal{T}(0)$  we have

$$\mathcal{T}(0) = \frac{X^2}{\sigma^2} \left( \frac{1 + \sum_{j=1}^{\infty} \eta_{t,j} Z^j}{1 + \sum_{j=1}^{\infty} \xi_{t,j} Z^j} \right) = \mathcal{T}_0 \left( \frac{1 + \sum_{j=1}^{\infty} \eta_{t,j} (2\kappa \mathcal{T}_0)^j}{1 + \sum_{j=1}^{\infty} \xi_{t,j} (2\kappa \mathcal{T}_0)^j} \right)$$

At  $a = 0$  or, equivalently,  $\kappa = 0$ , we have  $\mathcal{T}(0) = \mathcal{T}_0$ .

Now consider the case  $\nu = 0$ . Here the conditions we add to [equation \(85\)](#) are

$$a\beta_1 = 3\sigma^2\beta_3 - 1 \text{ and } a\gamma_0 = \sigma^2\gamma_2 + 1 \tag{86}$$

Plugging them into the boundary conditions  $m(X) = \mathcal{T}(X) = 0$ ,

$$-\beta_1 = \frac{X^2 \sum_{j=0}^{\infty} \eta_{m,j} Z^j}{3\sigma^2 \sum_{j=0}^{\infty} \xi_{m,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{m,j} Z^j}$$

$$\beta_3 = \frac{\sum_{j=0}^{\infty} \xi_{m,j} Z^j}{3\sigma^2 \sum_{j=0}^{\infty} \xi_{m,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{m,j} Z^j}$$

$$\gamma_0 = \frac{X^2 \sum_{j=0}^{\infty} \eta_{t,j} Z^j}{\sigma^2 \sum_{j=0}^{\infty} \xi_{t,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{t,j} Z^j}$$

$$-\gamma_2 = \frac{\sum_{j=0}^{\infty} \xi_{t,j} Z^j}{\sigma^2 \sum_{j=0}^{\infty} \xi_{t,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{t,j} Z^j}$$

The functional forms in this case are

$$m(x) = -x \frac{X^2 \left( \sum_{j=0}^{\infty} \eta_{m,j} Z^j \right) \left( \sum_{j=0}^{\infty} \xi_{m,j} y^{-j} x^{j(\nu+2)} \right)}{3\sigma^2 \sum_{j=0}^{\infty} \xi_{m,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{m,j} Z^j}$$

$$+ x^3 \frac{\left( \sum_{j=0}^{\infty} \xi_{m,j} Z^j \right) \left( \sum_{j=0}^{\infty} \eta_{m,j} y^{-j} x^{j(\nu+2)} \right)}{3\sigma^2 \sum_{j=0}^{\infty} \xi_{m,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{m,j} Z^j}$$

$$\mathcal{T}(x) = \frac{X^2 \left( \sum_{j=0}^{\infty} \eta_{t,j} Z^j \right) \left( \sum_{j=0}^{\infty} \xi_{t,j} y^{-j} x^{j(\nu+2)} \right)}{\sigma^2 \sum_{j=0}^{\infty} \xi_{t,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{t,j} Z^j}$$

$$- x^2 \frac{\left( \sum_{j=0}^{\infty} \xi_{t,j} Z^j \right) \left( \sum_{j=0}^{\infty} \eta_{t,j} y^{-j} x^{j(\nu+2)} \right)}{\sigma^2 \sum_{j=0}^{\infty} \xi_{t,j} Z^j + aX^2 \sum_{j=0}^{\infty} \eta_{t,j} Z^j}$$

Observe that in this case for  $\mathcal{T}(0)$  we have

$$\begin{aligned}\mathcal{T}(0) &= \frac{X^2}{\sigma^2} \left( \frac{\sum_{j=0}^{\infty} \eta_{t,j} Z^j}{\sum_{j=0}^{\infty} \xi_{t,j} Z^j + \frac{aX^2}{\sigma^2} \sum_{j=0}^{\infty} \eta_{t,j} Z^j} \right) \\ &= \mathcal{T}_0 \left( \frac{1 + \sum_{j=1}^{\infty} \eta_{t,j} (2\kappa \mathcal{T}_0)^j}{1 + \kappa \mathcal{T}_0 + \sum_{j=1}^{\infty} \xi_{t,j} (2\kappa \mathcal{T}_0)^j + \sum_{j=1}^{\infty} \eta_{t,j} (2\kappa \mathcal{T}_0)^j} \right)\end{aligned}$$

When  $\kappa = 0$ , we have  $\mathcal{T}(0) = \mathcal{T}_0$ .

We know that the adjustment frequency is given by

$$N_a = \frac{1}{\mathcal{T}(0)}$$

Hence, the adjustment frequency can be represented as a function of  $\kappa$  and  $\mathcal{T}_0$ . The same is true for the kurtosis of price changes. From [equation \(17\)](#),

$$\begin{aligned}\text{Kurt}(\Delta p) &= \frac{2 \left[ \int_0^X x^4 \Lambda(x) f(x) dx - X^4 \frac{\sigma^2}{2} f'(X) \right]}{N_a} \frac{1}{[\text{Var}(\Delta p)]^2} \\ &= \frac{2N_a \left[ \int_0^X x^4 \Lambda(x) f(x) dx - X^4 \frac{\sigma^2}{2} f'(X) \right]}{\sigma^4} = \frac{12N_a}{\sigma^2} \int_0^X f(x) x^2 dx \\ &= 6N_a \frac{X^2 \sum_{j=0}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+4} \left( \sum_{j=0}^{\infty} \xi_{p,j} Z^j \right) - \sum_{j=0}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+3} \left( \sum_{j=0}^{\infty} \eta_{p,j} Z^j \right)}{\sum_{j=0}^{\infty} \frac{\eta_{p,j} Z^j}{j(\nu+2)+2} \left( \sum_{j=0}^{\infty} \xi_{p,j} Z^j \right) - \sum_{j=0}^{\infty} \frac{\xi_{p,j} Z^j}{j(\nu+2)+1} \left( \sum_{j=0}^{\infty} \eta_{p,j} Z^j \right)} \\ &= 6N_a \mathcal{T}_0 \frac{\sum_{j=0}^{\infty} \varphi_{K,j} (2\kappa \mathcal{T}_0)^j}{\sum_{j=0}^{\infty} \chi_{K,j} (2\kappa \mathcal{T}_0)^j}\end{aligned}$$

Here the coefficients  $\{\varphi_{K,j}, \chi_{K,j}\}_{j=0}^{\infty}$  are given by

$$\begin{aligned}\varphi_{K,j} &= \sum_{i=0}^j \left( \frac{\xi_{p,j-i} \eta_{p,i}}{i(\nu+2)+4} - \frac{\eta_{p,j-i} \xi_{p,i}}{i(\nu+2)+3} \right) \\ \chi_{K,j} &= \sum_{i=0}^j \left( \frac{\xi_{p,j-i} \eta_{p,i}}{i(\nu+2)+2} - \frac{\eta_{p,j-i} \xi_{p,i}}{i(\nu+2)+1} \right)\end{aligned}$$

As expected, when  $\kappa = 0$  we have  $N_a = 1/\mathcal{T}_0$  and

$$\text{Kurt}(\Delta p) = 1.$$

The coefficients  $\{\varphi_{N,j}, \chi_{N,j}\}$  for  $N_a$  are taken from the corresponding formula for  $\mathcal{T}(0)$  in the cases  $\nu = 0$  and  $\nu \geq 1$ . In both cases  $\varphi_{N,0} = \chi_{N,0} = 1$ . To verify  $\varphi_{K,0} = -1/12$  and  $\psi_{K,0} = 1/2$ , plug

$\xi_{p,0} = \eta_{p,0} = 1$ . For  $\varphi_{K,1}$  and  $\chi_{K,1}$ , recall that

$$\xi_{p,1} = \frac{1}{(\nu+2)(\nu+1)} \text{ and } \eta_{p,1} = \frac{1}{(\nu+2)(\nu+3)}$$

The first derivative of  $Kurt(\Delta p)/(6N_a)$  evaluated at  $\kappa = 0$  is

$$\left. \frac{\partial}{\partial \kappa} \left( \frac{Kurt(\Delta p)}{6N_a} \right) \right|_{\kappa=0} = \mathcal{T}_0 \frac{\chi_{K,0}\varphi_{K,1} - \varphi_{K,0}\chi_{K,1}}{\chi_{K,0}^2} = -C(6\varphi_{K,1} - \chi_{K,1})$$

for some positive constant  $C$ . Plugging the terms,

$$\begin{aligned} \varphi_{K,1} &= -\frac{1}{12(\nu+5)(\nu+6)} \\ \chi_{K,1} &= -\frac{1}{2(\nu+3)(\nu+4)} \end{aligned}$$

Hence,

$$\left. \frac{\partial}{\partial \kappa} \left( \frac{Kurt(\Delta p)}{6N_a} \right) \right|_{\kappa=0} = \frac{C}{2} \left( \frac{1}{(\nu+5)(\nu+6)} - \frac{1}{(\nu+3)(\nu+4)} \right) < 0$$

This proves the fact that  $Kurt(\Delta p)/(6N_a)$  decreases for small  $\kappa$ .

## K Special Cases of Interest

### K.1 $m$ and $f$ in the discrete, unbounded case

We assume that we can divide  $[0, \infty)$  into  $N$  segments, each one where  $\Lambda(x)$  is constant at the value  $\rho_k > 0$  and with thresholds  $\{\bar{x}_k\}_{k=0}^N$  as follows. The values of  $\{\bar{x}_k\}$  and  $\{\rho_k\}$  are given. We let

$$0 = \bar{x}_0 < \bar{x}_1 < \bar{x}_2 < \dots < \bar{x}_{N-1} < \bar{x}_N = \infty$$

The function  $\Lambda(x)$  takes  $N$  different strictly positive values denoted by  $\{\rho_k\}_{k=1}^N$ , so that:

$$\begin{aligned} \Lambda(x) &= \rho_k \text{ for } x \in [\bar{x}_{k-1}, \bar{x}_k) \text{ for } k = 1, 2, \dots, N \\ 0 &< \rho_1 < \rho_2 < \dots < \rho_N. \end{aligned}$$

Since  $m(\cdot)$  and  $f(\cdot)$  solve Kolmogorov equations (backward for  $m(\cdot)$  and forward for  $f(\cdot)$ ), on each segment they can be parametrized by a pair of unknown constants:

$$\begin{aligned} m(x) &= M_k(x) = -\frac{x}{\rho_k} + u_k e^{\eta_k x} + v_k e^{-\eta_k x} \text{ for } x \in [\bar{x}_{k-1}, \bar{x}_k] \\ f(x) &= \bar{P}_k(x) = p_k e^{\eta_k x} + q_k e^{-\eta_k x} \text{ for } x \in [\bar{x}_{k-1}, \bar{x}_k] \\ \eta_k &= \sqrt{\frac{2\rho_k}{\sigma^2}} \end{aligned}$$

for  $k = 1, 2, \dots, N$ . We require that  $f(\cdot)$  and  $m(\cdot)$  be continuously differentiable on  $(0, \infty)$ . This implies that

$$M_k(\bar{x}_k) = M_{k+1}(\bar{x}_k) \text{ and } M'_k(\bar{x}_k) = M'_{k+1}(\bar{x}_k) \text{ for all } k = 1, 2, \dots, N-1 \quad (87)$$

$$\bar{P}_k(\bar{x}_k) = \bar{P}_{k+1}(\bar{x}_k) \text{ and } \bar{P}'_k(\bar{x}_k) = \bar{P}'_{k+1}(\bar{x}_k) \text{ for all } k = 1, 2, \dots, N-1 \quad (88)$$

In addition we have the following conditions. Since  $m$  is antisymmetric around zero we require  $m(0) = 0$ . Since  $f$  is a density, it must integrate to one, and since it symmetric it must integrate to one half over positive  $x$ . Finally, both  $m$  and  $f$  should converge to  $-x/\rho_N$  and 0 as  $x \rightarrow \infty$ . These conditions are sometimes referred as no-bubble conditions. Hence:

$$M_1(0) = 0, \quad \frac{1}{2} = \int_0^\infty f(x)dx = \sum_{k=1}^N \int_{\bar{x}_{k-1}}^{\bar{x}_k} \bar{P}_k(x)dx, \text{ and } p_N = u_N = 0$$

Overall, we have  $2N$  unknowns, namely  $\{u_k, v_k\}_{k=1}^N$ , and  $2N$  linear equations for  $m(\cdot)$ , namely  $2(N-1)$  from [equation \(87\)](#), that  $m(0) = 0$ , and the no-bubble condition. Likewise for  $f(\cdot)$ . We can write these equations and solve for the constants. Once we have them we can evaluate:

$$\int_0^\infty x^2 f(x)dx = \sum_{k=1}^N \int_{\bar{x}_{k-1}}^{\bar{x}_k} x^2 \bar{P}_k(x)dx \text{ and}$$

$$\int_0^\infty m(x)f(x)dx = \sum_{k=1}^N \int_{\bar{x}_{k-1}}^{\bar{x}_k} M'_k(x)\bar{P}_k(x)dx.$$

and check if:

$$\sum_{k=1}^N \int_{\bar{x}_{k-1}}^{\bar{x}_k} x^2 \bar{P}_k(x)dx = -\sigma^2 \sum_{k=1}^N \int_{\bar{x}_{k-1}}^{\bar{x}_k} M'_k(x)\bar{P}_k(x)dx.$$

Now we will determine the coefficients  $\{p_k, q_k\}_{k=1}^N$  and  $\{u_k, v_k\}_{k=1}^N$ . Start with the ones for  $\bar{p}(\cdot)$ . Combining the continuity and differentiability conditions, we can write the coefficients recursively for  $k = 1, 2, \dots, N-1$ :

$$p_k = \frac{1}{2} \left( 1 + \frac{\eta_{k+1}}{\eta_k} \right) e^{(\eta_{k+1}-\eta_k)x_k} p_{k+1} + \frac{1}{2} \left( 1 - \frac{\eta_{k+1}}{\eta_k} \right) e^{-(\eta_{k+1}+\eta_k)x_k} q_{k+1}$$

$$q_k = \frac{1}{2} \left( 1 + \frac{\eta_{k+1}}{\eta_k} \right) e^{(\eta_k-\eta_{k+1})x_k} q_{k+1} + \frac{1}{2} \left( 1 - \frac{\eta_{k+1}}{\eta_k} \right) e^{(\eta_{k+1}+\eta_k)x_k} p_{k+1}$$

We also have the terminal condition  $p_N = 0$  and the normalization (the density must integrate to one half over positives). Observe that the coefficients are in fact linear in  $q_N$ , so  $q_N$  can easily be found from the normalization. The integral is

$$\frac{1}{2} = \int_0^\infty f(x)dx = \sum_{k=0}^{N-1} p_{k+1} \frac{e^{\eta_{k+1}x_{k+1}} - e^{\eta_{k+1}x_k}}{\eta_{k+1}} - \sum_{k=0}^{N-1} q_{k+1} \frac{e^{-\eta_{k+1}x_{k+1}} - e^{-\eta_{k+1}x_k}}{\eta_{k+1}}$$



We can use linearity: letting  $p_k = \hat{p}_k q_N$  and  $q_k = \hat{q}_k q_N$  and plugging this into the normalization, we can write

$$\frac{1}{2} = \sum_{k=0}^{N-1} \left( \hat{p}_{k+1} \frac{e^{\eta_{k+1} x_{k+1}} - e^{\eta_{k+1} x_k}}{\eta_{k+1}} - \hat{q}_{k+1} \frac{e^{-\eta_{k+1} x_{k+1}} - e^{-\eta_{k+1} x_k}}{\eta_{k+1}} \right) q_N \quad (89)$$

The numbers  $\{\hat{p}_k, \hat{q}_k\}_{k=1}^{N-1}$  are easily obtained from  $\{p_k, q_k\}_{k=1}^{N-1}$  computed recursively for some presupposed value of  $q_N$ . Knowing them, we can recover the real  $q_N$  from [equation \(89\)](#) and recompute the real  $\{p_k, q_k\}_{k=1}^{N-1}$ .

Now we will determine the coefficients for  $m(\cdot)$ . The continuity and differentiability conditions lead to the following recursive representation:

$$\begin{aligned} u_k &= \frac{1}{2} \left( 1 + \frac{\eta_{k+1}}{\eta_k} \right) e^{(\eta_{k+1} - \eta_k) x_k} u_{k+1} + \frac{1}{2} \left( 1 - \frac{\eta_{k+1}}{\eta_k} \right) e^{-(\eta_{k+1} + \eta_k) x_k} v_{k+1} \\ &\quad + \frac{1}{2} \left( x + \frac{1}{\eta_k} \right) \left( \frac{1}{\rho_k} - \frac{1}{\rho_{k+1}} \right) e^{-\eta_k x_k} \\ v_k &= \frac{1}{2} \left( 1 + \frac{\eta_{k+1}}{\eta_k} \right) e^{(\eta_k - \eta_{k+1}) x_k} v_{k+1} + \frac{1}{2} \left( 1 - \frac{\eta_{k+1}}{\eta_k} \right) e^{(\eta_{k+1} + \eta_k) x_k} u_{k+1} \\ &\quad + \frac{1}{2} \left( x - \frac{1}{\eta_k} \right) \left( \frac{1}{\rho_k} - \frac{1}{\rho_{k+1}} \right) e^{-\eta_k x_k} \end{aligned}$$

We also have the terminal condition  $u_N = 0$  and the antisymmetry condition  $m(0) = 0$ . The latter one reduces to  $u_1 + v_1 = 0$ . Now we can observe that all  $u_k$  and  $v_k$  are in fact affine in  $v_N$ :  $u_k = \hat{u}_k v_N + \tilde{u}_k$  and  $v_k = \hat{v}_k v_N + \tilde{v}_k$ . The condition  $m(0) = 0$  can be written as

$$0 = u_1 + v_1 = (\hat{u}_1 + \hat{v}_1) v_N + (\tilde{u}_1 + \tilde{v}_1) \quad (90)$$

The coefficients  $\{\hat{u}_k, \hat{v}_k\}_{k=1}^{N-1}$  and  $\{\tilde{u}_k, \tilde{v}_k\}_{k=1}^{N-1}$  can be found from  $\{u_k, v_k\}_{k=1}^{N-1}$  computed recursively for two different presupposed values of  $v_N$  (we need two because the functions are affine, not linear). After that, we can recover the real  $v_N$  from [equation \(90\)](#) and recompute the real  $\{u_k, v_k\}_{k=1}^{N-1}$ .

## L Discrete Distribution of Fixed Costs

Let  $g_i > 0$  be the probability of drawing a fixed cost  $\psi_i$  for  $i = 1, \dots, n-1$ , conditional of drawing a low adjustment cost opportunity. We have  $0 < \psi_1 < \dots < \psi_{n-1}$ . A firm can always pay a fixed cost  $\Psi \equiv \psi_n$  and change prices, with  $\psi_n > \psi_{n-1}$ . At all points  $x$  where  $v$  is twice differentiable we have:

$$rv(x) = \min \left\{ Bx^2 + \frac{\sigma^2}{2} v''(x) + \kappa \sum_{j=1}^{n-1} \min \{ \psi_j + v(0) - v(x), 0 \} g_j, r(\psi_n + v(0)) \right\}$$

The optimal decision rule can be described by  $n$  thresholds  $0 < \bar{x}_1 < \bar{x}_2 < \dots < \bar{x}_n \equiv X$ . The optimal decision rule is that conditional on drawing the adjustment cost  $\psi_j$  an adjustment takes

place if  $|x| \geq \bar{x}_j$  for  $j = 1, \dots, n$ . Note that this implies that:

$$v(\bar{x}_j) + \psi_j = v(0) \text{ for } j = 1, 2, \dots, n.$$

To simplify the notation we let:

$$\lambda_j \equiv \kappa g_j \text{ for } j = 1, \dots, n-1 \text{ and } \Lambda(x) = \sum_{k=1}^{n-1} \lambda_k 1_{\{x \geq \bar{x}_k\}}$$

To summarize the firm's problem is defined by parameters  $r, B, \sigma^2, \{\lambda_j\}_{j=1}^{n-1}, \{\psi_j\}_{j=1}^n$ . The solution is given by a set of thresholds  $\{\bar{x}_j\}_{j=1}^n$  with  $0 < \bar{x}_1 < \dots < \bar{x}_n$ .

We can write the value function for each segment  $j = 1, 2, \dots, n$ :

$$\left( r + \sum_{k=1}^{j-1} \lambda_k \right) v_j(x) = Bx^2 + \frac{\sigma^2}{2} v_j''(x) + \sum_{k=1}^{j-1} [v_1(0) + \psi_k] \lambda_k \text{ for } x \in (\bar{x}_{j-1}, \bar{x}_j]$$

where for convenience we define  $\bar{x}_0 = 0$ . The value function  $v$  must be differentiable at all  $x \in \mathbb{R}$ , and twice differentiable for all  $x \in \mathbb{R}$ , except  $x = \bar{x}_j$  for  $j = 1, \dots, n$ . Thus we have the boundary conditions:

$$v'(0) = v'(\bar{x}_n) = 0$$

## L.1 Value function for discrete $\psi$ distribution

The solution of the value function  $v$  is characterized by coefficients  $\{a_j, b_j, c_j\}_{j=1}^n$ , roots  $\{\eta_j\}_{j=1}^n$  and thresholds  $\{\bar{x}_j\}_{j=1}^n$ . In particular, given the thresholds  $\{\bar{x}_j\}_{j=1}^n$  we write a linear o.d.e. for each segment  $[\bar{x}_{j-1}, \bar{x}_j]$  for  $j = 1, \dots, n$ . This o.d.e. is parametrized by three constants  $a_j, b_j, c_j$  as follows:

$$v_j(x) = a_j + b_j x^2 + c_j (e^{\eta_j x} + e^{-\eta_j x}) \text{ for } x \in [\bar{x}_{j-1}, \bar{x}_j] \text{ and } j = 1, \dots, n$$

where  $\eta_j$  is given by:

$$\eta_j = \sqrt{\frac{(r + \sum_{k=1}^{j-1} \lambda_k)}{\sigma^2/2}}$$

Replacing the non-homogenous solution  $a_j + b_j x^2$  into the o.d.e. in each segment we have:

$$\left( r + \sum_{k=1}^{j-1} \lambda_k \right) (a_j + b_j x^2) = Bx^2 + \frac{\sigma^2}{2} 2b_j + \sum_{k=1}^{j-1} [v_1(0) + \psi_k] \lambda_k \text{ for } x \in [\bar{x}_{j-1}, \bar{x}_j] \text{ and } j = 1, \dots, n$$

Matching the terms quadratic in  $x$ , and using that  $v_1(0) = a_1 + 2c_1$ , we get:

$$\left( r + \sum_{k=1}^{j-1} \lambda_k \right) b_j = B \text{ for } j = 1, \dots, n \quad (91)$$

Matching the constant we have:

$$\left( r + \sum_{k=1}^{j-1} \lambda_k \right) a_j = \sigma^2 b_j + \sum_{k=1}^{j-1} [a_1 + 2c_1 + \psi_k] \lambda_k \text{ for } j = 1, \dots, n \quad (92)$$

The continuity and (once) differentiability at  $x = \bar{x}_j$  for  $j = 1, \dots, n-1$  gives:

$$a_{j+1} + b_{j+1}(\bar{x}_j)^2 + c_{j+1}(e^{\eta_{j+1}\bar{x}_j} + e^{-\eta_{j+1}\bar{x}_j}) = a_j + b_j(\bar{x}_j)^2 + c_j(e^{\eta_j\bar{x}_j} + e^{-\eta_j\bar{x}_j}) \quad \text{for } j = 1, \dots, n-1 \quad (93)$$

and

$$2b_{j+1}\bar{x}_j + c_{j+1}\eta_{j+1}(e^{\eta_{j+1}\bar{x}_j} - e^{-\eta_{j+1}\bar{x}_j}) = 2b_j\bar{x}_j + c_j\eta_j(e^{\eta_j\bar{x}_j} - e^{-\eta_j\bar{x}_j}) \quad \text{for } j = 1, \dots, n-1 \quad (94)$$

value matching and smooth pasting at  $\bar{x}_n$  gives:

$$\psi_n + a_1 + 2c_1 = a_n + b_n(\bar{x}_n)^2 + c_n(e^{\eta_n\bar{x}_n} + e^{-\eta_n\bar{x}_n}) \quad (95)$$

$$0 = 2b_n\bar{x}_n + c_n\eta_n(e^{\eta_n\bar{x}_n} - e^{-\eta_n\bar{x}_n}) \quad (96)$$

The optimal return point conditions,  $v'(0) = 0$ , is automatically satisfied by symmetry of the value functions.

Thus we have  $4 \times n$  unknowns, namely  $\{\bar{x}_j, a_j, b_j, c_j\}_{j=1}^n$ , and  $4 \times n$  equations, namely  $n$  equations matching quadratic terms, i.e. [equations \(91\)](#),  $n$  equations matching constants, i.e. [equations \(92\)](#),  $n-1$  equations enforcing continuity, i.e. [equations \(93\)](#),  $n-1$  equations enforcing differentiability, i.e. [equations \(94\)](#), and two more equations on the boundary  $\bar{x}_n$  enforcing value matching, i.e. [equation \(95\)](#), and smooth pasting, i.e. [equation \(96\)](#).

## L.2 Inverse problem: recovering the cost function

We now solve an inverse problem, namely how to recover the menu cost values  $\psi_j$  that underlie a given observed hazard function  $\Lambda(x)$  at given thresholds  $\{\bar{x}_j\}$ . The main result is summarized by the next proposition:

**PROPOSITION 16.** Fix a discount rate, curvature and variance  $r, B, \sigma^2 > 0$ , and a step function  $\Lambda$  giving the probability per unit of time of a price adjustment for  $|x| < x_n$ . The function  $\Lambda$  is described by a set of probability rates for costs  $\{\lambda_j\}_{j=1}^{n-1} \in \mathbb{R}_+^n$  for  $n \geq 1$ , and a set of  $n$  thresholds  $\{\bar{x}_j\}_{j=1}^n$  with  $0 = \bar{x}_0 < \bar{x}_1 < \dots < \bar{x}_n$ . Then there is a unique set of  $n$  fixed costs  $0 = \psi_0 < \psi_1 < \dots < \psi_n$  so that the  $n$  thresholds  $\{\bar{x}_j\}_{j=1}^n$  solve the firm's problem defined by  $r, B, \sigma^2, \{\lambda_j\}_{j=0}^{n-1}, \{\psi_j\}_{j=1}^n$ . Moreover, the fixed costs  $\{\psi_j\}_{j=1}^n$  and the coefficients of the value function  $\{a_j, b_j, c_j\}_{j=1}^n$  solve a system of linear equations.

**Proof.** (of [Proposition 16](#)) We first solve for each of the coefficients  $b_j$  using [equation \(91\)](#) for each  $j = 1, \dots, n$ .

We note that the thresholds  $\{\bar{x}_j\}_{j=1}^n$  are given and that roots  $\{\eta_j\}_{j=1}^n$  can be computed as functions of given parameters.

Using the coefficients  $\{b_j\}_{j=1}^n$ , we solve for the coefficients  $\{c_j\}_{j=1}^n$ . First we solve for  $c_n$  enforcing smooth pasting at  $\bar{x}_n$  given by [equation \(96\)](#). Using  $c_n$  we recursively use  $c_{j+1}$  to solve for  $c_j$  imposing differentiability between adjacent segments, i.e. [equations \(94\)](#) for  $j = n-1, n-2, \dots, 1$ .

Next we solve for the  $\{a_j\}_{j=1}^n$ , given  $\{b_j, c_j\}_{j=1}^n$ . First, use  $rv(0) = \frac{\sigma^2}{2}v''(0) = \frac{\sigma^2}{2}(2b_1 + (\eta_1)^2 2c_1)$  and  $v(0) = a_1 + 2c_1$  to solve for  $a_1$ , namely  $a_1 = \frac{\sigma^2}{r}(b_1 + \eta_1^2 c_1) - 2c_1$ . Next, use [equations \(93\)](#) to solve recursively for  $\{a_j\}_{j=2}^n$ .

Finally, we solve for the fixed costs  $\{\psi_j\}_{j=1}^n$  using value matching and the values of  $\{a_j, b_j, c_j\}_{j=1}^n$ . They give:

$$\psi_j = v(\bar{x}_j) - v(0) = a_j + b_j(\bar{x}_j)^2 + c_j(e^{\eta_j\bar{x}_j} + e^{-\eta_j\bar{x}_j}) - a_1 - 2c_1$$

for  $j = 1, \dots, n$ .  $\square$