Aggregate Dynamics in Lumpy Economies*

Isaac Baley†
UPF, CREi, Barcelona GSE

Andrés Blanco ‡
University of Michigan

March 20, 2019

Abstract

We develop tools to analyze the aggregate implications of lumpiness in microeconomic adjustment. We derive a set of structural relationships between the steady-state moments and the business cycle dynamics of cross-sectional distributions, and we show how to discipline these relationships using micro panel data. As an application, we implement our machinery in a standard framework of firm investment with non-convex adjustment costs and random opportunities of free adjustment. We demonstrate analytically that, in order to explain aggregate capital dynamics, that model must match two steady-state moments related to capital misallocation and the time since the last investment. Using plant–level data from Chile and Colombia, we compute these two moments, and discover that there does not exist a calibration of the lumpy investment model that is consistent with the data.

JEL: D30, D80, E20, E30

Keywords: inaction, lumpiness, transitional dynamics, fixed adjustment costs, aggregate shocks, sufficient statistics, firm investment, Ss models.

*We thank Adrien Auclert, Rudi Bachmann, Fernando Broner, Andrea Caggese, Javier Cravino, Andrés Drenik (discussant), John Leahy, Kurt Mitman, Matthias Meier, Pablo Ottonello, Jaume Ventura, Edouard Schaal, and seminar participants at CREi, UPF, Michigan, Banque de France, Bocconi, Cleveland FED, PSE, EM^{3}C Conference, Transpyrenean Macro Workshop 2019, and XXII Workshop in International Economics and Finance for extremely useful feedback and suggestions. Lauri Esala provided outstanding research assistance. Isaac Baley gratefully acknowledges the support of the Spanish Ministry of the Economy and Competitiveness, through the Seed Grant (Aggregate Dynamics in Lumpy Economies) of the Barcelona Graduate School of Economics Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563).

†Universitat Pompeu Fabra, CREi, Barcelona GSE, isaac.baley@upf.edu. Ramón Trias Fargas 25–27, Barcelona, 08015.

‡University of Michigan, jablanco@umich.edu.

1 Introduction

Lumpiness in microeconomic adjustment is pervasive in many economic environments. Capital investment, labor hiring and firing, inventories, consumption of durable goods, price setting, portfolio management, and many other economic decisions faced by firms and households are characterized by periods of inaction followed by bursts of activity. How does lumpiness in microeconomic adjustment affect aggregate dynamics? After a policy change or an aggregate shock, how long do transitions last until the lumpy economy reaches its new long-run equilibrium? Understanding these issues is key for the design and implementation of policies aiming at stabilizing the business cycle or generating long-run growth.

This paper presents new tools to study aggregate dynamics in lumpy economies. We consider environments with ex-ante identical agents that make decisions subject to idiosyncratic shocks and adjustment frictions. These frictions may take the form of non-convex adjustment costs, random opportunities of adjustment, fixed dates of adjustments, among many others. A common challenge that arises when studying this type of environments is that aggregate dynamics depend on the cross-sectional distribution of agents, a highly dimensional and complicated object. The standard method to tackle this challenge is through state-of-the-art numerical techniques. We propose a complementary approach by deriving a set of structural relationships between the steady-state cross-sectional moments and microdata on adjustments, and between the steady-state moments and business cycle dynamics.

Our first result establishes a structural relationship between steady-state moments and microdata on adjustments. The idea behind this mapping consists in assuming a process for idiosyncratic shocks that hit agents during periods of inaction and a minimal structure for the adjustment policy, and then use the information revealed through agents' actions—frequency and size of adjustments—to back out the cross-sectional moments. Let us stress that the relationship between ergodic moments and data is model-free, as we do not assume any particular structure for the adjustments frictions; the only requirement is that policies are "memoryless", in the sense that upon adjustment, the agent fully incorporates all the accumulated idiosyncratic shocks into her policy.[1] An attractive feature of this approach is that it allows to quantify cross-sectional moments of unobserved variables, for example, markup dispersion or skewness of marginal products, using information on observable actions available in panel data, such as price changes or investment rates, with a minimum set of assumptions.

Our second result establishes a structural relationship between the steady-state moments and transitional dynamics of the distribution. The logic is that steady-state moments are informative about agents' responsiveness to idiosyncratic shocks, and therefore, are good predictors of the effect of lumpiness on the speed of adjustment. For example, an economy that features substantial ex-post heterogeneity relative to the volatility of idiosyncratic shocks must necessarily have important adjustment frictions, which in turn generate unresponsive policies and persistent transitional dynamics.

In this context, business cycle dynamics are analyzed in the following way. Starting from any arbitrary initial distribution, we characterize the convergence of any moment of such distribution towards its ergodic counterpart. Following Alvarez, Le Bihan and Lippi (2016), we measure transitional dynamics through

[1] Formally, we require certain degree of history independence in the stochastic processes and policies in order to collapse all the ex-post heterogeneity due to idiosyncratic shocks and frictions to the problem of a representative agent. Importantly, this aggregation result *does not imply* that heterogeneity is irrelevant for aggregate dynamics; it says that all heterogeneity can be summarized in a compact way.

the cumulative impulse-response function (CIR), which equals the cumulative deviations in the moment of interest with respect to its steady-state value, i.e. the area under the IRF. We show analytically that the CIR can be represented through a linear combination of cross-sectional moments. This mapping is not model-free anymore; still, we characterize the CIR for widely used models of inaction in a variety of economic applications.

With the two theoretical results at hand, we study aggregate investment dynamics in an economy with idiosyncratic productivity shocks, a common drift due to depreciation and productivity growth, potentially asymmetric policies, and non-convex capital adjustments costs. For this purpose, we set-up a canonical model of lumpy investment à la Khan and Thomas (2008), Bachmann, Caballero and Engel (2013) and Winberry (2016). This literature aims at understanding the role of lumpy investment at the plant-level for aggregate dynamics and has been marked by an active debate about the size of the adjustment frictions. Our approach considers a flexible formulation and lets the data inform us directly about the nature of frictions.

First, we compute various steady-state moments of interest (e.g. dispersion of marginal products of capital) using investment rates and frequency of adjustment from annual plant–level data from Chile and Colombia. Through the lens of our formulas, the data provides clear evidence against fully time-dependent and fully state-dependent models and in favor of hybrid models that combine both types of components.

Second, we study the transitional dynamics via the CIR. For this purpose, we assume a hybrid model for adjustment costs known as the "CalvoPlus" model in which firms pay a fixed cost to invest and get free random opportunities of adjustment.[2] We demonstrate analytically that, in order to explain aggregate capital dynamics, such a model must match two steady-state moments related to capital misallocation and capital age (the time since its last adjustment). Concretely, defining the capital gap as the log of capital to productivity ratio, the CIR of aggregate capital in that model is given by the steady-state dispersion of capital gaps and the covariance between capital gaps and capital age. We compute these two moments for the first time and discover that there does not exist a calibration of the CalvoPlus model that is consistent with the data. In this spirit, our tools can aid researchers in improving their models to be consistent with the empirical evidence on inaction.

Advantages and limitations. As stressed above, one key advantage of our framework is the limited assumptions on the structure of adjustment costs and policies. This permits us to remain agnostic about the true nature and size of adjustment costs and allows the microdata to inform us about them. In the same vein, we can accommodate very general (continuous) stochastic processes. While the mapping between data and ergodic moments clearly depends on the assumptions on the stochastic process (e.g. with drift, without drift, mean-reversion), the theory imposes cross-equation restrictions that allow us to validate such assumptions and discern across processes. Another advantage is that we can recover ergodic moments of variables that are hard to measure in the data, such as markups or capital gaps, with observable data on frequency and size of adjustments.

Regarding the analysis of transitional dynamics, we provide formulas to track any moment of the cross-sectional distribution as well as functions of these moments for a variety of family of models.

[2]This model is proposed by Nakamura and Steinsson (2010) in the context of price-setting.

This is useful as many applications require tracking the dynamics of higher moments of cross-sectional distributions, such as its skewness (Bloom, Guvenen and Salgado, 2016) or tails (Kozlowski, Veldkamp and Venkateswaran, 2015). Moreover, we can accommodate transitions starting from various initial conditions consisting on (small) perturbations around steady-state, such as horizontal shifts of the distribution or mean-preserving spreads, that can be interpreted as arising from aggregate shocks to first moments, second moments (Bloom, 2009), etc.

One limitation of our framework is that we can only recover the CIR of moments of the distribution, but not the complete IRF. The reason is that in order to exploit the ergodic properties of the environment (i.e. exchange the time series of the cross-sectional distribution with a cross-sectional distribution of individual stopping-time problems) we need to consider the complete transitional dynamics. Nevertheless, the CIR is a useful metric to summarize the persistence of aggregate dynamics, it eases the comparison across models, and it can be interpreted as a "multiplier" of aggregate shocks.[3]

Another limitation is that our analysis takes as a premise that the steady-state policies hold along the transition path. This assumption is valid as long as the general equilibrium feedback from the distribution to individual policies though prices is quantitatively insignificant. There are several general equilibrium frameworks in which this is the case.[4] But when general equilibrium effects are quantitative relevant, our framework does not fully characterize aggregate dynamics. Nevertheless, the tools developed in this paper are still informative about the role of lumpiness in richer general equilibrium models and serve as a guide to study one important dimension of the economic environment.[5]

Related literature. Aggregate dynamics in inaction models has been widely studied. The groundbreaking work of Caplin and Spulber (1987), Caballero and Engel (1991) and Caplin and Leahy (1991) provided theoretical guidelines in stylized models to understand the role of micro lumpiness in shaping aggregate dynamics. With the surge of microdata, more realistic models that incorporated idiosyncratic shocks were developed, such as Cooper and Haltiwanger (2006), Golosov and Lucas (2007), Midrigan (2011), Berger and Vavra (2015), Carvalho and Schwartzman (2015) and Álvarez, Lippi and Paciello (2016), with the objective of understanding how the interaction of heterogeneity and lumpiness mattered for aggregate dynamics. We contribute by providing novel theoretical insights and an empirical strategy that exploits the microdata to its maximum while imposing a minimum structure to the inaction model.

Our work is inspired by Alvarez, Le Bihan and Lippi (2016), who consider a multi-product menu cost model with random opportunities to freely adjust and Brownian innovations to markup gaps. In that setup, they study the real effects of monetary shocks. One of their striking results is that the CIR

[3] Álvarez and Lippi (2014); Alvarez, Le Bihan and Lippi (2016); Baley and Blanco (2019) use the CIR to evaluate the effect of monetary policy shocks on output in menu cost models.

[4] For the effect of monetary shocks, see Woodford (2009), Golosov and Lucas (2007), and the vast literature that builds on them. For real exchange dynamics, see Carvalho and Nechio (2011) and Kehoe and Midrigan (2008). Regarding investment models, Bachmann, Caballero and Engel (2013) and Winberry (2016), building on Khan and Thomas (2008), show that partial equilibrium dynamics are not undone by general equilibrium effects whenever the model is calibrated to match the cyclical properties of aggregate investment or interest rates. Web Appendix C describes some of these frameworks.

[5] A concrete example of this logic is found in the context of pricing literature with Calvo-type adjustments. In a model with negligible first order general equilibrium effects, Alvarez, Le Bihan and Lippi (2016) show analytically that the effectiveness of monetary policy is a function of the average duration of pricing spells, independent of any type of heterogeneity. Following this result, Blanco and Cravino (2018) reach a similar conclusion in a model with large general equilibrium effects (arising from real rigidities) in the context of real-exchange dynamics. Therefore, the role of heterogeneity and inaction in shaping aggregate dynamics is not altered by general equilibrium forces.

for average markup gaps—a measure of the real effects of a money shock—equals the kurtosis of price changes times the average duration of prices divided by 6. They show this result analytically for the case of one product ($n = 1$) and infinite products ($n = \infty$), and more generally, they construct power series of each of the terms in the equality and confirm that the relationship holds. Our contribution lies in proving an intermediate link between the CIR and observable actions, given by the steady-state moments. For example, in their menu cost model, the intermediate link between the CIR and the kurtosis of price changes is given by the variance of markups. This strategy allows us to study richer environments and provide new economic insights on the aggregate implications of lumpiness.

Our paper also relates to the pioneer work in Hamermesh (1989), where it is shown how firms' labor decision can be rationalized in a fixed adjustment cost model in which the adjustment trigger depends on the labor gap, i.e. difference between current and static optimal labor. This strategy has been applied in Caballero and Engel (1993) and Caballero, Engel and Haltiwanger (1997) for analyzing the consequences of lumpiness for macroeconomic fluctuations. More recently, a similar approach has been applied in the analysis of cross-country productivity differences due to capital misallocation across firms (see Restuccia and Rogerson (2013) for a survey). While in theory this methodology appears adequate, the challenge lies in finding the optimal target to construct the gap. The standard approach, as in Hsieh and Klenow (2009), consists in specifying a particular production function at the micro level that allows to recover the optimal input demand, therefore the gap. We propose an alternative way that consists in directly assuming a stochastic process for the unobserved marginal product of an input, thus its optimal static demand, and then adding discipline to the parameters of the stochastic process using observable microdata on investment that holds for *all* lumpy adjustment models.

Structure of the paper. Section 2 presents a standard model of lumpy investment that allows us to introduce the objects of interest. Section 3 develops the theory. Section 4 applies the theory using micro-level data. Section 5 generalizes and extends the results.

2 A Model of Lumpy Investment

This section describes the economic environment on which we build and apply the theory. We build a lumpy investment model in the spirit of Khan and Thomas (2008), Bachmann, Caballero and Engel (2013) and Winberry (2016), with a few simplifications that are discussed below.

2.1 Environment

Time is continuous and infinite. A representative household and a continuum of ex-ante identical firms live in the economy. There is no aggregate uncertainty and firms face idiosyncratic shocks to productivity and capital adjustment costs. We denote with $\omega \in \Omega$ the full history of shocks and consider $(\Omega, P, \mathcal{F})$ to be a probability space equipped with the filtration $\mathcal{F} = \{\mathcal{F}_t : t \geq 0\}$. We use the notation $g_{\omega,t} : \Omega \times \mathbb{R} \to \mathbb{R}$ to denote an adapted process (a function $\mathcal{F}_t$-measurable for any $t \geq 0$) and $\mathbb{E}[g_{\omega,t}]$ to denote its expectation under P.

Household. The household chooses the stochastic process for consumption to maximize its expected utility subject to a budget constraint. The household problem is given by

$$\int_0^\infty e^{-\rho t} C_t dt, \quad \text{subject to} \quad \int_0^\infty Q_t \left(C_t - \Pi_t\right) dt \;=\; 0, \tag{1}$$

where $\Pi_t \;\equiv\; \mathbb{E}[\pi_{\omega,t}]$ denotes aggregate firm's profits and C_t denotes household's consumption.

Firms. Firms operate in competitive markets. They produce output y using capital k as the only input through a decreasing returns technology:

$$y_{\omega,t} \;=\; e_{\omega,t}^{1-\alpha} k_{\omega,t}^{\alpha}, \quad \alpha < 1, \tag{2}$$

where the log of idiosyncratic productivity e evolves according to a Brownian motion with drift μ and volatility σ:

$$d\log(e_{\omega,t}) \;=\; \mu dt + \sigma dW_{\omega,t}, \quad W_{\omega,t} \sim Wiener. \tag{3}$$

A firm chooses capital to maximize its expected stream of profits, discounted at the Arrow–Debreu time-zero price Q_t. Capital between adjustments depreciates at a constant rate ψ. For every capital adjustment, a firm pays an adjustment cost proportional to its productivity $\kappa_{\omega,t} e_{\omega,t}$, where $\kappa_{\omega,t}$ is follows a stochastic process to be described below. With these elements, a firm's problem consists in choosing a sequence of adjustment dates $(\tau_{\omega,i})$ and investment rates $(\Delta k_{\tau_{\omega,i}} \;=\; k_{\tau_{\omega,i}} - k_{\tau^-_{\omega,i}})$ that jointly solve the following stopping-time problem:

$$\max_{\{\tau_{\omega,i}, \Delta k_{\tau_{\omega,i}}\}_{i\,=\,1}^{\infty}} \mathbb{E}\left[\int_0^\infty Q_s y_{\omega,s}\, ds - \sum_{i\,=\,1}^{\infty} Q_{\tau_{\omega,i}} \left(\kappa_{\omega,\tau_{\omega,i}} e_{\omega,\tau_{\omega,i}} + \Delta k_{\tau_{\omega,i}}\right)\right], \tag{4}$$

where output and productivity follow (2) and (3), respectively, and capital follows:

$$\log(k_{\omega,s}) \;=\; \log(k_{\omega,0}) - \psi s + \sum_{i:\tau_{\omega,i} \leq s} \Delta k_{\tau_{\omega,i}}. \tag{5}$$

Aggregate feasibility. Aggregate output Y_t is used for household's consumption C_t and firms' investments I_t, which includes capital adjustments adjustment costs:

$$\underbrace{\mathbb{E}\left[e_{\omega,t}^{1-\alpha} k_{\omega,t}^{\alpha}\right]}_{Y_t} \;=\; C_t + \underbrace{\mathbb{E}\left[\mathbb{1}_{\{\tau,t\}}\left[\kappa_{\omega,t} e_{\omega,t} + \Delta k_{\omega,t}\right]\right]}_{I_t}. \tag{6}$$

where $\mathbb{1}_{\{\tau,t\}} \;=\; \{\omega : \exists i \; s.t. \; \tau_{\omega,i} \;=\; t\}$ indicates the set of adjusters.

Equilibrium. Given an initial distribution of $\{k_{\omega,0}, e_{\omega,0}\}$, an equilibrium is a set of stochastic processes for prices $\{Q_t\}$, household's policy $\{C_t\}$, and firms' policies $\{\tau_{\omega,i}, \Delta k_{\omega,i}\}$ such that:

(i) Given prices $\{Q_t\}$, $\{C_t\}$ solves the household's problem (1).

(ii) Given $\{Q_t\}$ and the stochastic processes for productivity and capital adjustment costs, $\{\tau_{\omega,i}, \Delta k_{\omega,i}\}$ solve the firm's investment problem (4).

(iii) Goods market clears in (6).

Discussion of simplifying assumptions. Let us compare our environment with one of the benchmark lumpy investment models by Khan and Thomas (2008). First, in contrast to that paper, we do not consider labor as a production factor. Given that we consider a partial equilibrium setting, and the labor decision is static in their model, this assumption is innocuous as adding labor would only affect the value of the output-capital elasticity. Second, all the investments in our model regardless of their size are subject to the adjustment costs (in the language of that paper, we do not consider *unconstrained* investments). This assumption is quantitatively irrelevant for transitional dynamics, as in the calibration, most investments are constrained anyways due to the large size of idiosyncratic shocks relative to aggregate shocks. Lastly, we consider a random-walk process for idiosyncratic productivity instead of a mean-reverting process. This assumption is done to simplify the exposition at this stage but it is relaxed in Section 5. Moreover, this assumption is motivated by the empirical observation that plant-level investment rates are *iid*.[6]

2.2 Adjustment cost structures

The following specification for the stochastic process for capital adjustment costs $\kappa_{\omega,t}$ spans a wide set of models of adjustment costs. Consider two sequences of *iid* random variables $(u_{\omega,i}, \xi_{\omega,i})$, where $u_{\omega,i}$ is drawn from the distribution H_u over $\mathbb{R}^+$ and $\xi_{\omega,i}$ is drawn from the distribution H_ξ with support over $[0, \kappa]$, with $\kappa > 0$. Then the capital adjustment costs follows:

$$\kappa_{\omega,t} = \begin{cases} \xi_{\omega,i} & \text{if } \sum_{j=1}^{i} u_{\omega,i} = t \text{ for some } i, \\ \kappa & \text{otherwise.} \end{cases} \tag{7}$$

The first class of nested models refers to ***fully time-dependent models***, for which firms can adjust freely with *iid* probability; this class is generated by the assumptions $\kappa = \infty$ and $H_\xi(0) = 1$. If in addition, $H_u(u) = 0$ for all $u < T$ and $H_u(T) = 1$—the u is a constant—then firms adjust at the fixed date T, analogous to the Taylor (1980) model. If H_u is an exponential distribution with parameter λ, then firms adjust on random *iid* dates, analogous to the Calvo (1983) model.

The second class refers to ***fully state-dependent models*** , in which firms always face the same *constant fixed cost* $\xi = \kappa$, i.e. $H_\xi(\kappa) = 1$ and $H_\xi(\xi) = 0$ for all $\xi < \kappa$. This structure is considered by Caballero and Engel (1991) in investment.

The third class refers to hybrid models with both time- and state-dependence in the adjustment cost structure. Consider H_u as an exponential distribution with parameter λ. If H_ξ is degenerate at zero, then firms face either a positive or a zero adjustment cost; this case is known in the pricing literature as the ***CalvoPlus*** model (Nakamura and Steinsson, 2010). Finally, within this class, we call ***generalized***

[6] Section 4 shows that the first-order autocorrelation of investment rates in the Chilean data is zero; this observation is also true for US plants, see Cooper and Haltiwanger (2006). Mean-reverting shocks generate a negative autocorrelation in investment rates at the plant level. This observation also justifies the absence of convex adjustment costs in our specification, as they a generate a positive autocorrelation in investment rates.

hazard models those with non-degenerate H_ξ. For example, $H_\xi \sim Uniform[0, \kappa]$, which generates the same adjustment hazard as in Khan and Thomas (2008).

The advantage of considering a flexible formulation is that it permits us to remain agnostic about their true nature of adjustment costs and allows the data to inform us about it. In the next sections, we focus the analysis on the three families specified above.

2.3 Dynamics of the aggregate capital stock

Given the firms investment policy, we are interested in characterizing deviation of aggregate capital from the steady-state. Since there is growth, we work with log deviations of aggregate capital detrended by productivity. To this end, we define three variables. First, we define the individual capital gap $\hat{k}_{\omega,t} \equiv \log(k_{\omega,t}/e_{\omega,t})$ as the log of the ratio of a firm's capital to its productivity. Second, we define the average of capital gap in the steady-state $\hat{k}_{ss} \equiv \mathbb{E}\left[\log(k_\omega/e_\omega)\right]$. The notation without time index t refers to moments computed with the steady-state distribution. Lastly, we define the *normalized capital gap* $x_{\omega,t} \equiv \hat{k}_{\omega,t} - \hat{k}_{ss}$ as a firm's capital gap minus the steady-state average.

With these definitions, we compute the aggregate capital detrended by productivity denoted with $\hat{K}_t := \mathbb{E}\left[\log(k_{\omega,t})\right] - \mathbb{E}\left[\log(e_{\omega,t})\right] = \mathbb{E}\left[\log(k_{\omega,t}/e_{\omega,t})\right]$. Then, the aggregate capital deviation from steady-state is equal to the average normalized capital gap $\mathbb{E}[x_{\omega,t}]$:

$$\hat{K}_t - \hat{K}_{ss} = \mathbb{E}\left[\log(k_{\omega,t}/e_{\omega,t})\right] - \mathbb{E}\left[\log(k_\omega/e_\omega)\right] = \mathbb{E}\left[\hat{k}_{\omega,t}\right] - \hat{k}_{ss} = \mathbb{E}[x_{\omega,t}], \tag{8}$$

Notice that, in this normalization, we first centralize the capital-gap distribution around its steady-state average and then we aggregate across firms. By the previous analysis, we may shift the focus from aggregate capital to moments of the normalized capital gaps. Finally, note that the dynamics of other aggregate variables, such as output deviations from steady-state, denoted by $\hat{Y}_t$, can also be expressed in terms of moments of normalized capital gaps: $\hat{Y}_t = \alpha\mathbb{E}[x_{\omega,t}]$.

Law of motion of capital gaps. To derive the law of motion of capital gaps, we use the firm policy. The *uncontrolled* capital gaps—not considering any investments—follow the process

$$d\tilde{x}_{\omega,t} = \nu dt + \sigma dW_{\omega,t} \tag{9}$$

where we use tildes to show explicitly that these variables evolve exogenously. For convenience, we define the total drift of capital gaps as $\nu \equiv -(\psi + \mu)$, which is negative and includes depreciation and the productivity trend. The initial conditions $\tilde{x}_{\omega,0}$ are exogenously given. By the discussion above, the initial condition of the uncontrolled capital gap is $\tilde{x}_{\omega,0} = \hat{k}_{\omega,0} - \hat{k}_{ss}$. In contrast, the *controlled* capital gaps—taking into account investments—evolve as

$$x_{\omega,t} = \tilde{x}_{\omega,t} + \sum_{\tau_{\omega,i} \leq t} \Delta x_{\tau_{\omega,i}}, \tag{10}$$

where the adjustment dates $(\tau_{\omega,i})$ and the investment rates $(\Delta x_{\tau_{\omega,i}})$ solve the firms' problem in (4). Finally, we define the reset capital gap $\hat{x}$ (the new capital gap, conditional on adjustment) and capital gap age (the time since its last adjustment) implicitly as:

$$\hat{x} \;=\; x_{\omega,\tau_{\omega,i}} \quad , \quad a_{\omega,t} \;=\; t - \max\{\tau_{\omega,i} : \tau_{\omega,i} \leq t\}. \tag{11}$$

A few things are worth noting. First, $\hat{x}$ is the same across firms and time, in other words, firms become identical after an adjustment. This is due to the fact that there is no fixed heterogeneity[7] and the adjustment costs are memoryless. Second, due to the normalization explained above, $\hat{x}$ is the capital–productivity ratio of adjusting firms *relative* to the average ratio in the economy. With respect to age, for some adjustment cost structures (e.g. Taylor), capital gap age is a relevant state for the investment decision. Even if it is not the case, the joint stochastic process of (x, a) does matter for aggregate dynamics. Therefore, we carry age as part of the firm's state. Lastly, we assume that steady-state policies hold along the transition path. This assumption is valid as long as the general equilibrium feedback from the aggregate distribution to the individual policies though prices is quantitatively insignificant. In this model, it is guaranteed by the linear preferences.

2.4 Steady-state and transitional dynamics

Now we define steady-state moments and our notion of transitional dynamics through the cumulative impulse response or CIR.

Steady-state moments. Consider the steady-state distribution of capital gaps and age, denoted by $F(x, a)$. For any two numbers $k, l \in \mathbb{N}$, we define the ergodic cross-sectional moment of capital gaps and age as

$$\mathbb{E}[x^k a^l] \;\equiv\; \int_x \int_a x^k a^l \; dF(x, a) \quad \forall k, l \in \mathbb{N}, \quad \textit{with} \quad \mathbb{E}[x] \;=\; 0. \tag{12}$$

Transitional dynamics. Fix an initial distribution of the state $F_0(x, a) \;=\; \mathbb{E}\left[\mathbb{1}_{\{(x_{\omega,0}, a_{\omega,0}) \leq (x,a)\}}\right]$. We define the Impulse-Response Function for the m-th moment of the capital gap distribution under the initial distribution F_0, denoted by $IRF_{m,t}(F_0)$, as the difference between its time t value and its ergodic value:

$$IRF_{m,t}(F_0, t) \;\equiv\; \underbrace{\mathbb{E}\left[x^m_{\omega,t}\right]}_{\text{transition}} - \underbrace{\mathbb{E}[x^m]}_{\text{steady-state}}. \tag{13}$$

Following Alvarez, Le Bihan and Lippi (2016), we define the Cumulative Impulse-Response (CIR), denoted by $\text{CIR}_m(F_0)$, as the area under the $IRF_{m,t}(F_0)$ curve across all dates $t \in (0, \infty)$:

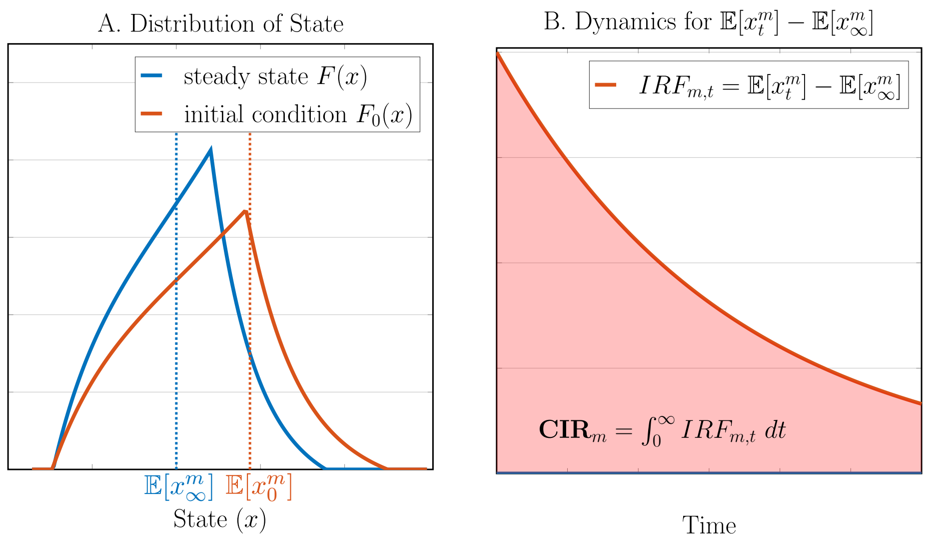
$$\text{CIR}_m(F_0) \;\equiv\; \int_0^\infty IRF_{m,t}(F_0, t) \; dt. \tag{14}$$

Figure I illustrates these two objects. In the left panel, we plot an initial marginal distribution $F_0(x)$ and the steady-state distribution $F(x)$, and also highlight the m-th moment of capital gaps $\mathbb{E}[x^m_0]$ which

[7] Section 5 introduces fixed heterogeneity.

will be tracked in its way towards steady-state.[8] In the right panel, the solid line represents the impulse-response of $\mathbb{E}[x_t^m]$, which is a function of time, and the area underneath it is the CIR. The CIR is our key measure of the convergence speed towards the steady-state. The smaller is the CIR, the faster the convergence.

Figure I – Cumulative Impulse-Response (CIR)



The following Lemma expresses the CIR in a recursive way, and it is a generalization of the result in Alvarez, Le Bihan and Lippi (2016) for any moment of the distribution $m > 1$, for an arbitrary Markovian stopping policy, and for any Markovian law of motion of the uncontrolled state.

Lemma 1. *The CIR can be written recursively as:*

$$CIR_m(F_0) \;=\; \int v_m(x,a) dF_0(x,a). \tag{15}$$

where the value function for an agent with initial state (x, a) is given by:

$$v_m(x,a) \;\equiv\; \mathbb{E}\left[\int_0^\tau \left(x_t^m - \mathbb{E}[x^m]\right) dt \Big| (x,a)\right] \tag{16}$$

The idea behind Lemma 1 is to exchange the integral across agents (the cross-section) with the infinite time integral (the time-series).[9] Then, it is key to recognize that the first time a firm adjusts its capital it incorporates all deviations from steady-state into its policy, and thus we only need to keep track of this firm until its *first* adjustment; any additional adjustments are driven purely by idiosyncratic conditions. The average of these additional adjustments equals the ergodic moment $\mathbb{E}[x^m]$, implying that the value function $v_m(x, a)$ equals zero after the first adjustment. For that reason, the infinite time integral gets substituted for an integral between $t = 0$ and the stopping date $t = \tau$.

[8] Abusing notation, we denote the marginal distributions as $F(x)$ and $F(a)$.

[9] This can be done due to the ergodic properties of the problem and the fact that moments are finite.

Initial distribution as δ-perturbations around steady-state. For ease of exposition, we interpret the initial distribution as a small perturbation of the steady-state distribution that can be parsimoniously described in terms of one parameter δ. As a baseline case, we focus on a particular type of perturbation that translates horizontally the distribution of capital gaps, i.e. a shock to the first moment of the distribution. If $f(x-\delta, a)$ denotes the initial density of capital gaps and $f_x(x,a)$ denotes its derivative with respect to x, we can approximate it as $f(x-\delta,a) \approx f(x,a) - \delta f_x(x,a)$. For $\delta < 0$, we observe that the initial density is equal to a right shift of the steady-state density. Afterwards, the distribution will evolve according to the agents' policies and will converge back to its steady-state. Under this interpretation, we will denote the $\text{CIR}_m(F_0)$ as $\text{CIR}_m(\delta)$.[10]

3 Theoretical Results

In this section, we establish the theoretical relationships between observable panel data, (possibly) unobservable steady-state moments and parameters, and the CIR. In particular, the CIR is expressed as a linear projection over the steady-state moments.

3.1 From microdata to steady-state moments

The first link connects the ergodic cross-sectional moments, the structural parameters of the stochastic process, and the reset state to the distribution of capital gaps Δx and adjustment dates τ in a panel of observations. The relevance of this result lies in that, in many applications, the state x is likely to be unobservable, but the adjustment sizes Δx and date τ are observable. This is the case in the investment model, as capital gaps are hard to observe but investment rates are readily available in the data.[11] Proposition 1 derives a set of relations between objects that we do observe—the distribution of investment and duration—with objects that we do not observe—the joint distribution of capital gaps and age.

Proposition 1. *Let $(\Delta x, \tau)$ be a panel of observations of adjustment size and inaction duration. Denote with $\mathbb{E}[\cdot], \mathbb{C}ov[\cdot]$ and $\mathbb{CV}[\cdot]^2 = \mathbb{V}[\cdot]/\mathbb{E}[\cdot]^2$, respectively, the cross-sectional average, the covariance, and the coefficient of variation squared, conditional on adjustment. Then, the following expressions hold:*

1. *The reset capital gap is given by:*

$$\hat{x} = \frac{\mathbb{E}[\Delta x]}{2}\left(1 - \mathbb{CV}[\tau]^2\right) + \frac{\mathbb{C}ov[\tau, \Delta x]}{\mathbb{E}[\tau]}. \tag{17}$$

2. *The drift and volatility of the capital gap process are recovered as:*

$$\nu = -\frac{\mathbb{E}[\Delta x]}{\mathbb{E}[\tau]} \quad ; \quad \sigma^2 = \frac{\mathbb{E}[\Delta x^2]}{\mathbb{E}[\tau]} + 2\nu\hat{x}. \tag{18}$$

[10] In Section 5 we consider more general perturbations around steady-state, such as mean-preserving spreads.

[11] Our formulas require us to compute the change in log capital gaps Δx in the data. Due to the continuity assumption for the idiosyncratic productivity, the changes in the capital-gap equal the observed investment rates: $\Delta x_{\omega,\tau} = \lim_{t\uparrow\tau_i} \log(K_{\omega,\tau_i}/K_{\omega,t}) - \lim_{t\uparrow\tau_i} \log(E_{\omega,\tau_i}/E_{\omega,t}) = \lim_{t\uparrow\tau_i} \log(K_{\omega,\tau_i}/K_{\omega,t})$. Therefore, we can compute the changes in the *capital gap* using changes in the *capital stock*.

3. If $\nu \neq 0$, the steady-state moments for any $m \geq 1$ are given by:

$$\mathbb{E}[x^m] = \frac{1}{m+1}\left\{\frac{\hat{x}^{m+1} - \mathbb{E}\left[(\hat{x}-\Delta x)^{m+1}\right]}{\mathbb{E}\left[\Delta x\right]} - \frac{\sigma^2}{2\nu}m(m+1)\mathbb{E}[x^{m-1}]\right\}, \tag{19}$$

$$\mathbb{E}[a] = \frac{\mathbb{E}[\tau]}{2}\left(1 + \mathbb{CV}[\tau]^2\right), \tag{20}$$

$$\mathbb{E}[x^m a] = \frac{1}{m+1}\left\{\frac{\mathbb{E}\left[\tau\left(\hat{x}-\Delta x\right)^{m+1}\right]}{\nu\mathbb{E}[\tau]} - \mathbb{E}[x^{m+1}] - \frac{\sigma^2}{2\nu}m(m+1)\mathbb{E}[x^{m-1}a]\right\}. \tag{21}$$

The proof is in the Appendix; it enumerates the formal assumptions for this proposition to hold. To show this result, we uses three tools: Ito's Lemma, the Optional Sampling Theorem (OST), and the equivalence between the cross-sectional distribution of agents and the occupancy measure.[12] To understand the logic of the proof, assume there is a single state x. Apply Ito's Lemma to x_t^{m+1} with the initial condition $\hat{x}$ (right after adjustment), integrate between 0 and τ, and use $x_\tau = \hat{x} - \Delta x_\tau$ to obtain:

$$\underbrace{(\hat{x}-\Delta x_\tau)^{m+1} - \hat{x}^{m+1}}_{\text{investment (observable)}} = \underbrace{\int_0^\tau \left[\nu(m+1)x_t^m + \frac{\sigma^2}{2}(m+1)m x_t^{m-1}\right] dt}_{\text{history of capital gaps (unobservable)}} + \sigma(m+1)\underbrace{\int_0^\tau x_t^m dW_t}_{\textit{noise}}. \tag{22}$$

Equation (22) shows that the distribution of investment is related to the history of capital gaps plus a noise term. While we cannot recover each individual history, we can recover the *average* history between adjustments. For this, take the expectation on both sides of (22), and observe that the noise term is a martingale with expectation zero by the OST, we have that

$$\underbrace{\mathbb{E}\left[(\hat{x}-\Delta x_\tau)^{m+1} - \hat{x}^{m+1}\right]}_{\text{moments of investment}} = \underbrace{\nu(m+1)\mathbb{E}\left[\int_0^\tau x_t^m dt\right] + \frac{\sigma^2}{2}(m+1)m\mathbb{E}\left[\int_0^\tau x_t^{m-1} dt\right]}_{\text{average capital gap during inaction}}. \tag{23}$$

The final step to relate the distributions of investment and capital gaps uses the occupancy measure. Intuitively, the average time that a single agent's state spends at a given value is proportional to the number of agents with a state equal to that same value, where the constant of proportionality is the expected time between adjustments $\mathbb{E}[\tau]$.[13] Therefore, instead of measuring the average capital gap between adjustments for an agent, we can measure the average capital gap across agents:

$$\underbrace{\mathbb{E}\left[(\hat{x}-\Delta x_\tau)^{m+1} - \hat{x}^{m+1}\right]/\mathbb{E}[\tau]}_{\text{moments of investment and duration}} = \underbrace{\nu(m+1)\mathbb{E}\left[x^m\right] + \frac{\sigma^2}{2}(m+1)m\mathbb{E}\left[x^{m-1}\right]}_{\text{average capital gap across agents}}. \tag{24}$$

Now we provide the economics behind the expressions in Proposition 1.

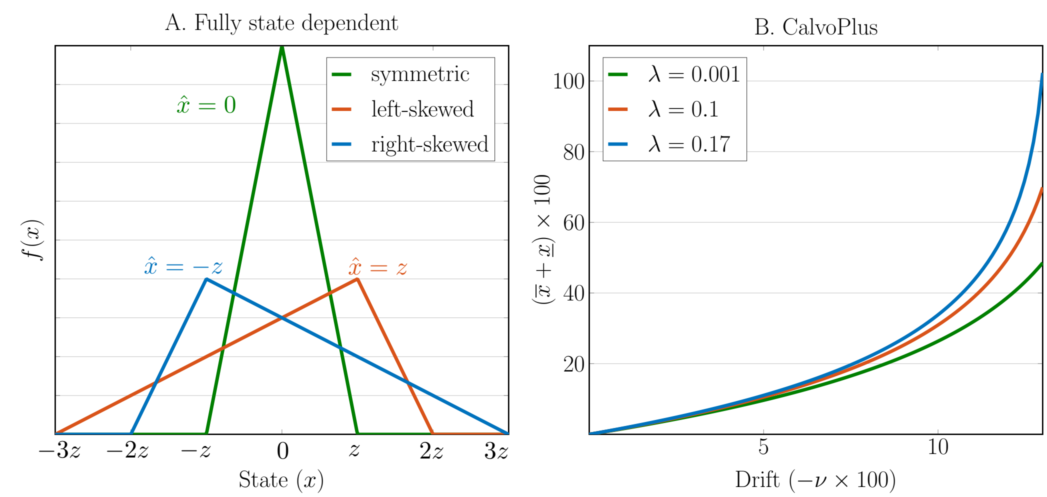
Reset state. Equation (17) shows how to recover the reset state $\hat{x}$ from the microdata; this expression is derived from the restriction imposed by the normalization of the ergodic mean to zero. It has two com-

[12]See Stokey (2009) for details.
[13]Formally, $\mathbb{E}\left[\int_0^\tau x_t^m dt\right] = \mathbb{E}[\tau]\mathbb{E}\left[x^m\right]$ for any m.

ponents that reflect how the reset state compensates for the drift, for the asymmetry in state-dependent policies, or a combination of both, ensuring that $\mathbb{E}[x] = 0$. To illustrate the compensation for drift, consider the family of fully time-dependent costs, which by construction do not exhibit asymmetric policies. In such models, $\mathbb{C}ov[\tau, \Delta x] = -\nu \mathbb{V}ar[\tau]$ and equation (17) collapses to $\hat{x} = -\nu \mathbb{E}[a]$, so that the reset state compensates the average accumulated drift between adjustments, centralizing the ergodic mean at zero.[14]

To illustrate the compensation for asymmetric policies, consider a driftless state and fully state-dependent costs, where the widths of the lower and upper inaction triggers relative to the reset point are $|-\underline{x} - \hat{x}|$ and $|\overline{x} - \hat{x}|$, respectively. Panel A of Figure II plots three distributions of capital gaps for different types of policies. First, symmetric policies (in green) necessarily imply $\hat{x} = 0$. Now consider an asymmetric inaction region such that the upper trigger is closer to $\hat{x}$ than the lower trigger, for example, $4z = |-\underline{x} - \hat{x}| > |\overline{x} - \hat{x}| = z$ for $z > 0$ (in red). In this case, the capital gap distribution is left-skewed and the covariance $\mathbb{C}ov[\tau, \Delta x]$ is positive: the longer the duration of inaction, the larger the probability of a positive investment. This implies a positive reset state $\hat{x} = \mathbb{C}ov[\tau, \Delta x]/\mathbb{E}[\tau] = z > 0$. Analogously, the reset state is negative for right-skewed distributions (in blue).

Figure II – Reset state for asymmetric policies in fully state-dependent models



Notes: Panel A describes the distribution of the state x in a fully state-dependent model. The symmetric distribution (green line) has widths of $(|\hat{x} - \underline{x}|, |\overline{x} - \hat{x}|) = (z, z)$, the left-skewed distribution $(|\hat{x} - \underline{x}|, |\overline{x} - \hat{x}|) = (4z, z)$, and right-skewed distribution $(|\hat{x} - \underline{x}|, |\overline{x} - \hat{x}|) = (z, 4z)$. The reset states are 0, z and $-z$, respectively. Panel B shows the levels of asymmetry in policy and drift that together imply a zero reset state $\hat{x} = 0$, for fixed parameters $(\sigma, \underline{x}) = (0.275, -0.49)$ and three values of $\lambda \in \{0.001, 0.1, 0.17\}$.

Lastly, regarding the interactions, policy asymmetry may dampen or amplify the effect of the drift in the reset state. For illustration, consider the CalvoPlus adjustment costs and fix a set of parameters. Panel B in Figure II shows the combination of values for $(-\nu \times 100, \overline{x} + \underline{x})$ such that the reset state is zero $\hat{x} = 0$. When $\nu = 0$, only symmetric policies generate a zero reset $\hat{x}$. As the drift increases, the upper limit $\overline{x}$ increases as well to compensate the drift. The covariance term informs about this interaction, as in this case $\hat{x} = -\nu\left(\mathbb{E}[\tau] - \mathbb{E}[a]\right) + \mathbb{C}ov[\tau, \Delta x]/\mathbb{E}[\tau]$. Note that the asymmetry is also increasing in the

[14] Proof: $\mathbb{C}ov[\tau, \Delta x] = \mathbb{E}[\tau \Delta x] - \mathbb{E}[\tau]\mathbb{E}[\Delta x] = \mathbb{E}\left[\tau(-\nu\tau - \sigma \int_0^\tau W_t)\right] + \nu \mathbb{E}[\tau]^2 = -\nu \mathbb{E}[\tau^2] + \nu \mathbb{E}[\tau]^2 = -\nu \mathbb{V}ar[\tau]$, where we have used the OST to kill the martingale $\mathbb{E}\left[\tau \int_0^\tau W_t\right] = 0$.

parameter λ (the arrival of free adjustment opportunities). This is because as λ increases, the fraction of state-dependent (asymmetric) investments decreases, and it must be compensated by a more asymmetric policy.

Mean and variance of capital gap growth. Expressions in (18), which extend those in Alvarez, Le Bihan and Lippi (2016) for the case with drift, provide a guide to infer the parameters of the stochastic process. The first expression shows how to infer the drift ν from the average investment rate in the data, scaled by the adjustment frequency. Intuitively, the average depreciation has to be equal the investment rate to have an ergodic distribution. The second expression shows how to infer the volatility σ from the dispersion in investment rates, scaled by the frequency and corrected by the drift. With a zero drift, higher idiosyncratic volatility reduces the average duration or increases the average investment size; with non-zero drift, investment's second moment also reflects the drift and (18) shows how to correct for it.

One interesting application of the results for non-zero drift would be to revisit the role of positive inflation for price-setting, as studied by Alvarez, Beraja, Gonzalez-Rozada and Neumeyer (2018) in Argentina, Gagnon (2009) for Mexico, and Blanco (2018) for the US.

Ergodic moments. Equation (20) relates average age to the average and the dispersion in duration, measured through the coefficient of variation. The relationship with the average duration is straightforward. To understand why the duration dispersion affects average age, recall a basic property in renewal theory: the probability that a random firm has an expected duration of inaction of τ is increasing in τ, i.e. larger stopping times are more representative in the capital-gap distribution.[15] Therefore, dispersion in duration reflects that there are firms that take a long time to adjust, and on top of that, those firms are more representative in the economy; this rises the average age.

Equations (19) and (21) provide recursive formulas to compute the ergodic moments using observed investment rates. For a given reset capital gap, the ergodic moments of capital gaps only depend on moments of investment, and we can ignore moments of duration. For example, set $\hat{x} = 0$ and $m = 2$, then equation (19) reads $\mathbb{E}[x^2] = \mathbb{E}\left[(\Delta x)^3\right]/3\mathbb{E}\left[\Delta x\right]$, relating the dispersion of capital gaps in the LHS to the skewness of investment rates in the RHS. For $m > 2$, the ratio of drift to idiosyncratic volatility also matters. This is not the case for the moments that interact x and a, as equation (21) shows.

Key assumptions and limitations. The first key assumption underlying Proposition 1 is the particular stochastic process for capital gaps. This is not a limitation, as different stochastic processes generate different mappings between data and steady-state moments that can be tested. For example, if one assumes an Ornstein–Uhlenbeck process instead, there is an additional equation that pins down the mean reversion parameter (see Section 5). The second key assumption is that the state only consists of the capital gap and its age. This is also not a limitation, as the proof can be generalized to any Strong Markov state. In a similar way, parametric assumptions on firms' policies are not necessary, as the proof can be generalized to arbitrary Markovian policy. The third key assumption is that $(\hat{x}, \nu, \sigma)$ is a constant vector of numbers. If this is not the case, the theory can still recover some cross-sectional moments of

[15] This property has been widely studied in labor economics when thinking about long-term unemployment. For example, Mankiw (2014)'s textbook Principles of Macroeconomics mentions that: "[...] many spells of unemployment are short, but most weeks of unemployment are attributable to long-term unemployment".

these variables. For example, in a model with stochastic volatility and zero drift, Baley and Blanco (2019) show how to recover the average level as $\mathbb{E}[\sigma_i^2] = \mathbb{E}[\Delta x^2]/\mathbb{E}[\tau]$.

3.2 From steady-state moments to transitional dynamics

This sections shows the following properties. In the class of adjustment cost specifications described in Section 2.2, the CIR is a linear combination of ergodic moments, where the coefficients depend on the stochastic process parameters and adjustment cost specification.

3.2.1 Fully time-dependent models

With fully time dependent adjustment costs, the distribution of stopping times is independent of the capital-gap. Proposition 2 characterizes the CIR_m for this class of models.

Proposition 2. *In fully time-dependent models, for every $m \geq 1$:*

$$CIR_m(\delta)/m\delta \;=\; \sum_{i\,=\,1}^{m} \left(-\frac{\sigma^2}{2\nu}\right)^{m-i} \left(\mathbb{E}\left[x^{i-1}a\right] + \mathbb{1}_{\{i\geq 2\}}\frac{\sigma^2}{2\nu}(i-1)\mathbb{E}\left[x^{i-2}a\right]\right) \;+\; o(\delta). \tag{25}$$

We leave the discussion of the proof for later, as this is a special case of the generalized hazard model explained below. To understand the economics behind this relationship, consider first the case $m = 1$, where we have that

$$\text{CIR}_1(\delta)/\delta \;=\; \mathbb{E}\left[a\right] \;+\; o(\delta). \tag{26}$$

This expression for the CIR_1 was first obtained in Alvarez, Lippi and Paciello (2015) in the context of a pricing model of inattentive producers with time-dependent revelation of information. Average age provides information about the speed at which the average firm adjusts to the perturbation from the steady-state. Remember, average age is different to expected duration, and it also considers its dispersion. Thus the older is average capital in the economy and the more dispersed its duration, the longer the transition. Consider a frictionless limit in which all firms continuously invest to brings capital gaps to zero. Since capital in all firms would have age equal to zero, the economy reaches its steady-state immediately.

Now consider the case $m = 2$, where we have that

$$\text{CIR}_2(\delta)/2\delta \;=\; \mathbb{C}ov\left[x,a\right] - \frac{\sigma^2}{2\nu}\mathbb{E}[a] \;+\; o(\delta). \tag{27}$$

To fix ideas, assume $\nu < 0$ so that the second term is positive. It is clear that a positive δ-perturbation could in principle generate a positive or negative CIR for the second moment depending on the value of the covariance. In turn, the value of the covariance depends on the model. Using an alternative expression for CIR_2 in terms of the first three moments of the duration distribution,[16] we can derive expressions for the Taylor and the Calvo models: $CIR_2^{Taylor}(\delta)/\delta = -\nu\mathbb{E}[\tau]\left(\sigma^2/2\nu^2 - \mathbb{E}[\tau]/12\right)$ and $CIR_2^{Calvo}(\delta)/\delta = -\nu\mathbb{E}[\tau]\left(\sigma^2/2\nu^2 - \mathbb{E}[\tau]\right)$. For the same average duration, the two models may imply

[16] $\text{CIR}_2(\delta)/\delta = \nu\left[\frac{\mathbb{E}[\tau^3]}{3\mathbb{E}[\tau]} - \frac{\mathbb{E}[\tau^2]^2}{4\mathbb{E}[\tau]^2} - \frac{\sigma^2}{\nu^2}\frac{\mathbb{E}[\tau^2]}{4\mathbb{E}[\tau]}\right]$. In Taylor, we substitute $\mathbb{E}[\tau^m] = \mathbb{E}[\tau]^m$; while in Calvo we substitute $\mathbb{E}[\tau^m] = m!\mathbb{E}[\tau]^m$.

different signs for the transitional dynamics of the second moment; this is due to differences in the tail and the mode ($\hat{x} = -\nu\mathbb{E}[a] > 0$) of the duration distribution generated by each model.

3.2.2 CalvoPlus model

Consider CalvoPlus adjustment costs. As explained in Section 2.2, this adjustment cost structure nests the Calvo model, the fully state-dependent model, and their combination. This model also delivers a very simple expression for the CIR, as a linear combination of two simple moments of the joint distribution $F(x,a)$ for all cases.

Proposition 3. *In the CalvoPlus model, for every $m \geq 1$:*

$$CIR_m(\delta)/\delta \;=\; \frac{\mathbb{E}\left[x^{m+1}\right] - \nu\mathbb{C}ov\left[x^m, a\right]}{\sigma^2} \;+\; o(\delta). \tag{28}$$

The proof uses the induction hypothesis. The guess for the moments in which the CIR is projected is suggested by the generalized hazard model below. Equation (28) shows that, up to first oder, there exists a one-to-one mapping between ergodic moments and the CIR. To build the intuition for this result, let us focus in the case $m = 1$. Consider $\nu = 0$ so that the CIR_1 is given exclusively by $\mathbb{V}ar[x]/\sigma^2$, the dispersion of capital gaps normalized by the idiosyncratic volatility. This dispersion encodes information about agents' *responsiveness* to idiosyncratic shocks (the higher is the ratio the less responsive is the policy), and in turn, the responsiveness determines the speed of convergence to the steady-state. In the case $\nu \neq 0$, the covariance between capital gap age and the investment rate appears in the expression to correct for the dispersion generated by the drift (which is orthogonal to the dispersion due to idiosyncratic shocks).

Now consider $m = 2$. As with the time-dependent models, the dynamics of the second moment can have either sign. Assume $\nu < 0$ and σ sufficiently large. Then $\mathbb{C}ov\left[x^2, a\right] > 0$, since larger age is associated with larger accumulated gaps. Then the sign of the CIR_2 depends on the asymmetry of the capital gap distribution, measured through its third moment $\mathbb{E}[x^3]$, which in turn depends on how much an asymmetric policy compensates the drift (recall Figure II).

Relation to the literature. With zero drift and a symmetric policy ($\hat{x} = 0$), one can recover ergodic moments from data with an expression similar to (19) given by[17]

$$\mathbb{E}[x^m] \;=\; \frac{2}{(m+1)(m+2)} \frac{\mathbb{E}\left[(-\Delta x)^{m+2}\right]}{\mathbb{E}\left[\Delta x^2\right]}. \tag{29}$$

Combining the CIR_1 in the CalvoPlus model (28) with the expression for σ^2 in (18) and the new expression for $\mathbb{E}[x^2]$ in (29), we obtain the well-known "kurtosis" formula in Alvarez, Le Bihan and Lippi (2016):

$$\text{CIR}_1(\delta)/\delta \;=\; \frac{\mathbb{E}[x^2]}{\sigma^2} \;=\; \frac{\mathbb{E}[\tau]}{\mathbb{E}[\Delta x^2]}\mathbb{E}[x^2] \;=\; \frac{\mathbb{E}[\tau]}{6}\frac{\mathbb{E}\left[\Delta x^4\right]}{\mathbb{E}[\Delta x^2]^2} \;=\; \frac{\mathbb{E}[\tau]}{6}\mathbb{K}ur[\Delta x]. \tag{30}$$

We complement their analysis by introducing the link to the ergodic variance $\mathbb{V}ar[x]/\sigma^2$.

[17] See Web Appendix for the observation formulas for $\nu = 0$.

3.2.3 Generalized hazard models

Finally, we proceed to characterize the CIR for the most general case, where we do not impose any structure to the adjustment hazard. The strategy is as follows. We set as an upper bound the value of the CIR obtained for fully time-dependent models. This value reflects the adjustments along the *intensive margin* exclusively, as agents cannot affect the duration of inaction. Then, we consider an additional term that captures adjustments through the *extensive margin* or changes in the duration, which are available in models with a state-dependent component, that can only reduce the value of the CIR.

There are two challenges involved in bringing discipline to the extensive margin. First, the extensive margin does not only depend on the immediate response of the aggregate duration, but it also reflects all current and future changes in duration. Second, even if we had the whole sequence of adjustments in duration that follows a perturbation, the extensive margin also depends on the capital gaps of the particular set of firms selected to invest. Next, we develop a way to discipline these margin in a general setting.

How to characterize the extensive margin? In order to characterize the extensive margin, we introduce the following auxiliary function:

$$g_m(x) \;\equiv\; \mathbb{E}[x_\tau^m | \hat{x} + x] - \mathbb{E}^{\hat{x}}\left[(x_\tau + x)^m | \, \hat{x}\right]. \tag{31}$$

The first term in (31) equals the expected capital gap at the moment of adjustment when the initial condition is $\hat{x} + x$; while the second term equals the expected capital gap at the moment of adjustment plus a deterministic increase of size x when the initial condition is $\hat{x}$. The difference between these two functions of x provides information of how the stopping time policy depends on the initial condition and how it correlates with the capital gap. To see this clearly, recall that the expected capital gap at the moment of adjustment is equal to $x_\tau = \hat{x} - \Delta x$ and we can re-express $g_m(x)$ in the following way:

$$g_m(x) \;=\; \mathbb{E}\left[(\hat{x} + x - \nu\tau^{\hat{x}+x} - \sigma W_{\tau^{\hat{x}+x}})^m\right] - \mathbb{E}\left[(\hat{x} + x - \nu\tau^{\hat{x}} - \sigma W_{\tau^{\hat{x}}})^m\right], \tag{32}$$

where τ^z is the stopping time with initial condition z. In equation (32) we observe that if τ is independent of the initial condition, i.e. $\tau^{\hat{x}+x} = \tau^{\hat{x}} = \tau$, as in time-dependent models, then $g_m(x) = 0$ for all x, implying that the extensive margin is null. For other models, the derivatives of g_m with respect to the initial condition, evaluated at zero, provides a micro-elasticity of firms' idiosyncratic response to the new initial conditions though changes in its stopping time τ. With this function at hand, we proceed to characterize the CIR_m for the general hazard model.

Proposition 4. *Assume $\nu \neq 0$. In the generalized hazard model, for every $m \geq 1$:*

$$CIR_m(\delta)/\delta \;=\; \mathcal{Z}_m - \mathbb{E}\left[x^m\right]\Theta_0 \;+\; o(\delta), \tag{33}$$

where the margins of adjustment are given by:

$$
\begin{aligned}
&\textit{(total)} && \mathcal{Z}_m &=& \; \Gamma_m + \Theta_m - \frac{\sigma^2 m}{2\nu}\mathcal{Z}_{m-1}, && (34)\\
&\textit{(intensive)} && \Gamma_m &=& \; m\mathbb{E}\left[x^{m-1}a\right] + \mathbb{1}_{\{m\geq 2\}}\frac{\sigma^2 m(m-1)}{2\nu}\mathbb{E}\left[x^{m-2}a\right], && (35)\\
&\textit{(extensive)} && \Theta_m &=& \; \sum_{j\,=\,0}^{\infty}\theta_{m,j}\mathbb{E}[x^j], \qquad \textit{with} && (36)\\
& && \theta_{m,j} &\equiv& \; \frac{1}{\nu}\sum_{k\geq j}^{\infty}\frac{\hat{x}^{k-j}}{k!j!}\left[\frac{d^{k+1}g_{m+1}(0)/m+1}{dx^{k+1}} - \frac{d^k g_m(0)}{dx^k}\right]. && (37)
\end{aligned}
$$

Equation (33) measures the total effect of the δ-perturbation as an area with height of δ and a base given by $\mathcal{Z}_m$ constructed with two components: Γ_m, which measures adjustments through the intensive margin and its expression is identical to that for time-dependent models in equation (25); and Θ_m, which measures adjustments through the extensive margin and its expression is a linear combination of ergodic moments. In turn, the weights $\theta_{m,j}$ or micro-elasticities are written in terms of derivatives of the auxiliary function g_m, evaluated at zero.

The proof of Proposition 4 is constructive and has two steps. The first step constructs two Bellman equations to characterize the intensive Γ_m and the extensive Θ_m margins of adjustment. The second step we proceeds similarly to Proposition 1 and expresses each Bellman equation as a function of ergodic moments using a combination of Ito's Lemma, Optional Sampling Theorem and the occupancy measure. We cannot solve analytically for the extensive margin, but we do it numerically.

Necessity of micro-elasticities. Proposition 4 uses steady-states moments plus micro-elasticities to discipline the dynamics of the distribution capture in the CIR, especially the extensive margin. A natural question arises: Are the micro-elasticities necessary to discipline the extensive margin? The answer is yes, as we show with the following counterexample.[18]

Let $T \equiv \mathbb{E}[\tau]$ denote average duration. Consider an inaction model with adjustments at fixed dates (Taylor-type) and a standard Ss model; assume away idiosyncratic shocks ($\sigma = 0$) and allow for a non-zero drift ($\nu \neq 0$). In these two models there exists a steady-state with a uniform distribution of capital-gaps and an investment distribution with an atom at $-\nu T$; thus they produce the same ergodic moments.[19] Now, let us study transitional dynamics for $\delta < 0$. As stated by the theory, in both models the intensive margin is equal to the average age: $\Gamma_1 = T/2$. Since the Taylor model is time-dependent, $\Theta_1 = 0$ and its CIR_1 equals: $\text{CIR}_1^{Taylor}(\delta)/\delta \approx T/2$. In the Ss model, the extensive margin is equal to $\Theta_1 = \theta_{1,0}\mathbb{E}[x^0] = -T/2$ (as $\theta_{1,j} = 0$ for $j > 0$), and its CIR_1 equals $\text{CIR}_1^{Ss}(\delta)/\delta \approx T/2 - T/2 = 0$. This result mirrors the classic money neutrality outcome in Caplin and Spulber (1987) in lumpy adjustment models.

The previous counterexample illustrates that two models may produce the same steady-state statistics, but nevertheless, they can exhibit completely different transitional dynamics. Our explanation lies in the

[18] See Web Appendix B.2 for the proof.

[19] These two models have several ergodic time-varying distributions that depend on the initial condition. To generate an unique ergodic distribution for any initial condition, we add a small and random probability of free adjustment. Besides generating a unique ergodic distribution, it gives differentiability to the CIR at $\delta = 0$ in the Ss model.

differences in micro-elasticities, zero in the Taylor-type model and $-\Gamma_1$ in the Ss model. Therefore, there are cases for which the micro-elasticities are relevant objects for characterizing the extensive margin, and our theory can guide researchers in finding experiments or exogenous variation to compute them.

In the following section we apply our theoretical results to the data.

4 Application: Investment Dynamics

In this section, we revisit the investment model from Section 2 and apply our tools using establishment-level data to gauge the magnitude of capital misallocation in steady-state as well as the transitional dynamics of aggregate capital.

4.1 Data description

Sources. We use yearly microdata on the cross-section of manufacturing plants in Chile from the Annual National Manufacturing Survey (*Encuesta Nacional Industrial Anual*) for the period 1979 to 2011.[20] Information on depreciation rates and price deflators used to construct the capital series comes from National Accounts and Penn World Tables. We report statistics for the total capital stock as well as for structures, a capital category that represents approximately 30% of all investment in the manufacturing sector and features the strongest lumpy behavior. We consider all plants that appear in the sample for at least 10 years (more than 60% of the sample).[21]

Capital stock and investment rates. We construct the capital stock series through the perpetual inventory method (PIM).[22] Let firm's ω stock of capital of type j on year t be given by:

$$K_{\omega,j,t} \;=\; (1-\delta_j)K_{\omega,j,t-1} + I_{\omega,j,t}/D_{j,t} \quad \text{for } K_{\omega,j,t_0} \text{ given,} \tag{38}$$

where the depreciation rate δ_j is a type-specific time-invariant depreciation rate; price deflators $D_{j,t}$ are gross fixed capital formation deflators by capital type; and initial capitals K_{i,j,t_0} are given by the firms' self-reported nominal stock of capital of type j at current prices on the first year in which they report a non-negative capital stock. Gross nominal investment $I_{\omega,j,t}$ is based on the information on purchases, reforms, improvements, and sales of fixed assets reported by each plant in the survey:

$$I_{\omega,j,t} \;=\; puchases_{\omega,j,t} + reforms_{\omega,j,t} + improvements_{\omega,j,t} - sales_{\omega,j,t} \tag{39}$$

[20]This data has been used by Liu (1993) to examine the role of turnover and learning on productivity growth; by Tybout (2000) to survey the state of the manufacturing sector in developing economies; and more recently, by Oberfield (2013) to study productivity and misallocation during crises.

[21]The Online Data Appendix presents all the details on the data, construction of variables, and analysis for each capital category separately: structures, machinery, equipment and vehicles. Additionally, we repeat all the analysis for Colombia, using the Annual Manufacturers Survey *Encuesta Anual Manufacturera* for the period 1995-2016. This data has been used by Eslava, Haltiwanger, Kugler and Kugler (2004, 2013) to study the effect of structural reforms and trade liberalization on aggregate productivity. Results are very similar across the two countries, so we focus here the analysis for the Chilean data.

[22]See Section A.2 for details on the PIM method and several checks on the data.

Once we construct the investment and capital stock series, we define the investment rate $i_{\omega,j,t}$ as the ratio of real gross investment to the capital stock:

$$i_{\omega,j,t} \equiv \frac{I_{\omega,j,t}/D_{j,t}}{K_{\omega,j,t}}. \tag{40}$$

Table I presents descriptive statistics on investment rates. Inaction is defined as investment below 1% in absolute value; positive spikes are investments above 20% and negative spikes below −20%. Besides the statistics for Chile, we include the numbers reported by Cooper and Haltiwanger (2006) for 7,000 US manufacturing plants between 1972 and 1988.[23] Structures presents an inaction rate of 76% and a large fraction of positive spikes of 31%. Investment rates are very asymmetric (the frequency of positive investments is larger than the frequency of negative investments) and serially uncorrelated, as in the US data. The zero correlation of investment rates in the data is consistent with the model we developed in the previous section.

Table I – Investment Rates Statistics

	Structures	**Total**	**US**
Average Investment	10.3	18.3	12.2
Positive Fraction ($i > 1\%$)	22.8	56.7	81.5
Negative Fraction ($i < -1\%$)	1.4	4.0	10.4
Inaction rate ($\lvert i\rvert <= 1\%$)	75.8	39.3	8.1
Spike rate ($\lvert i\rvert > 20\%$)	10.2	24.3	20.4
Positive spikes ($i > 20\%$)	9.7	23.2	18.6
Negative spikes ($i < -20\%$)	0.5	1.0	1.8
Serial correlation	0.0	0.0	0.1

Sources: Own calculations using plant-level data for Chile. The time period is 1979-2011 for plants that appear in the sample for at least 10 years. US refers to data in Cooper and Haltiwanger (2006) which covers manufacturing plants total investment in the US from 1972 to 1988. Investment rates reported in this table are computed as real investment divided by initial capital. We use perpetual inventories to compute capital stock. We eliminate investment rates below the 1st percentile and above the 99th percentile of the investment rate distribution.

4.2 Construction of capital gaps and duration

To apply the theory, for each firm ω and each inaction spell k, we record the capital gap change upon action $\Delta x_{\omega,k}$ and the spell's duration $\tau_{\omega,k}$. Recall that the capital gap change is given by the log difference in the capital stock between adjustment dates:

$$\Delta x_{\omega,k} = x_{t_{\omega,k}} - x_{t_{\omega,k-1}} = \log\left(K_{t_{\omega,k}}/K_{t_{\omega,k-1}}\right) = \log\left(1 + i_{\omega,t}\right) \tag{41}$$

[23] The Data Online Appendix shows numbers reported in Zwick and Mahon (2017) from tax records for US firms. In particular, the weighted inaction rate across firms is 30% in their unbalanced panel.

Using the information on investment rates, we construct capital gap changes as:

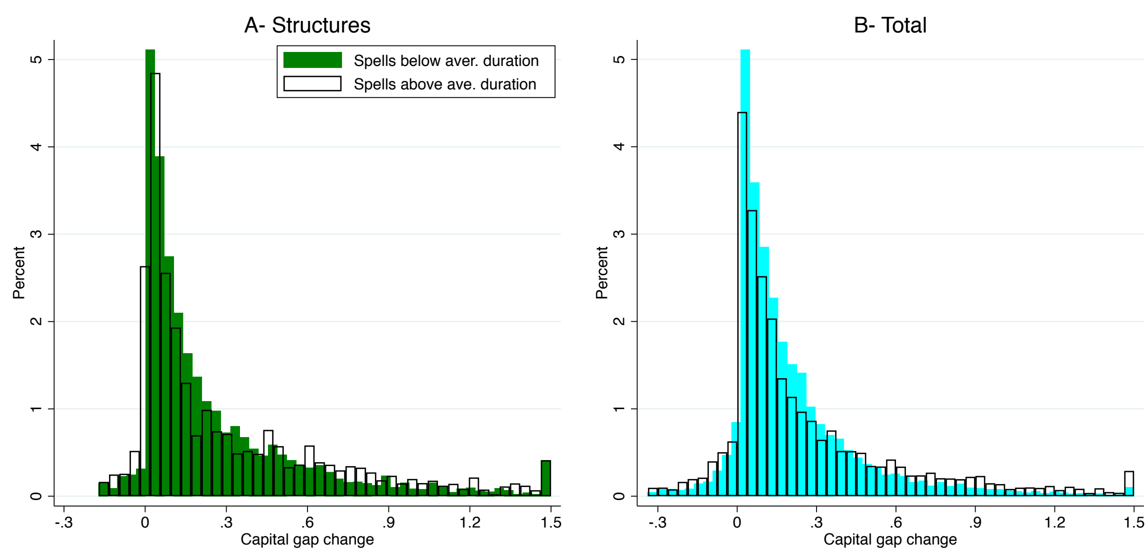
$$\Delta x_{\omega,k} = \begin{cases} \log(1+i_{\omega,k}) & \text{if} \quad |i_{\omega,k}| > \underline{i} \\ 0 & \text{if} \quad |i_{\omega,k}| < \underline{i}, \end{cases} \tag{42}$$

where $\underline{i} > 0$ is a parameter that captures the idea that small maintenance investments do not incur the fixed cost. Following Cooper and Haltiwanger (2006), we set $\underline{i} = 0.01$, such that all investments smaller than 1% in absolute value are excluded and considered as inaction. Finally, we compute a spell's duration as the difference between two adjacent adjustment dates:[24]

$$\tau_{\omega,k} = t_{\omega,k} - t_{\omega,k-1}. \tag{43}$$

Figure III plots the cross-sectional distribution of capital gap changes for structures and the total capital stock. In each figure, we show the distribution for two subsamples: observations with spell duration above the average (dark bars) and spell duration below the average (white bars). Interestingly, the two populations present the same behavior in terms of adjustment size: there is a large mass concentrated at low levels of capital gap changes and there are a few firms that have very large adjustments, and the distribution is asymmetric. Moreover, given that the two distributions lie on top of one another, there is no apparent correlation between adjustment size and duration.

Figure III – Distribution of capital gap changes by duration



As the next step, we put the theory to work by computing the cross-sectional statistics of the capital gap changes and duration in order to back out the parameters of the stochastic process as well as the ergodic moments. With these objects at hand, we use the theory to provide evidence against fully time-dependent models and in favor of hybrid models, and study the transitional dynamics via the CIR.

[24]See the Online Data Appendix for corrections due to duration dependence and unobserved heterogeneity.

4.3 Putting the theory to work

The relationships derived in Proposition 1 tell us how to use cross-sectional data on capital gaps and duration to pin down the parameters of the productivity process, the reset point, as well as the ergodic moments, which in turn are used to construct the CIR. Table II summarizes the statistics calculated from the microdata which serve as inputs into the formulas, as well as the theory's output.

Table II – Inputs from Micro Data and Outputs from the Theory

	Inputs from Data			**Outputs from Theory**	
	Structures	**Total**		**Structures**	**Total**
Frequency			**Parameters**		
$\mathbb{E}^x[\tau]$	2.441	1.714	ν	-0.111	-0.129
$\mathbb{CV}^2[\tau]$	1.093	0.855	σ^2	0.076	0.067
			$\hat{x}$	0.013	0.039
Capital Gaps			**Steady State Moments**		
$\mathbb{E}^x[\Delta x]$	0.271	0.220	$\mathbb{V}ar[x]$	0.228	0.149
$\mathbb{E}^{\hat{x}}[\Delta x^2]$	0.192	0.132	$\mathbb{E}[a]$	2.554	1.590
$\mathbb{E}^{\hat{x}}[(\hat{x}-\Delta x)^3]$	-0.186	-0.099	$\mathbb{C}ov[a,x]$	0.914	0.417
Covariances			**Transitional Dynamics**		
$\mathbb{C}ov^x[\tau,\Delta x]$	0.063	0.040	$\mathbb{V}ar[x]/\sigma^2$	3.014	2.240
$\mathbb{E}^{\hat{x}}[\tau(\hat{x}-\Delta x)^2]$	0.534	0.254	$-\nu\mathbb{C}ov[a,x]/\sigma^2$	1.340	0.803
			$\text{CIR}_1(\delta)$	4.354	3.043

Sources: Authors' calculations using establishment-level survey data for Chile.

The left part of Table II shows the inputs from the data: cross-sectional statistics for frequency, capital gaps, and covariances between them. The right part of the table shows the outputs from our theory: estimated parameters $\nu, \sigma^2, \hat{x}$ and ergodic moments $\mathbb{V}ar[x], \mathbb{C}ov[x,a]$, and the $\text{CIR}_1(\delta)$ for the CalvoPlus model.

4.3.1 Inputs from microdata

Consider first the distribution of expected times τ. We obtain an average expected time to adjustment of $\mathbb{E}[\tau] = 2.4$ years with a large dispersion, suggesting substantial heterogeneity in adjustment times across plants. Now consider the distribution of capital gaps; it has an average of $\mathbb{E}[\Delta x] = 0.27$ and a second moment of $\mathbb{E}[\Delta x^2] = 0.19$, and it is right-skewed. The covariance between adjustment size and expected time is almost zero $\mathbb{C}ov[\tau, \Delta x] = 0.06$. As we discuss below, this zero covariance is one of the key statistics that allows us to discern across inaction models.

4.3.2 Outputs from theory: parameters

Let us know explain the parameter values implied by our formulas. From (18), the implied drift, which captures the depreciation rate, productivity growth, and changes in relative prices, equals

$$\nu \;=\; -\frac{\mathbb{E}[\Delta x]}{\mathbb{E}[\tau]} \;=\; -\frac{0.271}{2.441} \;=\; -0.111, \tag{44}$$

and the volatility of idiosyncratic shocks equals

$$\sigma \;=\; \sqrt{\underbrace{\frac{\mathbb{E}[\Delta x^2]}{\mathbb{E}[\tau]}}_{0.08} + \underbrace{2\nu\hat{x}}_{0.002}} \;=\; 0.27. \tag{45}$$

Note that the main component that drives the volatility estimate is the second moment of capital gap changes, normalized by expected duration, whereas the drift term is negligible. The calibration for the volatility of innovations in the literature falls within a very wide range, from 0.052 in Khan and Thomas (2008) to 0.121 in Winberry (2016) to 0.202 in Bachmann, Caballero and Engel (2013).[25] It is worth noting that these calibrations are done jointly with the fixed adjustment cost within a particular inaction model. In contrast, our volatility estimate is pinned down directly through our model–independent mapping between data and parameters. Lastly, the observation formula for the reset capital gap in (17) implies that, upon adjustment, capital gaps are reset on average 1.3% above the average capital gap:

$$\hat{x} \;=\; \underbrace{\frac{\mathbb{E}[\Delta x]}{2}}_{0.14}\underbrace{\left(1 - \mathbb{CV}[\tau]^2\right)}_{-0.23} + \underbrace{\frac{\mathbb{C}ov[\tau, \Delta x]}{\mathbb{E}[\tau]}}_{0.02} \;=\; 0.013 \tag{46}$$

As with the covariance, the reset state together with the drift provide very useful information to tell apart families of inaction models, as we explain below.

4.3.3 Output from theory: ergodic moments.

According to the observation formula (19), the steady-state dispersion of capital gaps $\mathbb{V}ar[x]$—a notion of misallocation—can be expressed in terms of capital gap changes and the reset point as follows:

$$\mathbb{V}ar[x] \;=\; \frac{\hat{x}^3 - \mathbb{E}[(\hat{x} - \Delta x)^3]}{3\mathbb{E}[\Delta x]} \;=\; 0.23, \tag{47}$$

where the cubic powers capture asymmetries in the distribution. The average age $\mathbb{E}[a]$ is recovered using information about the average and the dispersion of adjustment times from (20). Following our

[25] The original numbers used in those papers are 0.022 and 0.049, respectively. Since we abstract from labor and our productivity is rescaled, we must adjust their volatilities by a factor $1/1-\alpha$ in order to make their numbers comparable to ours. We assuming a labor share of $\alpha = 0.58$ and obtain the numbers above. Additionally, for Bachmann, Caballero and Engel (2013) and Winberry (2016), we convert their quarterly volatilities $\sigma^q = 0.024, 0.047$ to yearly taking into account their yearly growth rate standard deviation: $\sigma^a = \sigma^q\sqrt{\frac{\rho^4-1}{1-\rho^2} + 1 + \rho^2 + \rho^4 + \rho^6}$), with $\rho = 0.94, 0.86$. Lastly, for Bachmann, Caballero and Engel (2013), we only consider the idiosyncratic shocks (excluding the sectorial shocks). Recall that we only consider here structures.

earlier discussion on renewal theory—larger stopping times are more representative in the sample—the heterogeneity in expected times increases the average age:

$$\mathbb{E}[a] \;=\; \underbrace{\mathbb{E}[\tau]/2}_{1.2}\underbrace{\left(1+\mathbb{CV}[\tau]^2\right)}_{2.1} \;=\; 2.52. \tag{48}$$

Lastly, equation (21) implies that the covariance between age and adjustment size suggested by the data is positive:

$$\mathbb{C}ov[x,a] \;=\; \frac{1}{2\nu}\left(\underbrace{\frac{\mathbb{E}\left[\tau\left(\hat{x}-\Delta x\right)^2\right]}{\mathbb{E}[\tau]}}_{0.22}-\underbrace{\mathbb{V}ar[x]}_{0.23}-\underbrace{\sigma^2\mathbb{E}[a]}_{0.20}\right) \;=\; 0.914. \tag{49}$$

This positive covariance between capital gaps and capital age means that the capital holdings of plants that have not adjusted in a long time (their capital is old) are *above* the gap of firms that have recently adjusted. In the next section, we discuss how this covariance as well as other objects compute above allows us to discern across models of inaction.

4.4 Discerning between inaction models

Let us recall the only two assumptions we have made to establish the link between data and ergodic moments and parameters: (i) capital gaps follow a Brownian motion, and (ii) the reset state $\hat{x}$ is constant across plants and time. We have not assumed any particular inaction model or parametric restriction of the plants' state besides those imposed to the capital gaps x. Clearly, plants may have other drivers of their investments besides capital gaps and we do not impose any structure on those.

Evidence against fully time-dependent models. A first source of evidence against this family is the zero covariance between duration and adjustment size. According to our formulas, in fully time dependent models it equals $\mathbb{C}ov[\tau,\Delta x] = \mathbb{V}ar[\tau] = 6.5$, which is extremely large compared to its tiny value in the data. A second source is the reset state: if the driver of inaction was fully time-dependent, then the reset state should be equal to $\hat{x}^{\text{time dep}} = -\nu\mathbb{E}[a] = -0.111 * 2.554 = 0.279$, twenty times larger that the reset state implied by the data $\hat{x}^{data} = 0.013$. We conclude that there exists an important state-dependent component that trades-off a lower reset state with an asymmetric policy in order to compensate the drift.

Mixed evidence of state-dependent models. Fully state dependent models are characterized by an increasing adjustment hazard: the covariance between duration of inaction and the adjustment size is positive. Even more so if there is a negative drift: the longer the inaction period, the larger the amount of idiosyncratic shocks that have accumulated. For this reason, one would expect a positive covariance: the longer the inaction period, the stronger the effect of the drift; consequently, upon taking action, the investment rate should be larger. Again, this is invalidated by the data, as the covariance is very small. Recall that the drift in capital gaps is equal to minus 11%. As firms that reset their capital gaps do it to a point 1.3% above the average, the part of the drift that is not compensated by the reset state (9.7%) must be accommodated through an asymmetric policy. This is in favor of state-dependent models.

Evidence in favor a hybrid model. The data points towards a hybrid model with both time and state-dependent components. A tiny covariance between adjustment sizes and duration, $\mathbb{C}ov[\tau, \Delta x] \approx 0$, which in turn implies a positive covariance between capital gaps and capital age of $\mathbb{C}ov[x, a] \approx 1$. This provides evidence that plants compensate part of the structural shocks they receive. However, compensation is not total. This can be seen with the ratio $\mathbb{V}ar[x]/\sigma^2 = 2.8$, which is quantitatively close to the expected duration $\mathbb{E}[\tau] = 2.4$. This suggests that the dispersion—a measure of ex-post heterogeneity—is almost equal to the fundamental volatility σ^2 times expected duration. In other words, there is a large passthrough of productivity shocks to capital misallocation, signaling inefficient capital adjustments. Moreover, comparing this result to (18), implies that $2\nu\hat{x} \approx 0$: the large negative drift of -11% is undone by firms through their adjustment policy.

While some of these moments appear counterintuitive, we want to stress again that they come directly from the data. Thus the empirical evidence suggests a hybrid model with both state and time dependent components. For this reason, we focus on a CalvoPlus structure for adjustment costs and proceed to analyze the dynamics of aggregate capital for this family.

4.5 Business cycle dynamics in CalvoPlus model

Now we assume the CalvoPlus model to analyze transitional dynamics of the first moment of capital gaps. We consider an unanticipated permanent aggregate productivity shock that shifts horizontally the distribution of idiosyncratic productivity of all firms. If this model is true, where both intensive and extensive margins are active, we obtain that

$$\mathrm{CIR}_1(\delta)/\delta \approx \underbrace{\frac{\mathbb{V}ar[x]}{\sigma^2}}_{3.014} - \underbrace{\frac{\nu\ \mathbb{C}ov[a,x]}{\sigma^2}}_{1.340} = 4.354. \tag{50}$$

The fact that the $\mathrm{CIR}_1(\delta)/\delta = 4.354$ is larger than its upper bound, that of a fully time dependent model given by $CIR_1(\delta)/\delta = \mathbb{E}[a] = 2.55$, suggests that their is no calibration such that the CalvoPlus generates the values of the ergodic moments in the data.

Calibration strategy. Now we show that the CalvoPlus model cannot generate jointly the four scalars that determine the CIR_1. The parameters for the stochastic process of the capital gaps $(\nu, \sigma^2) = (-0.111, 0.076)$ and the reset state $\hat{x} = 0.013$ are taken from the data through our formulas, see Table II. We are left with four parameters to calibrate: the upper and lower border of the inaction region $(\underline{x}, \overline{x})$, and the arrival rate of free adjustment opportunities λ. To set the remaining parameters, we follow the following strategy. First, we find the arrival rate of free adjustments λ to match the average expected time to adjustment $\mathbb{E}[\tau] = 2.44$. The implied λ is plotted in Figure IV. Then, we vary the normalized borders of inaction $(\underline{x} - \hat{x}, \overline{x} - \hat{x})$ and compute the implied capital gap variance $\mathbb{V}ar[x]$ and covariance with age $\mathbb{C}ov[x, a]$ for those parameters.

Figure IV – Calibration of arrival rate of free adjustment opportunities

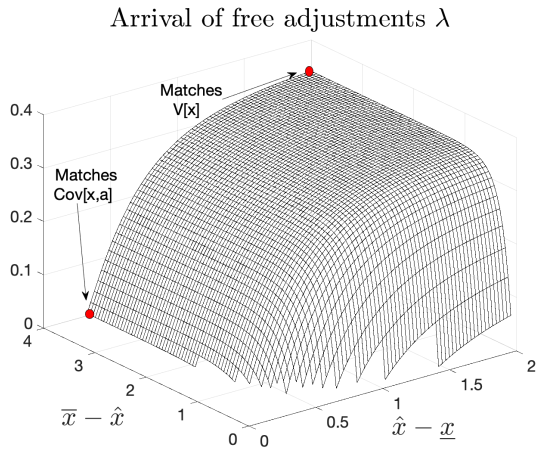
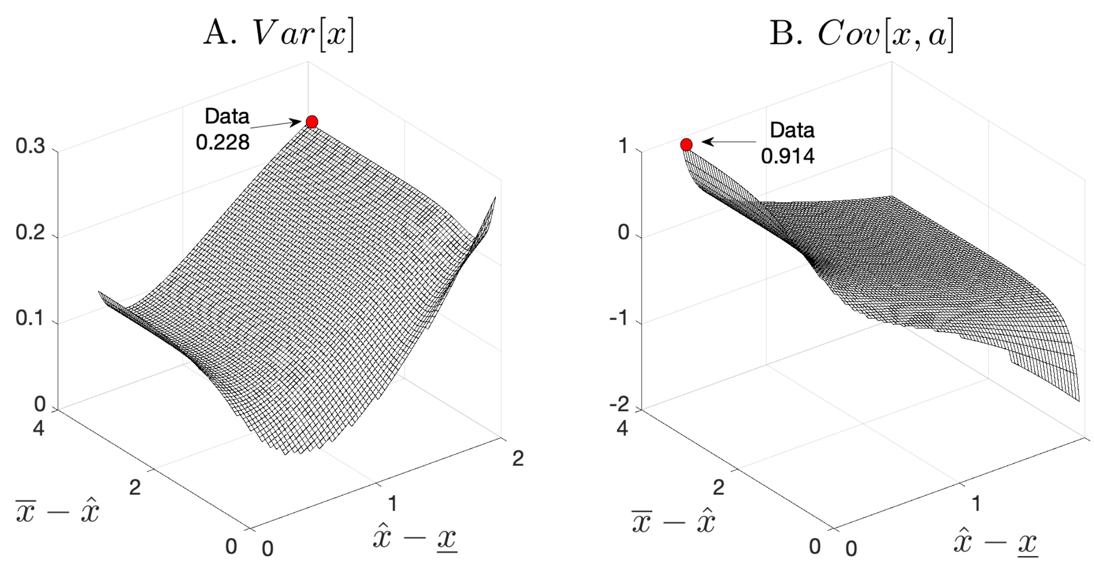


Figure V plots these key moments for different values of the normalized inaction regions (and the implied λ that matches average duration). The variance is increasing in both sides of the inaction region, while the covariance is decreasing in the lower band and increasing in the first. Given these shapes, we find that, in order to match the variance of capital gaps $\mathbb{V}ar[x]$, we need a wide upper band $\overline{x} - \hat{x} = 4$ and a wide lower band $\hat{x} - \underline{x} = 2$ (and a large $\lambda = 0.34$); in other words, the variance calls for wide inaction region with frequent free adjustment opportunities. In this case, the adjustments would be *primarily* driven by the Poisson arrival rate, as in a Calvo model. In contrast, to match the covariance of capital gap and its age, we still need a wide upper band $\overline{x} - \hat{x} = 4$ but a very narrow lower band $\hat{x} - \underline{x} = 0.3$ (as well as a low $\lambda = 0.005$); thus, the covariance calls for a one-sided inaction region and infrequent opportunities to adjust. This is close to a standard menu cost model with very asymmetric inaction regions.

Figure V – Key moments generated by the random fixed cost model



In summary, the random fixed cost model cannot match simultaneously the two key moments that it must satisfy for the CIR of average capital.

4.6 Heterogeneity concerns

One obvious concern that arises regarding the empirical analysis is that different layers of heterogeneity may affect the computation and the interpretation of cross-sectional statistics. To mitigate this concern, we repeat the analysis considering eight 2-digit subsectors within manufacturing and also different plant sizes in terms of their number of workers.[26]

Table III – Inputs from Micro Data and Outputs from the Theory by Size and Sector

Inputs from Data				**Outputs from Theory**			
	Plants size		**Excluding**		**Plants size**		**Excluding**
	Small	**Large**	**Textile/Chemical**		**Small**	**Large**	**Textile/Chemical**
Frequency				**Parameters**			
$\mathbb{E}^{x}[\tau]$	2.862	2.300	2.445	ν	-0.111	-0.111	-0.110
$\mathbb{CV}^2[\tau]$	1.167	1.012	1.096	σ^2	0.077	0.078	0.075
				$\hat{x}$	0.004	0.013	0.011
Capital Gaps				**Steady State Moments**			
$\mathbb{E}^{x}[\Delta x]$	0.317	0.256	0.270	$\mathbb{V}ar[x]$	0.220	0.244	0.228
$\mathbb{E}^{\hat{x}}[\Delta x^2]$	0.222	0.185	0.190	$\mathbb{E}[a]$	3.101	2.314	2.563
$\mathbb{E}^{\hat{x}}[(\hat{x}-\Delta x)^3]$	-0.209	-0.187	-0.184	$\mathbb{C}ov[a,x]$	0.849	1.016	0.922
Covariances				**Transitional Dynamics**			
$\mathbb{C}ov^{x}[\tau,\Delta x]$	0.088	0.033	0.059	$\mathbb{V}ar[x]/\sigma^2$	2.870	3.145	3.021
$\mathbb{E}^{\hat{x}}[\tau(\hat{x}-\Delta x)^2]$	0.769	0.455	0.531	$-\nu\mathbb{C}ov[a,x]/\sigma^2$	1.227	1.454	1.350
				$\text{CIR}_1(\delta)$	4.097	4.599	4.371

Sources: Authors' calculations using establishment-level survey data for Chile.

Regarding the sectoral composition, we find that besides textiles and chemicals, all other sectors present very similar investment patters. Table III recomputes our formulas excluding textiles and chemicals and we find that neither the inputs from the data nor the outputs from the theory differ from the numbers that include these sectors. We conclude that heterogeneity across sectors should not be a concern once these two subsectors are excluded. Regarding plant size, things are more interesting. Across all capital categories, average investment, the frequency of non-zero investments, and the fraction of spikes are increasing in plant size; in contrast, the inaction rate decreases with size. Table III recomputes our formulas splitting the sample in small and large plants. Not surprisingly, the inputs from the data are different for small and large firms. What is really striking is that the parameters of the stochastic process that we recovered are identical same across sizes: firms are hit with the same type of shocks. In terms

[26]See Data Online Appendix for details. Table XIV reports cross-sectional statistics by subsectors: (1) Food and beverages; (2) Textile, clothing and leather; (3) Wood and furniture; (4) Paper and printing; (5) Chemistry, petroleum, rubber and plastic; (6) Manufacture of non-metallic mineral products; (7) Basic metal; (8) Metal products, machinery and equipment. Table XVI reports cross-sectional statistics by quartiles of the average number of workers during the sample period: small plants (0-25%, S), medium plants (25-50%, M), large plants (50-75%, L), and very large plants (75-100%, XL).

of ergodic moments, larger plants have more dispersed capital gaps, capital is younger on average, and have a larger covariance; all of these forces increase the CIR.

Overall, heterogeneity across sectors and across plant size does not change the conclusion.

5 Extensions and Generalization

In the previous sections we specified parametric restrictions to the inaction model and to the firms' state space. Such assumptions exclude from our analysis models with fixed adjustment dates as in Taylor (1980), models with observation costs as in Álvarez, Lippi and Paciello (2011), and several others. Nevertheless, it is possible to extend our theory to accommodate richer models. In this section, we generalize our results to consider any stopping-time model or state space, explaining the assumptions on policies and processes that are key to apply our tools.

Second, we extend the analysis in three directions, to consider: (i) transitions of higher moments ($m > 1$) of the distribution; (ii) transitions starting from any general initial condition F_0; and (iii) transitions for a mean-reverting process. In each case, we focus on the one property that delivers the most interesting mechanism.[27] We denote conditional distributions as $Z|Y$, conditional expectations with initial condition z as $\mathbb{E}^z[Z]$, and the minimum between two stopping times as $t \wedge s \equiv \min\{t, s\}$.

5.1 Generalization

Let $(\Omega, P, \mathcal{F})$ be a probability space equipped with a filtration $\mathcal{F} = (\mathcal{F}_t; t \geq 0)$. We consider an economy populated by a continuum of agents indexed with $\omega \in \Omega$, where agent ω's information set at time t is the filtration $\mathcal{F}_t$. Each agent's uncontrolled state is given by $\tilde{S}_t(\omega) = [\tilde{x}_t(\omega), S_t^{-x}(\omega)] \in \mathbb{R}^{1+K_{-x}}$. The state is split between a *main state* $\tilde{x}$ and a set of *complementary states* $\tilde{S}_t^{-x}$. The main state follows a Brownian motion $d\tilde{x}_t(\omega) = \sigma dW_t(\omega)$. Agent's policies consist of a sequence of adjustment dates $\{\tau_k\}_{k=1}^{\infty}$ and adjustments sizes $\{\Delta S_{\tau_k}\}_{k=1}^{\infty}$, measurable with respect to $\mathcal{F}_t$. Given these policies $\{\tau_k(\omega), \Delta S_{\tau_k}(\omega)\}_{k=1}^{\infty}$, the controlled state $S_t(\omega)$ evolves as the sum of the uncontrolled state plus the adjustments: $S_t(\omega) = \tilde{S}_t(\omega) + \sum_{\tau_k(\omega) \leq t} \Delta S_{\tau_k(\omega)}$.

The first premise for our theory is a recursive representation of the conditional CIR, both between and within stopping dates. This demands $S_t(\omega)$ to be a sufficient statistic for the conditional CIR, which in turn requires that the policy is history independent. Formally, this mean that

$$\mathbb{E}\left[\int_{\tau_i \wedge t(\omega)}^{\tau_{i+1}} f(x_t) dt | \mathcal{F}_{\tau_i \wedge t(\omega)}\right] = \mathbb{E}\left[\int_0^{\tau} f(x_t) dt | S_{\tau_i \wedge t(\omega)}\right] = v^f(S_{\tau_i \wedge t(\omega)}), \quad \text{for all} \quad t(\omega) \leq \tau_{i+1}.$$

Since the main state follows a Brownian motion, the burden of this requirement falls completely on the complementary state and the policy. Assumption 1 and 2 formalize these requirements.

[27] The Web Appendix presents the full characterization and analysis of the three properties.

Assumption 1 (**Markovian complementary state**). *The complementary state $\tilde{S}_t^{-x}$ follows a Strong Markov process:*

$$\tilde{S}^{-x}_{(t\wedge\tau_k)+h}(\omega)|\mathcal{F}_{t\wedge\tau_k} \quad = \quad \tilde{S}^{-x}_h(\omega)|\tilde{S}_{(t\wedge\tau_k)}(\omega), \qquad \forall k. \tag{51}$$

To understand this assumption, consider a history ω such that $t < \tau_k(\omega)$. In this case, the complementary state's law of motion depends only on its current value; thus it is independent of its own history. Additionally, the complementary state is an homogenous process, since its law of motion at date t is equivalent to its law of motion at zero, given an initial condition. In the complementary case $t \geq \tau_k(\omega)$, these properties continue to hold, thus the stopping policy does not reveal new information about the complementary state's law of motion.

Assumption 2 (**Markovian policies**). *Policies satisfy the following conditions:*

$$\tau_{k+1}|\mathcal{F}_{\tau_k+h} \quad = \quad \tau_1|S_{\tau_k+h} \quad \text{for all} \quad h \in [0, \tau_{k+1}-\tau_k]. \tag{52}$$

A second premise in our theory is that we can characterize the CIR with the *first* stopping time of every agent. This means that, upon taking action, agents fully adjust to include any deviations from their steady-state behavior and come back to the steady-state process. This would imply that S_{τ_k} is *iid* across time and independent of the history previous to the adjustment. The challenge with stochastic *iid* resets is that is makes it more difficult to identify the parameters of the stochastic process, e.g. differentiating the fundamental volatility σ from the volatility arising from a random reset state. Therefore, in order for the reset state to be sufficiently informative, we ask that it is a constant $x_{\tau_i} \; = \; \hat{x}$.[28]

Assumption 3 (**Constant reset state**). *The reset state is constant: $x_{\tau_k} \; = \; \hat{x}$ for all k.*

It is straightforward to check that the previous assumptions hold in the investment example developed in Section 2. For Assumption 1, the complementary state is given by the arrival of free adjustment opportunities N_t, which is assumed to be a Poisson counter process and thus a Strong Markov process. The requirements in Assumption 2 and 3 are also satisfied. We showed that the reset capital gap is constant; and since the stopping policy is an inaction set with respect to the controlled state, the stopping policy is history independent within and between adjustments.

Finally, in order to apply the Optional Sampling Theorem, we require several stopping processes to be well-defined (finite moments at the stopping-time).[29]

Assumption 4 (**Well-defined stopping processes**). *The processes $\left(\left\{\int_0^t s^j x_s^m \; dB_s\right\}_t, \tau\right)$ for all m and $j \; = \; 0, 1$, are well-defined stopping processes.*

The previous Markovian requirements are enough in order to characterize the aggregation, representation of the intensive margin, and observation properties; however, in order to apply the representation property to the extensive margin, we must require one additional assumption. There must exist an equivalent representation of the extensive margin as a function exclusively of the main state x. For this,

[28] In this paper, we ignore ex-ante heterogeneity across agents (that could be reflected in different reset states and policies), but this can be relaxed. Nevertheless, it remains crucial that history is erased at the moment of reseting the state.

[29] See Web Appendix A for a formal definition of a well-defined stopping process.

we require that there exists a stopping policy τ^* that only depends on the main state x and can fully describe the extensive margin by itself. For instance, a stopping policy given by a Poisson counter with hazard $\Lambda(x)dt$ satisfies this requirement.

Assumption 5 (**Hazard**). *Assume that there exist a stopping policy τ^* s.t.*

$$\mathbb{E}\left[\int_0^{\tau}\left(\frac{\partial \mathbb{E}^S\left[x_\tau^{m+2}/m+2\right]}{\partial x}-\mathbb{E}^S\left[x_\tau^{m+1}\right]\right)dt\right] \;=\; \mathbb{E}^{\hat{x}}\left[\int_0^{\tau}\left(\frac{\partial \mathbb{E}^x\left[x_{\tau^*}^{m+2}/m+2\right]}{\partial x}-\mathbb{E}^x\left[x_{\tau^*}^{m+1}\right]\right)dt\right], \tag{53}$$

and there exist a smooth function $g_m(x)$ such that

$$g_m(x) \;=\; \mathbb{E}^{\hat{x}+x}\left[(\hat{x}-\Delta x)^m\right]-\mathbb{E}\left[(\hat{x}-\Delta x+x)^m\right], \quad \forall m. \tag{54}$$

where Δx is under the policy τ^*.

5.2 Extensions

Now that we have stated the formal requirements needed to apply our theory, we proceed to develop the three extensions. To highlight the new mechanisms, in all the extensions we focus on the driftless case $\nu \;=\; 0$, but the proofs are straightforward to extend to consider a non-zero drift.

Extensions I: CIR for arbitrary functions the capital-gaps. This section provides formulas for the CIR of the m-th moment for different classes of models. Additionally, these formulas have two useful applications. They can be used to (i) derive bounds for the dynamics of functions of the m-th moments, and (ii) study transitions of any arbitrary function of the state. To illustrate the first application, let us consider the transitional dynamics for the variance. Using Jensen's inequality, we derive an upper bound on the variance's CIR:[30]

$$\mathrm{CIR}(\mathbb{V}ar[x]) \;\equiv\; \int_0^\infty \left(\mathbb{V}_t\left[x\right]-\mathbb{V}\left[x\right]\right)dt \;\leq\; \mathrm{CIR}_2(\delta)-\mathrm{CIR}_1(\delta)^2. \tag{55}$$

To illustrate the second application, consider a smooth function of the state $f(x)$. For example, in many models the aggregate welfare criteria can be written in this form. Using a Taylor approximation around zero, we write the CIR of the $f(x)$ function in terms of the state's CIR, weighted by the Taylor factors.

$$\mathrm{CIR}(f(x)) \;=\; \int_0^\infty \mathbb{E}_t[f(x)]-\mathbb{E}[f(x)]dt \;=\; \sum_{j\,=\,1}^{\infty}\frac{df^j(0)}{dx^j}\frac{\mathrm{CIR}_j(\delta)}{j!}. \tag{56}$$

Extension II: Transitional dynamics for higher moments We first consider the transitional dynamics for higher moments of the distribution ($m \geq 1$). The initial condition remains to be a mean translation of the steady-state distribution. In this case, we focus the discussion on the representation of the intensive margin.

[30] $\mathrm{CIR}(\mathbb{V}ar[x]) \;\equiv\; \int_0^\infty \left(\mathbb{V}_t\left[x\right]-\mathbb{V}\left[x\right]\right)dt \;=\; \int_0^\infty \left(\mathbb{E}_t\left[x^2\right]-\mathbb{E}\left[x^2\right]\right)dt-\int_0^\infty \mathbb{E}_t\left[x\right]^2 dt \leq \mathrm{CIR}_2(\delta)-\mathrm{CIR}_1^2(\delta)$.

Proposition 5. *Assume $d\tilde{x}_t = \sigma dW_t$. To a first order, the transitional dynamics of the m-th moment are given by*

$$CIR_m(\delta)/\delta \;=\; \Gamma_m + \Theta_m - \mathbb{E}[x^m]\Theta_0 \;+\; o(\delta) \tag{57}$$

where the intensive margin relates to ergodic moments as follows:

$$\Gamma_m \;=\; m\mathbb{E}[x^{m-1}, a], \tag{58}$$

$$\mathbb{E}[x^{m-1}, a] \;=\; \frac{2}{m(m+1)}\left[\frac{\mathbb{E}\left[\tau\left(\hat{x} - \Delta x\right)^{m+1}\right]}{\mathbb{E}\left[\Delta x^2\right]} - \frac{\mathbb{E}[x^{m+1}]}{\sigma^2}\right]. \tag{59}$$

To focus on the intensive margin, assume $\Theta_m = 0$ for all m and consider the transitional dynamics for the state's first three moments by setting $m = 1, 2, 3$. We have that

$$\mathrm{CIR}_1(\delta)/\delta \;=\; \Gamma_1 \;=\; \mathbb{E}[a] \tag{60}$$

$$\mathrm{CIR}_2(\delta)/\delta \;=\; \Gamma_2 \;=\; 2\mathbb{E}[xa] \tag{61}$$

$$\mathrm{CIR}_3(\delta)/\delta \;=\; \Gamma_3 \;=\; 3\mathbb{E}[x^2a] \tag{62}$$

As discussed earlier, the dynamics of the first moment ($m = 1$)—average capital gaps—are fully driven by the state's average age. The dynamics of the second moment ($m = 2$)—dispersion of capital gaps or misallocation—are driven by the covariance between the age and the size of capital gaps. If this covariance is zero, then the distribution's second moment remains constant along the transition path. Asymmetry in in the agents' investment policy, which generates a skewed ergodic distribution, is one way to generate a non-zero covariance. This interaction between the business cycle dynamics of capital misallocation and the asymmetry of the ergodic capital distribution is studied by Ehouarne, Kuehn and Schreindorfer (2016) and Jo and Senga (2014). Finally, the dynamics of the third moment ($m = 3$)—skewness of capital gaps—are driven by the covariance between age and the square of capital gaps. Note that if the ergodic distribution features excess kurtosis, then the skewness of the distribution will change along the transition.

Proposition 5 provides formulas for the CIR of the m-th moment. Additionally, these formulas have two useful applications. They can be used to (i) derive bounds for the dynamics of functions of the m-th moments, and (ii) study transitions of any arbitrary function of the state. To illustrate the first application, let us consider the transitional dynamics for the variance. Using Jensen's inequality, we derive an upper bound on the variance's CIR:[31]

$$\mathrm{CIR}(\mathbb{V}ar[x]) \;\equiv\; \int_0^\infty \left(\mathbb{V}_t\left[x\right] - \mathbb{V}\left[x\right]\right) dt \;\leq\; \mathrm{CIR}_2(\delta) - \mathrm{CIR}_1^2(\delta). \tag{63}$$

To illustrate the second application, consider a smooth function of the state $f(x)$. For example, in many models the aggregate welfare criteria can be written in this form. Using a Taylor approximation around zero, we write the CIR of the $f(x)$ function in terms of the state's CIR, weighted by the Taylor

[31] $\mathrm{CIR}(\mathbb{V}ar[x]) \;\equiv\; \int_0^\infty \left(\mathbb{V}_t\left[x\right] - \mathbb{V}\left[x\right]\right) dt \;=\; \int_0^\infty \left(\mathbb{E}_t\left[x^2\right] - \mathbb{E}\left[x^2\right]\right) dt - \int_0^\infty \mathbb{E}_t\left[x\right]^2 dt \leq \mathrm{CIR}_2(\delta) - \mathrm{CIR}_1^2(\delta)$.

factors.

$$\text{CIR}(f(x)) \;=\; \int_0^\infty \mathbb{E}_t[f(x)] - \mathbb{E}[f(x)]dt \;=\; \sum_{j\,=\,1}^{\infty} \frac{df^j(0)}{dx^j}\frac{\text{CIR}_j(\delta)}{j!}. \tag{64}$$

Extension III: General initial conditions This extension considers transitional dynamics for general initial conditions. For instance, since the work on uncertainty shocks by Bloom (2009), there has been a large literature interested in the macroeconomic consequences of uncertainty in the business cycle. Within our framework, these aggregate uncertainty shocks can be studied by setting the initial distribution as a mean-preserving spread of the steady-state distribution. Moreover, the interaction between first and second moment shocks, as studied by Aastveit, Natvik and Sola (2013), Vavra (2014), Caggiano, Castelnuovo and Nodari (2014), Castelnuovo and Pellegrino (2018), and Baley and Blanco (2019), can be accommodated as well.

For simplicity, we consider perturbations that can be expressed via a single parameter δ. The initial distribution is described through a function $\mathcal{G}(x,\delta)$, such that $F_0(x) = F(\mathcal{G}^{-1}(x,\delta))$. To make progress, we impose certain smoothness and differentiability properties to the function $\mathcal{G}$.Additionally, we focus on perturbations to the first and second moments. Since this extension does not affect steady-state moments, we omit the characterization of the observation property as it remains as before.

Proposition 6. *Assume $d\tilde{x}_t = \sigma dW_t$ and let $\mathcal{G}(x,\delta)$ be a function that satisfies the following properties:*

1. $\mathcal{G}(x,0) = x$.
2. $\exists z > 0$ *such that* $\forall \epsilon \in (-z,z)$, $\mathcal{G}(\cdot,\epsilon)$ *is bijective.*
3. $\frac{\partial \mathcal{G}(\mathcal{G}^{-1}(y,0),0)}{\partial \delta} = -(\mathcal{G}_0 + \mathcal{G}_1 y)$ *with* $\mathcal{G}_0^2 + \mathcal{G}_1^2 = 1$.

To a first order, the CIR is given by:

$$\begin{aligned}
CIR_1(\mathcal{G})/\delta &= \underbrace{\mathcal{G}_0\left(\Gamma_{1,0} + \Theta_{1,0}\right)}_{\text{1st moment shock}} + \underbrace{\mathcal{G}_1\left(\Gamma_{1,1} + \Theta_{1,1}\right)}_{\text{2nd moment shock}} + o(\delta) &(65)\\
\Gamma_{1,i} &= (i+1)\mathbb{E}[x^i a] &(66)\\
\Theta_{1,i} &= \sum_{j\,=\,0}^{\infty} \theta_{1,j}\mathbb{E}[x^{j+i}] &(67)
\end{aligned}$$

with $\theta_{1,j}$ are the micro-elasticities.

Proposition 6 points towards the moments that are crucial to characterize the dynamics for a particular type of initial condition. As long as there exists enough differentiability in the perturbation of the initial condition, we can find ergodic moments that perfectly describe the dynamics of the model. Interestingly, the micro-elasticities needed to compute the extensive margin are independent of the number of moments that are shocked.

As an example, consider $\mathcal{G}$ to be a mean preserving spread of the steady-state distribution F_0. This means that $\mathcal{G}(x,\delta) = x(1+\delta)$ and therefore $\mathcal{G}_0 = 0$ and $\mathcal{G}_1 = 1$. Again, let us focus only in the intensive margin by setting $\Theta_{1,i} = 0$ for all i. Then the CIR is approximated as:

$$\frac{\text{CIR}_1(\delta)}{\delta} \approx \mathcal{G}_1\Gamma_{1,1} = \mathbb{C}ov[x,a].$$

Thus mean-preserving perturbations have first order effects if and only if the covariance between age and the state is different from zero. A non-zero covariance is consistent with the data presented in Section 4. Therefore, suggesting that uncertainty shocks (in the form of mean-preserving spreads of the capital gap distribution) would have effects on average investment.

Extension IV: Mean-reversion. This extension considers a mean-reverting process for the uncontrolled state. This type of process is wildly used due to its empirical relevance and because it ensures the existence of an ergodic distribution. For this application, we focus on the observation properties.

Proposition 7. *Assume the uncontrolled state follows a Ornstein–Uhlenbeck process $d\tilde{x}_t = \rho\tilde{x}_t dt + \sigma dW_t$. Then, the reset state and structural parameters are recovered through a system of equations:*

$$\hat{x} = \frac{\mathbb{E}[e^{-\rho\tau}\Delta x]}{\mathbb{E}[e^{-\rho\tau}] - 1} \tag{68}$$

$$\frac{\sigma^2}{\rho} = 2\frac{\hat{x}^2 - \mathbb{E}\left[e^{-2\rho\tau}(\hat{x}-\Delta x)^2\right]}{\mathbb{E}\left[e^{-2\rho\tau}\right] - 1} \tag{69}$$

$$erf\left(\frac{\hat{x}}{\sqrt{\sigma^2/\rho}}\right) = \mathbb{E}\left[erf\left(\frac{\hat{x}-\Delta x}{\sqrt{\sigma^2/\rho}}\right)\right] \tag{70}$$

where $erf(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ is the Gauss error function.

To gain some intuitions about the observation formulas above, let us consider the limiting case $\rho \to 0$. Using the approximation $e^{-\rho\tau} \approx 1 - \rho\tau$, it is easy to show that equations (68) and (69) converge to our baseline observations expressions in (17) and (18) with $\nu = 0$ (no mean-reversion):

$$\hat{x} \to_{\rho\to 0} \frac{\mathbb{E}[\tau\Delta x]}{\mathbb{E}[\tau]}, \qquad \sigma^2 \to_{\rho\to 0} \frac{\mathbb{E}\left[\Delta x^2\right]}{\mathbb{E}[\tau]}. \tag{71}$$

Therefore, as long mean reversion is "sufficiently small", the mappings between the data and the reset state, and between the data and idiosyncratic volatility do not change.

Let us make a deeper comparison of how $\hat{x}$ is determined with and without mean-reversion. With *iid* shocks, we can write (17) as a weighted sum of investment rates across firms:

$$\hat{x}^{iid} = \mathbb{E}[\eta(\tau)\Delta x], \quad \textit{with} \quad \eta(\tau) \equiv \frac{\tau}{\mathbb{E}[\tau]} > 0, \quad \mathbb{E}[\eta(\tau)] = 1,$$

where the weights $\eta(\tau)$ are increasing in τ, i.e. more weight is given to the investment rate of firms with large periods of inaction (with "old" capital). In order to understand this result, note that conditional of surviving, the distribution of the state is more centered around the reset state for "young" capital vintages, which cannot reflect policy asymmetries. The opposite happens for firms with "old" vintages, as the distribution of the state is more centered around *the domain's middle point*, reflecting the policy asymmetries. Thus investment rates associated with large stopping times are more informative about these asymmetries. The opposite happens when we consider a mean-reverting process. An analogous decomposition yields

$$\hat{x}^{mr} = \mathcal{R}\mathbb{E}[\eta'(\tau)\Delta x], \quad \textit{with} \quad \eta'(\tau) \equiv \frac{e^{-\rho\tau}}{\mathbb{E}[e^{-\rho\tau}]} > 0, \quad \mathbb{E}[\eta'(\tau)] = 1, \quad \mathcal{R} \equiv \frac{\mathbb{E}[e^{-\rho\tau}]}{\mathbb{E}[e^{-\rho\tau}] - 1} < 0,$$

where now the weights are decreasing in duration and it is preceded by a negative number. As the inaction period of increases, the mean-reverting productivity process goes back to its zero long-run mean, and the distribution gets centered *around zero* on its own, so there is no need to correct for policy asymmetries with the initial condition.

6 Conclusion

This paper provides a structural relation in model of inaction between the CIR (a measure of persistence for aggregate dynamics) and microdata. This relation holds for any moment of the distribution, any inaction model, and any initial condition. In the same way we apply our tools to a model of lumpy investment, we foresee applications in models with labor adjustment costs, inventory models, portfolio management, government debt management, among others.

For developing our theory, we assume that upon taking action, agents fully adjust to include any deviations from their steady-state behavior. Thus our results do not accommodate partial adjustments which are due, for instance, to imperfect information or convex adjustment costs. One example of these frameworks is the menu cost model with information frictions in Baley and Blanco (2019). We leave for future research the application of the tools developed here to that type of frameworks.

References

Aastveit, K. A., Natvik, G. J. and Sola, S. (2013). Economic uncertainty and the effectiveness of monetary policy. *Working Paper*.

Alvarez, F., Beraja, M., Gonzalez-Rozada, M. and Neumeyer, P. A. (2018). From hyperinflation to stable prices: Argentina?s evidence on menu cost models. *The Quarterly Journal of Economics*, **134** (1), 451–505.

—, Le Bihan, H. and Lippi, F. (2016). The real effects of monetary shocks in sticky price models: a sufficient statistic approach. *American Economic Review*, **106** (10), 2817–51.

Álvarez, F. and Lippi, F. (2014). Price setting with menu cost for multiproduct firms. *Econometrica*, **82** (1), 89–135.

—, Lippi, F. and Paciello, L. (2011). Optimal price setting with observation and menu costs. *Quarterly Journal of Economics*.

—, — and Paciello, L. (2016). Monetary shocks in models with inattentive producers. *The Review of Economic Studies*, **83** (2), 421–459.

Alvarez, F. E., Lippi, F. and Paciello, L. (2015). Monetary shocks in models with inattentive producers. *The Review of economic studies*, **83** (2), 421–459.

Bachmann, R., Caballero, R. J. and Engel, E. M. (2013). Aggregate implications of lumpy investment: new evidence and a dsge model. *American Economic Journal: Macroeconomics*, **5** (4), 29–67.

Baley, I. and Blanco, A. (2019). Firm uncertainty cycles and the propagation of nominal shocks. *AEJ: Macro*, **11** (1), 276–337.

Berger, D. and Vavra, J. (2015). Consumption dynamics during recessions. *Econometrica*, **83** (1), 101–154.

Blanco, A. (2018). Optimal inflation target in an economy with menu costs and an occasionally binding zero lower bound. *Working Paper*.

— and Cravino, J. (2018). Price rigididity and the relative ppp. *NBER Working Paper*.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, **77** (3), 623–685.

—, Guvenen, F. and Salgado, S. (2016). Skewed business cycles. In *2016 Meeting Papers*, 1621, Society for Economic Dynamics.

Caballero, R. J. and Engel, E. M. R. A. (1991). Dynamic (s, s) economies. *Econometrica*, **59** (6), 1659–1686.

— and — (1993). Microeconomic adjustment hazards and aggregate dynamics. *The Quarterly Journal of Economics*, **108** (2), 359–383.

—, — and Haltiwanger, J. (1997). Aggregate employment dynamics: Building from microeconomic evidence. *The American Economic Review*, **87** (1), 115–137.

Caggiano, G., Castelnuovo, E. and Nodari, G. (2014). *Uncertainty and Monetary Policy in Good and Bad Times*. Tech. rep., Dipartimento di Scienze Economiche" Marco Fanno".

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

Caplin, A. and Leahy, J. (1991). State-Dependent Pricing and the Dynamics of Money and Output. *The Quarterly Journal of Economics*, **106** (3), 683–708.

Caplin, A. S. and Spulber, D. F. (1987). Menu costs and the neutrality of money. *The Quarterly Journal of Economics*, **113**, 287–303.

Carvalho, C. and Nechio, F. (2011). Aggregation and the ppp puzzle in a sticky-price model. *American Economic Review*, **101** (6), 2391–2424.

— and Schwartzman, F. (2015). Selection and monetary non-neutrality in time-dependent pricing models. *Journal of Monetary Economics*, **76**, 141 – 156.

Castelnuovo, E. and Pellegrino, G. (2018). Uncertainty-dependent effects of monetary policy shocks: A new keynesian interpretation. *Journal of Economic Dynamics and Control*, **Forthcoming**.

Cooper, R. W. and Haltiwanger, J. C. (2006). On the nature of capital adjustment costs. *The Review of Economic Studies*, **73** (3), 611–633.

Ehouarne, C., Kuehn, L.-A. and Schreindorfer, D. (2016). *Misallocation cycles*. Tech. rep., Working Paper.

Eslava, M., Haltiwanger, J., Kugler, A. and Kugler, M. (2004). The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from colombia. *Journal of Development Economics*, **75** (2), 333 – 371, 15th Inter American Seminar on Economics.

—, —, — and — (2013). Trade and market selection: Evidence from manufacturing plants in colombia. *Review of Economic Dynamics*, **16** (1), 135 – 158, special issue: Misallocation and Productivity.

Feenstra, R. C., Inklaar, R. and Timmer, M. P. (2015). The next generation of the penn world table. *American Economic Review*, **105** (10), 3150–82.

Gagnon, E. (2009). Pricing setting during low and high inflation: Evidence from mexico. *Quarterly Journal of Economics*, p. 126.

Golosov, M. and Lucas, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, pp. 171–199.

Hamermesh, D. S. (1989). Labor demand and the structure of adjustment costs. *The American Economic Review*, **79** (4), 674–689.

Henríquez, C. G. (2008). *Stock de Capital en Chile (1985-2005): Metodologa y Resultados*. Economic Statistics Series 63, Central Bank of Chile.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, **124** (4), 1403–1448.

Jo, I. H. and Senga, T. (2014). Size distribution and firm dynamics in an economy with credit shocks.

Kehoe, P. and Midrigan, V. (2008). Sticky prices and real exchange rates in the cross-section, mimeo, University of Minnesota and New York University.

Khan, A. and Thomas, J. K. (2008). Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, **76** (2), 395–436.

Kozlowski, J., Veldkamp, L. and Venkateswaran, V. (2015). The tail that wags the economy: Belief-driven business cycles and persistent stagnation.

Liu, L. (1993). Entry-exit, learning, and productivity change evidence from chile. *Journal of Development Economics*, **42** (2), 217 – 242.

Mankiw, N. G. (2014). *Principles of macroeconomics*. Cengage Learning.

Midrigan, V. (2011). Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica*, **79** (4), 1139–1180.

Nakamura, E. and Steinsson, J. (2010). Monetary non-neutrality in a multisector menu cost model. *The Quarterly Journal of Economics*, **125** (3), 961–1013.

Oberfield, E. (2013). Productivity and misallocation during a crisis: Evidence from the chilean crisis of 1982. *Review of Economic Dynamics*, **16** (1), 100–119.

Oksendal, B. (2007). *Stochastic Differential Equations*. Springer, 6th edn.

Pombo, C. (1999). Productividad industrial en colombia: una aplicación de números índices. *Revista de economía del Rosario*, **2** (1), 107–139.

Restuccia, D. and Rogerson, R. (2013). Misallocation and productivity. *Review of Economic Dynamics*, **1** (16), 1–10.

STOKEY, N. (2009). *The Economics of Inaction*. Princeton University Press.

TAYLOR, J. B. (1980). Aggregate dynamics and staggered contracts. *The Journal of Political Economy*, pp. 1–23.

TYBOUT, J. R. (2000). Manufacturing firms in developing countries: How well do they do, and why? *Journal of Economic Literature*, **38** (1), 11–44.

VAVRA, J. (2014). Inflation dynamics and time-varying volatility: New evidence and an ss interpretation. *The Quarterly Journal of Economics*, **129** (1), 215–258.

WINBERRY, T. (2016). Lumpy investment, business cycles, and stimulus policy.

WOODFORD, M. (2009). Information-constrained state-dependent pricing. *Journal of Monetary Economics*, **56**, S100–S124.

ZWICK, E. and MAHON, J. (2017). Tax policy and heterogeneous investment behavior. *American Economic Review*, **107** (1), 217–48.

A Appendix: Proofs

Proof of Lemma 1. Here we extend the result in Alvarez, Le Bihan and Lippi (2016) for higher order moments and arbitrary state state. Let $S_t = [x_t, a_t]$. Next, we enumerate the three assumptions needed for this Lemma:

1. $\{S_t\}_{t\in[0,\infty)}$ is an Strong Markov Process with first element x_t.
2. $S_{\tau_i} = \hat{S}$.
3. τ_i is a stopping time w.r.t. the filtration generated by $\{S_t\}$.

Fix an $m \in \mathbb{N}$. Start from the CIR's definition:

$$\text{CIR}_m = \mathbb{E}\left[\int_0^\infty (x_t(\omega)^m - \mathbb{E}[x^m])\, dt\right], \tag{A.1}$$

where the expectation is taken across agents ω. Using the strong Markov property and law of iterated expectations

$$\text{CIR}_m = \mathbb{E}\left[\mathbb{E}\left[\int_0^\infty (x_t(\omega)^m - \mathbb{E}[x^m])\, dt\middle| \mathcal{F}_0\right]\right] = \int_S \mathbb{E}\left[\int_0^\infty (x_t(\omega)^m - \mathbb{E}[x^m])\, dt\middle| S_0 = S\right] dF_0(S) =: \text{CIR}_m(F_0), \tag{A.2}$$

Let $\{\tau_i\}_{i=1}^\infty$ be the sequence of stopping times. In (1), we write the CIR as the cumulative deviations between time $t = 0$ and the first stopping time τ_1 plus the sum of deviations between all future stopping times. In (2), we use the Law of Iterated Expectations to condition on the information set $\mathcal{F}_{\tau_i}$. In (3), we use the Strong Markov Property of S_t, the assumption of homogenous resets and that $\hat{S}$ is constant for $i \geq 1$ to change the conditioning from $S_{\tau_i+h}|\mathcal{F}_{\tau_i}$ to $S_h|\hat{S}$ and write the problem recursively. In (4), we show that *every element* inside the infinite sum is equal to zero. For this purpose, recall the relationship between ergodic moments and expected duration derived in Auxiliary Theorem 2, $\mathcal{M}_m[x] = \mathbb{E}^{\hat{S}}\left[\int_0^\tau x_t(\delta|\omega)^m\right] / \mathbb{E}^{\hat{S}}[\tau]$, and thus we are left with the simple expression in the fourth line (we also relabel τ_1 as τ):

$$\begin{aligned}
\text{CIR}_m(F_0) &= \int_S \mathbb{E}\left[\int_0^\infty (x_t(\omega)^m - \mathbb{E}[x^m])\, dt\middle| S_0 = S\right] dF_0(S), \\
&=^{(1)} \int_S \mathbb{E}\left[\int_0^{\tau_1} (x_t(\omega)^m - \mathbb{E}[x^m])\, dt + \sum_{i=1}^\infty \int_{\tau_i}^{\tau_{i+1}} (x_t(\omega)^m - \mathbb{E}[x^m])\, dt\middle| S_0 = S\right] dF_0(S) \\
&=^{(2)} \int_S \mathbb{E}\left[\int_0^{\tau_1} (x_t(\omega)^m - \mathbb{E}[x^m])\, dt + \sum_{i=1}^\infty \mathbb{E}\left[\int_{\tau_i}^{\tau_{i+1}} (x_t(\omega)^m - \mathbb{E}[x^m])\, dt\middle|\mathcal{F}_{\tau_i}\right]\middle| S_0 = S\right] dF_0(S) &\text{(A.3)} \\
&=^{(3)} \int_S \mathbb{E}\left[\left[\int_0^{\tau_1} (x_t(\omega)^m - \mathbb{E}[x^m])\, dt\right] + \mathbb{E}\left[\sum_{i=1}^\infty \underbrace{\mathbb{E}\left[\int_0^\tau x_t(\omega)^m\, dt\middle|\hat{S}\right] - \mathcal{M}_m[x]\mathbb{E}[\tau|\hat{S}]}_{=0}\right| S_0 = S\right] dF_0(S) &\text{(A.4)} \\
&=^{(4)} \int_S \left[\mathbb{E}\left[\int_0^\tau (x_t(\omega)^m)\, dt|S\right] - \mathbb{E}[x^m]\mathbb{E}[\tau|S]\middle| S_0 = S\right] dF_0(S) &\text{(A.5)}
\end{aligned}$$

As a final step, define the following value function conditional on a particular initial condition $S = [x, S^{-x}]$:

$$v^m(S) \equiv \mathbb{E}^S\left[\int_0^\tau [x_t(\omega)^m - \mathbb{E}[x^m]]\, dt\right], \tag{A.6}$$

$$\text{CIR}_m(F_0) = \int_S v^m(S) dF_0(S) \tag{A.7}$$

□

Proof of Proposition 1. We continue with the notation $S_t = [x_t, a_t]$ and assume $S_{\tau_i} = \hat{S}$. Additionally, we assume that $\left(\left\{\int_0^t x_s^m s^n\, dW_s\right\}_t, \tau\right)$ are a well-defined stopping processes for any m and $n = 0, 1$.

- **Average adjustment size.** From the law of motion $x_t = \hat{x} + \nu t + \sigma W_t$, we find the following equalities: $\sigma W_\tau = -\nu\tau + x_\tau - \hat{x} = -\nu\tau - \Delta x$. Taking expectations on both sides, we have $\sigma\mathbb{E}^{\hat{S}}[W_\tau] = -\nu\mathbb{E}^{\hat{S}}[\tau] - \mathbb{E}^{\hat{S}}[\Delta x]$. Since W_τ is a martingale, $\mathbb{E}^{\hat{S}}[W_\tau] = W_0 = 0$ by the OST. Therefore, $\nu = -\frac{\mathbb{E}^{\hat{S}}[\Delta x]}{\mathbb{E}^{\hat{S}}[\tau]}$ as well.
- **Observation of fundamental volatility:** For characterizing σ define $Y_t = x_t - \nu t$ with initial condition $Y_0 = \hat{x}$. With similar steps as before we have that

$$\sigma^2 = \frac{\mathbb{E}^{\hat{S}}\left[\Delta Y_\tau^2\right]}{\mathbb{E}^{\hat{S}}[\tau]} = \frac{\mathbb{E}^{\hat{S}}\left[(x_\tau - \nu\tau - x_\tau + \hat{x})^2\right]}{\mathbb{E}^{\hat{S}}[\tau]} = \frac{\mathbb{E}^{\hat{S}}\left[(\nu\tau + \Delta x)^2\right]}{\mathbb{E}^{\hat{S}}[\tau]} \tag{A.8}$$

or equivalently

$$\sigma^2 = \frac{\mathbb{E}^{\hat{S}}[\Delta x^2]}{\mathbb{E}^{\hat{S}}[\tau]} + 2\nu \left(\frac{\mathbb{E}^{\hat{S}}[\Delta x \tau]}{\mathbb{E}^{\hat{S}}[\tau]} + \nu \frac{\mathbb{E}^{\hat{S}}[\tau^2]}{\mathbb{E}^{\hat{x}}[\tau]} \right)$$

Applying the formula for $\hat{x}$ below (A.10) we have the result.

- **Observation of reset state:** For the reset state $\hat{x}$, we apply Itō's lemma to x_t^2 to obtain $d(x_t^2) = 2x_t dx_t + (dx_t)^2 = \left(2\nu x_t + \sigma^2\right) dt + 2\sigma x_t dW_t$. Using the OST $\mathbb{E}^{\hat{S}}[\int_0^\tau x_s dW_s] = 0$. Moreover, given that $\mathbb{E}^{\hat{S}}[\int_0^\tau x_s ds] = \mathbb{E}[x]\mathbb{E}^{\hat{S}}[\tau] = 0$, we have that

$$\mathbb{E}^{\hat{S}}[x_\tau^2] = \hat{x}^2 + \sigma^2 \mathbb{E}^{\hat{S}}[\tau] \tag{A.9}$$

Completing squares $\mathbb{E}^{\hat{S}}[x_\tau^2] = \mathbb{E}^{\hat{S}}[(\hat{x} - (\hat{x} - x_\tau))^2] = \mathbb{E}^{\hat{S}}[\Delta x^2] - 2\hat{x}\mathbb{E}^{\hat{S}}[\Delta x] + (\hat{x})^2$, we get

$$\begin{aligned} \hat{x} &= \frac{1}{2\mathbb{E}^{\hat{S}}[\Delta x]} \left[\mathbb{E}^{\hat{S}}[\Delta x^2] - \sigma^2 \mathbb{E}^{\hat{S}}[\tau] \right] \\ &= \frac{1}{2\mathbb{E}^{\hat{S}}[\Delta x]} \left[\mathbb{E}^{\hat{S}}[\Delta x^2] - \left(\mathbb{E}^{\hat{S}}[\Delta x^2] + 2\frac{\mathbb{E}^{\hat{S}}[\Delta x]\mathbb{E}^{\hat{x}}[\Delta x \tau]}{\mathbb{E}^{\hat{S}}[\tau]} + \frac{\mathbb{E}^{\hat{S}}[\Delta x]^2 \mathbb{E}^{\hat{S}}[\tau^2]}{\mathbb{E}^{\hat{S}}[\tau]^2} \right) \right] \\ &= \frac{\mathbb{E}^{\hat{S}}[\Delta x \tau]}{\mathbb{E}^{\hat{S}}[\tau]} - \frac{\mathbb{E}^{\hat{S}}[\Delta x]\mathbb{E}^{\hat{S}}[\tau^2]}{2\mathbb{E}^{\hat{S}}[\tau]^2}. \end{aligned} \tag{A.10}$$

Applying the formula for the covariance $\mathbb{E}^{\hat{S}}[\tau \Delta x] + \mathbb{E}^{\hat{S}}[\tau]\mathbb{E}^{\hat{x}}[\Delta x] = Cov^{\hat{S}}[\tau, \Delta x]$ and coefficient of variation square $\mathbb{CV}^2[\tau] = \frac{\mathbb{V}^{\hat{S}}[\tau]}{\mathbb{E}^{\hat{S}}[\tau]^2}$, we have the result.

- **Observation of ergodic moments with respect to the state:** For observability of ergodic moments of x, apply Itō's lemma to x^{m+1} and get $dx_t^{m+1} = (m+1)x_t^m \nu dt + (m+1)x_t^m \sigma dW_t + \frac{\sigma^2}{2} m(m+1)x_t^{m-1} dt$. Integrating from 0 to τ, using the OST to eliminate martingales, and rearranging:

$$\mathbb{E}^{\hat{S}} \left[\int_0^\tau x_t^m dt \right] = \frac{1}{\nu(m+1)} \left(\mathbb{E}^{\hat{S}}[x_\tau^{m+1}] - \hat{x}^{m+1} \right) - \frac{\sigma^2}{2\nu} m \mathbb{E}^{\hat{S}} \left[\int_0^\tau x_t^{m-1} dt \right] \tag{A.11}$$

Substituting the equivalences $\mathbb{E}[x^m] = \mathbb{E}^{\hat{S}} \left[\int_0^\tau x_t^m dt \right] / \mathbb{E}^{\hat{S}}[\tau]$ and $\mathbb{E}^{\hat{S}}[\Delta x] = -\nu \mathbb{E}^{\hat{S}}[\tau]$ yields:

$$\mathbb{E}[x^m] = \frac{\hat{x}^{m+1} - \mathbb{E}^{\hat{x}}[(\hat{x} - \Delta x)^{m+1}]}{\mathbb{E}^{\hat{x}}[\Delta x](m+1)} - \frac{\sigma^2 m}{2\nu} \mathbb{E}[x^{m-1}], \qquad \mathbb{E}[x] = 0 \tag{A.12}$$

- **Observation of ergodic moments with respect to the joint moments of state and age:** For observability of ergodic moments of $x^m a$, where a stand for the duration of the last action, we use Itō's lemma and the OST on $x_t^{m+1} t$:

$$\mathbb{E}^{\hat{S}} \left[\tau \left(\hat{x} - \Delta x \right)^{m+1} \right] = \mathbb{E}^{\hat{S}} \left[\int_0^\tau x_t^{m+1} dt \right] + (m+1)\nu \mathbb{E}^{\hat{S}} \left[\int_0^\tau x_t^m t dt \right] + \frac{\sigma^2 m(m+1)}{2} \mathbb{E}^{\hat{S}} \left[\int_0^\tau x_t^{m-1} t dt \right] \tag{A.13}$$

and therefore

$$\mathbb{E}[x^m a] = \frac{\mathbb{E}^{\hat{S}} \left[\tau \left(\hat{x} - \Delta x \right)^{m+1} \right]}{\nu(m+1)\mathbb{E}^{\hat{S}}[\tau]} - \frac{\mathbb{E}[x^{m+1}]}{\nu(m+1)} - \frac{\sigma^2 m}{2\nu} \mathbb{E}[x^{m-1} a] \tag{A.14}$$

with initial condition $\mathbb{E}[a] = \frac{\mathbb{E}^{\hat{S}}[\tau^2]}{2\mathbb{E}^{\hat{S}}[\tau]}$.

□

Proof of Proposition 2. We collapse this proof with the proof of proposition 3. □

Proof of Proposition 3. The proof is divided into Lemmas 2 to 7 for clarity. Some of the steps are developed in Online Appendix. □

Lemma 2. *[CIR as a function of $v_m'(x)$]*
Define $v_m(x) \equiv \mathbb{E}^x \left[\int_0^\tau x_t^m dt \right]$, then

$$CIR_m(\delta)/\delta = \int_{\underline{x}}^{\overline{x}} \left(v_m'(x) - \mathbb{E}[x^m] v_0'(x) \right) dF(x) + o(\delta) \tag{A.15}$$

Proof. The proof comes directly from Lemma 1. □

Lemma 3. *[Characterization of ergodic moments] Define the following values:*

$$v_m(x) \equiv \mathbb{E}^x\left[\int_0^\tau x_t^m dt\right], \qquad V_m^1(x) \equiv \mathbb{E}^x\left[\int_0^\tau e^{\frac{-\nu-\sqrt{\nu^2+2\lambda\sigma^2}}{\sigma^2}x_t} dt\right], \qquad V_m^2(x) \equiv \mathbb{E}^x\left[\int_0^\tau e^{\frac{-\nu+\sqrt{\nu^2+2\lambda\sigma^2}}{\sigma^2}x_t} dt\right]$$

They can be expressed as

$$v_m(x) = \frac{-e^{\xi_1 x}\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right] - e^{\xi_2 x}\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right] + \kappa^m(x)}{\lambda} \tag{A.16}$$

$$V_m^1(x) = -\frac{-e^{\xi_1 x}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2 x}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_1 x}x}{\sigma^2(\tilde{\nu}+\xi_1)} \tag{A.17}$$

$$V_m^2(x) = -\frac{-e^{\xi_1 x}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2 x}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_2 x}x}{\sigma^2(\tilde{\nu}+\xi_2)} \tag{A.18}$$

where the parameters are given by

$$\tilde{\nu} = \frac{\nu}{\sigma^2} \quad ; \quad \tilde{\lambda} = \frac{\lambda}{\sigma^2} \tag{A.19}$$

$$\xi_1 = -\tilde{\nu} - \sqrt{\tilde{\nu}^2 + 2\tilde{\lambda}} \quad ; \quad \xi_2 = -\tilde{\nu} + \sqrt{\tilde{\nu}^2 + 2\tilde{\lambda}} \tag{A.20}$$

$$\overline{\alpha}_1 = \frac{e^{\xi_1\overline{x}}}{e^{\xi_1\underline{x}+\xi_2\overline{x}} - e^{\xi_2\underline{x}+\xi_1\overline{x}}} \quad ; \quad \underline{\alpha}_1 = \frac{e^{\xi_1\underline{x}}}{e^{\xi_1\underline{x}+\xi_2\overline{x}} - e^{\xi_2\underline{x}+\xi_1\overline{x}}} \tag{A.21}$$

$$\overline{\alpha}_2 = \frac{e^{\xi_2\overline{x}}}{e^{\xi_1\underline{x}+\xi_2\overline{x}} - e^{\xi_2\underline{x}+\xi_1\overline{x}}} \quad ; \quad \underline{\alpha}_2 = \frac{e^{\xi_2\underline{x}}}{e^{\xi_1\underline{x}+\xi_2\overline{x}} - e^{\xi_2\underline{x}+\xi_1\overline{x}}} \tag{A.22}$$

$$\kappa^m(x) \equiv \sum_{i=0}^m b_{i,m}x^i \text{ with } b_{i,m} = \frac{m!}{i!}\left[\frac{\xi_1 + \xi_1\xi_2\frac{\tilde{\nu}}{\tilde{\lambda}}}{\xi_1-\xi_2}(\xi_1)^{i-m} + \frac{\xi_2 + \xi_1\xi_2\frac{\tilde{\nu}}{\tilde{\lambda}}}{\xi_2-\xi_1}(\xi_2)^{i-m}\right] \tag{A.23}$$

Proof. The derivation of $v_m(x)$ is in proposition E.3 in the Online Appendix and the derivation of $V^1(x)$ and $V^2(x)$ are in proposition E.7 in the Online Appendix. □

Lemma 4. *[Decomposition I of CIR_m] The following relation holds*

$$CIR_m(\delta)/\delta = \frac{\mathcal{K}_{1m} - \tilde{\nu}\mathcal{K}_{2m} - \mathbb{E}[x^m]\left(\mathcal{K}_{10} - \tilde{\nu}\mathcal{K}_{20}\right)}{\lambda\sigma^2\mathbb{E}^{\hat{x}}\left[\tau\right]} + o(\delta), \tag{A.24}$$

where

$$\mathcal{K}_{1m} = \left[-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right]\dots$$
$$\dots + \left[-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right] + \sigma^2\sum_{i=1}^m iv_{i-1}(\hat{x})b_{i,m} \tag{A.25}$$

$$\mathcal{K}_{2m} = \frac{-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_1\hat{x}}\hat{x}}{\tilde{\nu}+\xi_1}\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right]\dots$$
$$\dots + \frac{-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_2\hat{x}}\hat{x}}{\tilde{\nu}+\xi_2}\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right] \tag{A.26}$$

Proof. Using Theorem 2 and Lemma 3, we have that

$$\mathrm{CIR}_m(\delta)/\delta = \int_{\underline{x}}^{\overline{x}}\left[v_m'(x) - \mathbb{E}[x^m]v_0'(x)\right]dF(x) + o(\delta) \tag{A.27}$$

$$= \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau v_m'(x_t)dt\right] - \mathbb{E}[x^m]\mathbb{E}^{\hat{x}}\left[\int_0^\tau v_0'(x_t)dt\right]}{\mathbb{E}^{\hat{x}}[\tau]} + o(\delta). \tag{A.28}$$

Operating over $\sigma^2\lambda\mathbb{E}^{\hat{x}}\left[\int_0^\tau v_m'(x_t)dt\right]$ we have that

$$
\begin{aligned}
&\sigma^2\lambda\mathbb{E}^{\hat{x}}\left[\int_0^\tau v_m'(x_t)dt\right] \\
&= -\xi_1\sigma^2\mathbb{E}^{\hat{x}}\left[\int_0^\tau e^{\xi_1 x_t}dt\right]\left[\overline{\alpha}_2\kappa^m(\underline{x})-\underline{\alpha}_2\kappa^m(\overline{x})\right]-\xi_2\sigma^2\mathbb{E}^{\hat{x}}\left[\int_0^\tau e^{\xi_2 x_t}dt\right]\left[\underline{\alpha}_1\kappa^m(\overline{x})-\overline{\alpha}_1\kappa^m(x)\right]+\sigma^2\sum_{i=1}^m iv_{i-1}(\hat{x})b_{i,m} \\
&= \frac{\xi_1}{(\tilde{\nu}+\xi_1)}\left[-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x})-\underline{\alpha}_2\kappa^m(\overline{x})\right]\dots \\
&\dots+\frac{\xi_2}{(\tilde{\nu}+\xi_2)}\left[-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x})-\overline{\alpha}_1\kappa^m(\underline{x})\right]+\sigma^2\sum_{i=1}^m iv_{i-1}(\hat{x})b_{i,m} \\
&= \left[-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x})-\underline{\alpha}_2\kappa^m(\overline{x})\right]\dots \\
&\dots+\left[-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x})-\overline{\alpha}_1\kappa^m(\underline{x})\right]+\sigma^2\sum_{i=1}^m iv_{i-1}(\hat{x})b_{i,m} \\
&\dots-\frac{\tilde{\nu}}{\tilde{\nu}+\xi_1}\left[-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x})-\underline{\alpha}_2\kappa^m(\overline{x})\right]\dots \\
&\dots-\frac{\tilde{\nu}}{\tilde{\nu}+\xi_2}\left[-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x})-\overline{\alpha}_1\kappa^m(\underline{x})\right] \\
&= \mathcal{K}_{1m}-\tilde{\nu}\mathcal{K}_{2m}
\end{aligned}
\tag{A.29}
$$

where $\mathcal{K}_{1m}$ and $\mathcal{K}_{2m}$ are defined in equations above. □

Lemma 5. ***[Characterization of $\mathcal{K}_{1m}$ as a function of ergodic moments with respect to x]*** *The following relation holds*

$$
\mathcal{K}_{1m}=\lambda\left(v_{m+1}(\hat{x})-\hat{x}v_m(\hat{x})-\frac{\nu}{\lambda}\sum_{i=0}^m b_{i,m}v_i(\hat{x})\right)
\tag{A.30}
$$

Proof. For this proof it would be useful to define $\mathcal{T}=(\overline{\alpha}_2\underline{\alpha}_1-\overline{\alpha}_1\underline{\alpha}_2)^{-1}$. The following relation holds $\mathcal{T}$

$$
\mathcal{T}=(\overline{\alpha}_2\underline{\alpha}_1-\overline{\alpha}_1\underline{\alpha}_2)^{-1}=\frac{(e^{\xi_1\underline{x}+\xi_2\overline{x}}-e^{\xi_2\underline{x}+\xi_1\overline{x}})^2}{e^{\xi_1\underline{x}+\xi_2\overline{x}}-e^{\xi_2\underline{x}+\xi_1\overline{x}}}=e^{\xi_1\underline{x}+\xi_2\overline{x}}-e^{\xi_2\underline{x}+\xi_1\overline{x}}.
\tag{A.31}
$$

Thus

$$
e^{\xi_1\underline{x}}=\mathcal{T}\underline{\alpha}_1;e^{\xi_1\overline{x}}=\mathcal{T}\overline{\alpha}_1;\;e^{\xi_2\underline{x}}=\mathcal{T}\underline{\alpha}_2;e^{\xi_1\overline{x}}=\mathcal{T}\overline{\alpha}_2.
\tag{A.32}
$$

Departing from the definition of $\mathcal{K}_{1m}$ we have

$$
\begin{aligned}
\mathcal{K}_{1m}&=\left[-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x})-\underline{\alpha}_2\kappa^m(\overline{x})\right]\dots \\
&\dots+\left[-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x})-\overline{\alpha}_1\kappa^m(\underline{x})\right]\dots \\
&\dots+\sigma^2\sum_{i=1}^m iv_{i-1}(\hat{x})b_{i,m} \\
&=\left[-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\sum_{i=0}b_{i,m}\underline{x}^i-\underline{\alpha}_2\sum_{i=0}b_{i,m}\overline{x}^i\right]\dots \\
&\dots+\left[-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2-e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2)-e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1-e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right)+e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\sum_{i=0}b_{i,m}\overline{x}^i-\overline{\alpha}_1\sum_{i=0}b_{i,m}\underline{x}^i\right]\dots \\
&\dots+\sigma^2\sum_{i=1}^m i\mathbb{E}^{\hat{x}}\left[\int_0^\tau x^{i-1}\right]b_{i,m}.
\end{aligned}
\tag{A.33}
$$

Using the definition of $\mathcal{T}$

$$
\begin{aligned}
\mathcal{K}_{1m} &= \hat{x}[-v_m(\hat{x})\lambda + \kappa^m(\hat{x})] + \frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} ib_{i,m}\left[-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\kappa^{i-1}(\underline{x}) - \underline{\alpha}_2\kappa^{i-1}(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\kappa^{i-1}(\overline{x}) - \overline{\alpha}_1\kappa^{i-1}(\underline{x})\right] + \kappa^{i-1}(\hat{x})\right]\dots \\
&\dots - e^{\xi_1\hat{x}}\mathcal{T}\left[(\underline{\alpha}_1\underline{x}\overline{\alpha}_2 - \overline{\alpha}_1\overline{x}\underline{\alpha}_2)\left[\overline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\underline{x}^i - \underline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\overline{x}^i\right] + (\underline{\alpha}_2\underline{x}\overline{\alpha}_2 - \overline{\alpha}_2\overline{x}\underline{\alpha}_2)\left[\underline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\overline{x}^i - \overline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\underline{x}^i\right]\right]\dots \\
&\dots - e^{\xi_2\hat{x}}\mathcal{T}\left[(\overline{\alpha}_1\overline{x}\underline{\alpha}_1 - \underline{\alpha}_1\underline{x}\overline{\alpha}_1)\left[\overline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\underline{x}^i - \underline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\overline{x}^i\right] + (\overline{\alpha}_2\overline{x}\underline{\alpha}_1 - \underline{\alpha}_2\underline{x}\overline{\alpha}_1)\left[\underline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\overline{x}^i - \overline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\underline{x}^i\right]\right] \\
&= \hat{x}[-v_m(\hat{x})\lambda + \kappa^m(\hat{x})] + \frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} ib_{i,m}\left[-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\kappa^{i-1}(\underline{x}) - \underline{\alpha}_2\kappa^{i-1}(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\kappa^{i-1}(\overline{x}) - \overline{\alpha}_1\kappa^{i-1}(\underline{x})\right] + \kappa^{i-1}(\hat{x})\right]\dots \\
&\dots - e^{\xi_1\hat{x}}\mathcal{T}\left[[\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2]\left[\overline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\underline{x}^{i+1} - \underline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\overline{x}^{i+1}\right] + \sum_{i=0}^{m} b_{i,m}\underline{x}^i\overline{x}\underbrace{[\overline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2 - \overline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2]}_{=0} + \sum_{i=0}^{m} b_{i,m}\overline{x}^i\underline{x}\underbrace{[\underline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2 - \underline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2]}_{=0}\right]\dots \\
&\dots - e^{\xi_2\hat{x}}\mathcal{T}\left[[\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2]\left[\underline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\overline{x}^{i+1} - \overline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\underline{x}^{i+1}\right] + \sum_{i=0}^{m} b_{i,m}\underline{x}^i\overline{x}\underbrace{[\overline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1 - \overline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1]}_{=0} + \sum_{i=0}^{m} b_{i,m}\overline{x}^i\underline{x}\underbrace{[\underline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1 - \underline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1]}_{=0}\right] \\
&= \hat{x}[-v_m(\hat{x})\lambda + \kappa^m(\hat{x})] + \frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} ib_{i,m}\left[-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\kappa^{i-1}(\underline{x}) - \underline{\alpha}_2\kappa^{i-1}(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\kappa^{i-1}(\overline{x}) - \overline{\alpha}_1\kappa^{i-1}(\underline{x})\right] + \kappa^{i-1}(\hat{x})\right]\dots \\
&\dots - e^{\xi_1\hat{x}}\left[\left[\overline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\underline{x}^{i+1} - \underline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\overline{x}^{i+1}\right]\right] - e^{\xi_2\hat{x}}\left[\left[\underline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\overline{x}^{i+1} - \overline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\underline{x}^{i+1}\right]\right]. \qquad (A.34)
\end{aligned}
$$

$$(A.35)$$

Define

$$\Psi_i(x) \equiv -e^{\xi_1\hat{x}}\left[\overline{\alpha}_2 x^i - \underline{\alpha}_2 x^i\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1 x^i - \overline{\alpha}_1 x^i\right] + x^i, \qquad (A.36)$$

and rewrite as

$$
\begin{aligned}
\mathcal{K}_{1m} &= -\hat{x}v_m(\hat{x})\lambda + \frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} ib_{i,m}\left[-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\kappa^{i-1}(\underline{x}) - \underline{\alpha}_2\kappa^{i-1}(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\kappa^{i-1}(\overline{x}) - \overline{\alpha}_1\kappa^{i-1}(\underline{x})\right] + \kappa^{i-1}(\hat{x})\right]\dots \\
&\dots - e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\underline{x}^{i+1} - \underline{\alpha}_2\sum_{i=0}^{m} b_{i,m}\overline{x}^{i+1}\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\overline{x}^{i+1} - \overline{\alpha}_1\sum_{i=0}^{m} b_{i,m}\underline{x}^{i+1}\right] + \sum_{i=0}^{m} b_{i,m}\hat{x}^{i+1} \\
&= -\hat{x}\lambda v_m(\hat{x}) + \frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} ib_{i,m}\sum_{j=0}^{i-1} b_{j,i-1}\Psi(j) + \sum_{i=0}^{m} b_{i,m}\Psi(i+1)
\end{aligned}
\qquad (A.37)
$$

Using Proposition E.8 in the Online Appendix and $\kappa^m(x) \equiv \sum_{i=0}^{m} b_{i,m}x^i$

$$\frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} ib_{i,m}\sum_{j=0}^{i-1} b_{j,i-1}\Psi(j) + \sum_{i=0}^{m} b_{i,m}\Psi(i+1) = \sum_{i=0}^{m+1} b_{i,m+1}\Psi(i) - \frac{\nu}{\lambda}\sum_{i=0}^{m} b_{i,m}\sum_{j=0}^{i} b_{j,i}\Psi(j), \qquad (A.38)$$

we have

$$\mathcal{K}_{1,m} = \lambda\left(v_{m+1}(\hat{x}) - \hat{x}v_m(\hat{x}) - \frac{\nu}{\lambda}\sum_{i=0}^{m} b_{i,m}v_i(\hat{x})\right) \qquad (A.39)$$

$\square$

Lemma 6. ***[Characterization of ergodic moments with respect to age]*** *The following relation holds*

$$\mathbb{E}[ax^m] = \frac{v_m'(\hat{x}) + \mathcal{C}_{1m} + \mathcal{C}_{2m}}{\mathbb{E}^{\hat{x}}[\tau]\,\lambda} \qquad (A.40)$$

$$\mathcal{C}_{1m} = \frac{\mathcal{K}_{2m}}{\sigma^2} \qquad (A.41)$$

$$\mathcal{C}_{2m} = \frac{1}{\lambda}\sum_{i=0}^{m-1} b_{i,m}v_i(\hat{x})\lambda \qquad (A.42)$$

Proof. From proposition E.6, we have that

$$\mathbb{E}[ax^m] = \frac{h'_m(0)}{\mathbb{E}^{\hat{x}}[\tau]} \tag{A.43}$$

$$h_m(\varphi) = \frac{-e^{\xi_1(\varphi)\hat{x}}\left[\overline{\alpha}_2(\varphi)\kappa^m(\underline{x},\varphi) - \underline{\alpha}_2(\varphi)\kappa^m(\overline{x},\varphi)\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1(\varphi)\kappa^m(\overline{x},\varphi) - \overline{\alpha}_1(\varphi)\kappa^m(\underline{x},\varphi)\right] + \kappa^m(\hat{x})}{\lambda - \varphi} \tag{A.44}$$

where

$$\tilde{\lambda}(\varphi) = \frac{\lambda - \varphi}{\sigma^2} \tag{A.45}$$

$$\xi_1(\varphi) = -\tilde{\nu} - \sqrt{\tilde{\nu}^2 + 2\tilde{\lambda}(\varphi)}\ ;\ \xi_2(\varphi) = -\tilde{\nu} + \sqrt{\tilde{\nu}^2 + 2\tilde{\lambda}(\varphi)} \tag{A.46}$$

$$\overline{\alpha}_1(\varphi) = \frac{e^{\xi_1(\varphi)\overline{x}}}{e^{\xi_1(\varphi)\underline{x}+\xi_2(\varphi)\overline{x}} - e^{\xi_2(\varphi)\underline{x}+\xi_1(\varphi)\overline{x}}}\ ;\ \underline{\alpha}_1(\varphi) = \frac{e^{\xi_1(\varphi)\underline{x}}}{e^{\xi_1(\varphi)\underline{x}+\xi_2(\varphi)\overline{x}} - e^{\xi_2(\varphi)\underline{x}+\xi_1(\varphi)\overline{x}}} \tag{A.47}$$

$$\overline{\alpha}_2(\varphi) = \frac{e^{\xi_2(\varphi)\overline{x}}}{e^{\xi_1(\varphi)\underline{x}+\xi_2(\varphi)\overline{x}} - e^{\xi_2(\varphi)\underline{x}+\xi_1(\varphi)\overline{x}}}\ ;\ \underline{\alpha}_2(\varphi) = \frac{e^{\xi_2(\varphi)\underline{x}}}{e^{\xi_1(\varphi)\underline{x}+\xi_2(\varphi)\overline{x}} - e^{\xi_2(\varphi)\underline{x}+\xi_1(\varphi)\overline{x}}} \tag{A.48}$$

$$\kappa_j^m(x,\varphi) = \sum_{i=0}^{j}(x)^i\,\frac{m!}{i!}\left[\frac{\xi_1(\varphi)+\xi_1(\varphi)\xi_2(\varphi)\frac{\tilde{\nu}}{\tilde{\lambda}(\varphi)}}{\xi_1(\varphi)-\xi_2(\varphi)}\xi_1(\varphi)^{i-m} + \frac{\xi_2(\varphi)+\xi_1(\varphi)\xi_2(\varphi)\frac{\tilde{\nu}}{\tilde{\lambda}(\varphi)}}{\xi_2(\varphi)-\xi_1(\varphi)}\xi_2(\varphi)^{i-m}\right]. \tag{A.49}$$

Taking the derivative of $h_m(\varphi)$, and using the result that and that

$$\frac{de^{\xi_1(\varphi)\hat{x}}}{d\varphi} = -\frac{e^{\xi_1(\varphi)\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_1(\varphi))} \tag{A.50}$$

$$\frac{de^{\xi_2(\varphi)\hat{x}}}{d\varphi} = -\frac{e^{\xi_2(\varphi)\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_2(\varphi))}, \tag{A.51}$$

we have that

$$\begin{aligned}
h'_m(\varphi) = {} & \frac{-e^{\xi_1(\varphi)\hat{x}}\left[\overline{\alpha}_2(\varphi)\kappa^m(\underline{x},\varphi) - \underline{\alpha}_2(\varphi)\kappa^m(\overline{x},\varphi)\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1(\varphi)\kappa^m(\overline{x},\varphi) - \overline{\alpha}_1(\varphi)\kappa^m(\underline{x},\varphi)\right] + \overline{\kappa}^m_m(\hat{x},\varphi)}{(\lambda-\varphi)^2}\dots \\
& \dots + \frac{\frac{e^{\xi_1(\varphi)\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_1(\varphi))}\left[\overline{\alpha}_2(\varphi)\kappa^m(\underline{x},\varphi) - \underline{\alpha}_2(\varphi)\kappa^m(\overline{x},\varphi)\right] + \frac{e^{\xi_2(\varphi)\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_2(\varphi))}\left[\underline{\alpha}_1(\varphi)\kappa^m(\overline{x},\varphi) - \overline{\alpha}_1(\varphi)\kappa^m(\underline{x},\varphi)\right]}{(\lambda-\varphi)}\dots \\
& \dots + \frac{-e^{\xi_1(\varphi)\hat{x}}\left[\frac{d\overline{\alpha}_2(\varphi)}{d\varphi}\kappa^m(\underline{x},\varphi) - \frac{\underline{\alpha}_2(\varphi)}{d\varphi}\kappa^m(\overline{x},\varphi)\right] - e^{\xi_2(\varphi)\hat{x}}\left[\frac{\underline{\alpha}_1(\varphi)}{d\varphi}\kappa^m(\overline{x},\varphi) - \frac{\overline{\alpha}_1(\varphi)}{d\varphi}\kappa^m(\underline{x},\varphi)\right]}{(\lambda-\varphi)}\dots \\
& \dots + \frac{-e^{\xi_1(\varphi)\hat{x}}\left[\overline{\alpha}_2(\varphi)\frac{d\kappa^m(\underline{x},\varphi)}{d\varphi} - \underline{\alpha}_2(\varphi)\frac{\kappa^m(\overline{x},\varphi)}{d\varphi}\right] - e^{\xi_2(\varphi)\hat{x}}\left[\underline{\alpha}_1(\varphi)\frac{\kappa^m(\overline{x},\varphi)}{d\varphi} - \overline{\alpha}_1(\varphi)\frac{\kappa^m(\underline{x},\varphi)}{d\varphi}\right] + \frac{\overline{\kappa}^m(\hat{x},\varphi)}{d\varphi}}{(\lambda-\varphi)},
\end{aligned} \tag{A.52}$$

and evaluating at zero, we have that

$$\begin{aligned}
h'_m(0) = {} & \frac{-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right] + \kappa^m(\hat{x})}{\lambda^2}\dots \\
& \dots + \frac{\frac{e^{\xi_1\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right] + \frac{e^{\xi_2\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right]}{\lambda}\dots \\
& \dots + \frac{-e^{\xi_1\hat{x}}\left[\left.\frac{d\overline{\alpha}_2(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\underline{x}) - \left.\frac{\underline{\alpha}_2(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\left.\frac{\underline{\alpha}_1(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\overline{x}) - \left.\frac{\overline{\alpha}_1(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\underline{x})\right]}{\lambda}\dots \\
& \dots + \frac{-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\left.\frac{d\kappa^m_m(\underline{x},\varphi)}{d\varphi}\right|_{\varphi=0} - \underline{\alpha}_2\left.\frac{\kappa^m_m(\overline{x},\varphi)}{d\varphi}\right|_{\varphi=0}\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\left.\frac{\kappa^m_m(\overline{x},\varphi)}{d\varphi}\right|_{\varphi=0} - \overline{\alpha}_1\left.\frac{\kappa^m_m(\underline{x},\varphi)}{d\varphi}\right|_{\varphi=0}\right] + \left.\frac{\kappa^m(\hat{x},\varphi)}{d\varphi}\right|_{\varphi=0}}{\lambda}, \\
= {} & \frac{v'_m(\hat{x}) + \mathcal{C}_{1m} + \mathcal{C}_{2m}}{\lambda}
\end{aligned} \tag{A.53}$$

where we have defined

$$\begin{aligned}
\mathcal{C}_{1m} &= \frac{e^{\xi_1 \hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right] + \frac{e^{\xi_2 \hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right]\dots \\
&\quad \dots + -e^{\xi_1\hat{x}}\left[\left.\frac{d\overline{\alpha}_2(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\underline{x}) - \left.\frac{\underline{\alpha}_2(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\left.\frac{\underline{\alpha}_1(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\overline{x}) - \left.\frac{\overline{\alpha}_1(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\underline{x})\right]\dots \\
\mathcal{C}_{2m} &= -e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\left.\frac{d\kappa^m(\underline{x},\varphi)}{d\varphi}\right|_{\varphi=0} - \underline{\alpha}_2\left.\frac{\kappa^m(\overline{x},\varphi)}{d\varphi}\right|_{\varphi=0}\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\left.\frac{\kappa^m(\overline{x},\varphi)}{d\varphi}\right|_{\varphi=0} - \overline{\alpha}_1\left.\frac{\kappa^m(\underline{x},\varphi)}{d\varphi}\right|_{\varphi=0}\right] + \left.\frac{\kappa^m(\hat{x},\varphi)}{d\varphi}\right|_{\varphi=0}.
\end{aligned} \tag{A.54}$$

Now we show that $\frac{\mathcal{K}_{2m}}{\sigma^2} = \mathcal{C}_{1m}$. Taking derivative of $\underline{\alpha}_i(\varphi)$ and $\overline{\alpha}_i(\varphi)$

$$\sigma^2 \left.\frac{d\overline{\alpha}_2(\varphi)}{d\varphi}\right|_{=0} = \overline{\alpha}_2\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right] - \frac{\overline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_2} \tag{A.55}$$

$$\sigma^2 \left.\frac{d\overline{\alpha}_1(\varphi)}{d\varphi}\right|_{=0} = \overline{\alpha}_1\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right] - \frac{\overline{\alpha}_1\overline{x}}{\tilde{\nu}+\xi_1} \tag{A.56}$$

$$\sigma^2 \left.\frac{d\underline{\alpha}_2(\varphi)}{d\varphi}\right|_{=0} = \underline{\alpha}_2\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right] - \frac{\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2} \tag{A.57}$$

$$\sigma^2 \left.\frac{d\underline{\alpha}_1(\varphi)}{d\varphi}\right|_{=0} = \underline{\alpha}_1\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right] - \frac{\underline{\alpha}_1\underline{x}}{\tilde{\nu}+\xi_1}. \tag{A.58}$$

Operating and using the definition $\mathcal{T}_2 = (\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2)^{-1}$

$$\sigma^2 \left.\frac{d\overline{\alpha}_2(\varphi)}{d\varphi}\right|_{=0} = \overline{\alpha}_2\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\overline{\alpha}_1\underline{\alpha}_2\left(\overline{x} - \underline{x}\right)}{\tilde{\nu}+\xi_2}\right] \tag{A.59}$$

$$\sigma^2 \left.\frac{d\overline{\alpha}_1(\varphi)}{d\varphi}\right|_{=0} = \overline{\alpha}_1\mathcal{T}\left[\frac{-\underline{\alpha}_1\overline{\alpha}_2(\underline{x} - \overline{x})}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right] \tag{A.60}$$

$$\sigma^2 \left.\frac{d\underline{\alpha}_2(\varphi)}{d\varphi}\right|_{=0} = \underline{\alpha}_2\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\left(\overline{x} - \underline{x}\right)}{\tilde{\nu}+\xi_2}\right] \tag{A.61}$$

$$\sigma^2 \left.\frac{d\underline{\alpha}_1(\varphi)}{d\varphi}\right|_{=0} = \underline{\alpha}_1\mathcal{T}\left[\frac{-\overline{\alpha}_1\underline{\alpha}_2(\overline{x} - \underline{x})}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right]. \tag{A.62}$$

Using (A.59)-(A.62), and departing from

$$\begin{aligned}
\mathcal{C}_{1m} &= \left[\frac{e^{\xi_1 \hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right] + \frac{e^{\xi_2 \hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right]\right]\dots \\
&\quad \dots + \left[-e^{\xi_1\hat{x}}\left[\left.\frac{d\overline{\alpha}_2(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\underline{x}) - \left.\frac{\underline{\alpha}_2(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\overline{x})\right] - e^{\xi_2\hat{x}}\left[\left.\frac{\underline{\alpha}_1(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\overline{x}) - \left.\frac{\overline{\alpha}_1(\varphi)}{d\varphi}\right|_{\varphi=0}\kappa^m(\underline{x})\right]\right] \\
&= \frac{e^{\xi_1 \hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right] + \frac{e^{\xi_2 \hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right]\dots \\
&\quad \dots - \frac{e^{\xi_1\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right] - \frac{e^{\xi_1\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_2)}\overline{\alpha}_2\underline{\alpha}_2\left(\overline{x} - \underline{x}\right)\left(\overline{\alpha}_1\kappa^m(\underline{x}) - \underline{\alpha}_1\kappa^m(\overline{x})\right) \\
&\quad \dots - \frac{e^{\xi_2\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_1)}\overline{\alpha}_1\underline{\alpha}_1(\underline{x} - \overline{x})\left[\underline{\alpha}_2\kappa^m(\overline{x}) - \overline{\alpha}_2\kappa^m(\underline{x})\right] - \frac{e^{\xi_2\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right] \\
&= \frac{1}{\sigma^2(\tilde{\nu}+\xi_1)}\left[-e^{\xi_1\hat{x}}\mathcal{T}\left[\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}\right] - e^{\xi_2\hat{x}}\mathcal{T}\overline{\alpha}_1\underline{\alpha}_1(\underline{x} - \overline{x}) + e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right]\dots \\
&\quad \dots + \frac{1}{\sigma^2(\tilde{\nu}+\xi_2)}\left[-e^{\xi_1\hat{x}}\mathcal{T}\overline{\alpha}_2\underline{\alpha}_2\left(\overline{x} - \underline{x}\right) - e^{\xi_2\hat{x}}\mathcal{T}\left[\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}\right] + e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right] \\
&= \frac{1}{\sigma^2(\tilde{\nu}+\xi_1)}\left[-e^{\xi_1\hat{x}}(e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2\hat{x}}\left(e^{\xi_1\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_1\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_1\hat{x}}\hat{x}\right]\left[\overline{\alpha}_2\kappa^m(\underline{x}) - \underline{\alpha}_2\kappa^m(\overline{x})\right]\dots \\
&\quad \dots \frac{1}{\sigma^2(\tilde{\nu}+\xi_2)}\left[-e^{\xi_1\hat{x}}(e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_2 - e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_2) - e^{\xi_2\hat{x}}\left(e^{\xi_2\overline{x}}\overline{x}\underline{\alpha}_1 - e^{\xi_2\underline{x}}\underline{x}\overline{\alpha}_1\right) + e^{\xi_2\hat{x}}\hat{x}\right]\left[\underline{\alpha}_1\kappa^m(\overline{x}) - \overline{\alpha}_1\kappa^m(\underline{x})\right].
\end{aligned} \tag{A.63}$$

Thus,

$$\frac{\mathcal{K}_{2m}}{\sigma^2} = \mathcal{C}_{1m}. \tag{A.64}$$

Finally, we need to show that

$$\mathcal{C}_{2m} = \frac{1}{\lambda}\sum_{i=0}^{m-1} b_{i,m} v_i(\hat{x})\lambda. \tag{A.65}$$

Departing form the definition of $\mathcal{C}_{2m}$ and using Proposition E.9 in the Online Appendix, we have that

$$\begin{aligned}
\mathcal{C}_{2m} &= -e^{\xi_1 \hat{x}}\left[\overline{\alpha}_2 \left.\frac{d\kappa^m(\underline{x},\varphi)}{d\varphi}\right|_{\varphi=0} - \underline{\alpha}_2 \left.\frac{\kappa^m(\overline{x},\varphi)}{d\varphi}\right|_{\varphi=0}\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1 \left.\frac{\kappa^m(\overline{x},\varphi)}{d\varphi}\right|_{\varphi=0} - \overline{\alpha}_1 \left.\frac{\kappa^m(\underline{x},\varphi)}{d\varphi}\right|_{\varphi=0}\right] + \left.\frac{\kappa^m(\hat{x},\varphi)}{d\varphi}\right|_{\varphi=0} \\
&= \sum_{i=0}^{m-1} \left.\frac{dB_{i,m}(\varphi)}{d\varphi}\right|_{\varphi=0} \Psi(i) = \frac{1}{\lambda}\sum_{i=0}^{m-1}\sum_{j=i}^{m-1} b_{i,j} b_{j,m}\Psi(i) = \frac{1}{\lambda}\sum_{j=0}^{m-1}\sum_{i=0}^{j} b_{j,m} b_{i,j}\Psi(i) = \frac{1}{\lambda}\sum_{j=0}^{m-1} b_{j,m} v_j(\hat{x})\lambda
\end{aligned}$$

Thus, we have the result. □

Lemma 7. ***[Final characterization of the CIR]*** *The following relation holds*

$$CIR_m(\delta)/\delta = \frac{\mathbb{E}\left[x^{m+1}\right] - \nu \mathbb{C}ov\left[a, x^m\right]}{\sigma^2} + o(\delta), \tag{A.66}$$

Proof. To show the previous Lemma, first, we need to characterize

$$\frac{\mathcal{K}_{1m} - \tilde{\nu}\mathcal{K}_{2m}}{\lambda\sigma^2\mathbb{E}^{\hat{x}}\left[\tau\right]} \tag{A.67}$$

Using Lemma 5 and Lemma 6

$$\begin{aligned}
\frac{\mathcal{K}_{1m} - \tilde{\nu}\mathcal{K}_{2m}}{\lambda\sigma^2\mathbb{E}^{\hat{x}}\left[\tau\right]} &= \frac{\lambda\left(v_{m+1}(\hat{x}) - \hat{x}v_m(\hat{x}) - \frac{\nu}{\lambda}\sum_{i=0}^{m} b_{i,m}v_i(\hat{x})\right) - \nu\left[\mathbb{E}^{\hat{x}}[\tau]\lambda\mathbb{E}^{\hat{x}}\left[ax^m\right] - v_m(\hat{x}) - \frac{\sum_{i=0}^{m-1} b_{i,m}v_i(\hat{x})\lambda}{\lambda}\right]}{\lambda\sigma^2\mathbb{E}^{\hat{x}}\left[\tau\right]}, \qquad &\text{(A.68)} \\
&= \frac{\mathbb{E}^{\hat{x}}\left[x^{m+1}\right] - \hat{x}\mathbb{E}^{\hat{x}}\left[x^m\right] - \frac{\nu}{\lambda}b_{m,m}\mathbb{E}[x^m] + \frac{\nu}{\lambda}\mathbb{E}[x^m]}{\sigma^2} - \nu\frac{\mathbb{E}^{\hat{x}}\left[ax^m\right]}{\sigma^2}. &\text{(A.69)}
\end{aligned}$$

Since $b_{m,m} = 1$, we have that

$$\frac{\mathcal{K}_{1m} - \tilde{\nu}\mathcal{K}_{2m}}{\lambda\sigma^2\mathbb{E}^{\hat{x}}\left[\tau\right]} = \frac{\mathbb{E}^{\hat{x}}\left[x^{m+1}\right] - \hat{x}\mathbb{E}^{\hat{x}}\left[x^m\right] - \nu\mathbb{E}^{\hat{x}}\left[ax^m\right]}{\sigma^2} \tag{A.70}$$

Since $\mathbb{E}[x] = 0$

$$\begin{aligned}
\text{CIR}_m(\delta)/\delta &= \frac{\mathcal{K}_{1m} - \tilde{\nu}\mathcal{K}_{2m} - \mathbb{E}[x^m]\left(\mathcal{K}_{10} - \tilde{\nu}\mathcal{K}_{20}\right)}{\lambda\sigma^2\mathbb{E}^{\hat{x}}\left[\tau\right]} \\
&= \left[\frac{\mathbb{E}\left[x^{m+1}\right] - \hat{x}\mathbb{E}\left[x^m\right] - \nu\mathbb{E}^{\hat{x}}\left[ax^m\right]}{\sigma^2} - \mathbb{E}[x^m]\frac{\mathbb{E}\left[x\right] - \hat{x} - \nu\mathbb{E}\left[a\right]}{\sigma^2}\right] + o(\delta), \\
&= \frac{\mathbb{E}\left[x^{m+1}\right] - \nu\left[\mathbb{E}\left[ax^m\right] - \mathbb{E}[x^m]\mathbb{E}\left[a\right]\right]}{\sigma^2} + o(\delta) \\
&= \frac{\mathbb{E}\left[x^{m+1}\right] - \nu\mathbb{C}ov\left[a, x^m\right]}{\sigma^2} + o(\delta), \qquad \text{(A.71)}
\end{aligned}$$

where in the last step we use the covariance formula $\mathbb{C}ov^{\hat{x}}\left[a, x^m\right] = \mathbb{E}[x^m]\mathbb{E}\left[a\right]$. □

Proposition 4. *Assume that:*

- *The uncontrolled state follows* $d\tilde{x}_t = \nu dt + \sigma dW_t$*, with* W_t *a Wiener process;*
- $\left(\left\{\int_0^t x_s^m s^n \ dW_s\right\}_t, \tau\right)$ *are a well-defined stopping processes for any* m *and* $n = 0, 1$*; and*
- *Define the function*

$$g_m(x) \equiv \mathbb{E}^{\hat{x}+x}\left[(\hat{x} - \Delta x)^m\right] - \mathbb{E}^{\hat{x}}\left[(\hat{x} - \Delta x + x)^m\right],$$

1. ***Aggregation:*** *To a first order, the CIR is given by*

$$\mathcal{A}_m(\delta)/\delta = \mathcal{Z}_m - \mathcal{M}_m[x]\Theta_0 + o(\delta) \tag{A.72}$$

where the intensive and extensive margin are given by

$$\mathcal{Z}_m = \Theta_m + \Gamma_m - \frac{\sigma^2 m}{2\nu}\mathcal{Z}_{m-1} \tag{A.73}$$

$$\Gamma_m \equiv \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau \varphi_m^\Gamma(x_t)dt\right]}{\mathbb{E}^{\hat{x}}[\tau]} \quad ; \quad \varphi_m^\Gamma(x_t) \equiv \frac{1}{\nu}\left(\mathbb{E}^x\left[x_\tau^m\right] - x_t^m\right) \tag{A.74}$$

$$\Theta_m \equiv \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau \varphi_m^\Theta(x_t)dt\right]}{\mathbb{E}^{\hat{x}}[\tau]} \quad ; \quad \varphi_m^\Theta(S_t) \equiv \frac{1}{\nu}\left[\frac{\partial \mathbb{E}^x\left[x_\tau^{m+1}/(m+1)\right]}{\partial x} - \mathbb{E}^x\left[x_\tau^m\right]\right] \tag{A.75}$$

2. ***Representation for the intensive margin:***

$$\Gamma_m = m\mathcal{M}_{m-1,1}[x,a] + \frac{\mathbb{1}_{\{m\geq 2\}}\sigma^2 m(m-1)}{2\nu}\mathcal{M}_{m-2,1}[x,a] \tag{A.76}$$

3. ***Representation for the extensive margin:***

$$\Theta_m = \sum_{j=0}^{\infty}\theta_{m,j}\mathcal{M}_j[x] \quad \textit{with} \quad \theta_{m,j} \equiv \frac{2}{\sigma^2(m+1)}\sum_{k\geq j}^{\infty}\frac{\hat{x}^{k-j}}{k!j!}\left[\frac{d^{k+1}g_{m+2}(0)}{dx^{k+1}}/m + 2 - \frac{d^k g_{m+1}(0)}{dx^k}\right]. \tag{A.77}$$

   - *If* $\tau | x_t \sim \tau$*,* $g_m(x) = \theta(m,j) = 0$ *for all* m, i*.*

The proof is divided into 4 Lemmas for clarity.

Lemma 8. *[**Aggregation**] To a first order, the transitional dynamics of the m-th moment are given by*

$$\mathcal{A}_m(\delta)/\delta = \times\left(\mathcal{Z}_m - \mathcal{M}_m[x]\Theta_0\right) + o(\delta) \tag{A.78}$$

where the intensive and extensive margin are given by

$$\mathcal{Z}_m = \Theta_m + \Gamma_m - \frac{\sigma^2 m}{2\nu}\mathcal{Z}_{m-1} \tag{A.79}$$

$$\Gamma_m \equiv \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau \varphi_m^\Gamma(x_t)dt\right]}{\mathbb{E}^{\hat{x}}[\tau]} \quad ; \quad \varphi_m^\Gamma(x_t) \equiv \frac{1}{\nu}\left(\mathbb{E}^x\left[x_\tau^m\right] - x_t^m\right) \tag{A.80}$$

$$\Theta_m \equiv \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau \varphi_m^\Theta(x_t)dt\right]}{\mathbb{E}^{\hat{x}}[\tau]} \quad ; \quad \varphi_m^\Theta(S_t) \equiv \frac{1}{\nu}\left[\frac{\partial\mathbb{E}^x\left[x_\tau^{m+1}/(m+1)\right]}{\partial x} - \mathbb{E}^x\left[x_\tau^m\right]\right] \tag{A.81}$$

Proof. The proof consists of 6 steps.

Characterization of CIR as the recursive problem of a representative agent. Fix an $m \in \mathbb{N}$. Start from the CIR's definition:

$$\mathcal{A}_m(\delta) = \mathbb{E}\left[\int_0^\infty \left(x_t(\omega|\delta)^m - \mathcal{M}_m[x]\right)dt\right], \tag{A.82}$$

where the expectation is taken across agents ω. Let $\{\tau_i\}_{i=1}^\infty$ be the sequence of stopping times after the arrival of the perturbation. In (1), we write the CIR as the cumulative deviations between time $t=0$ and the first stopping time τ_1 plus the sum of deviations between all future stopping times. In (2), we use the Law of Iterated Expectations to condition on the information set $\mathcal{F}_{\tau_i}$. In (3), we use the Strong Markov Property of the Brownian motion, the assumption of homogenous resets and that $\hat{x}$ is independent of δ for $i \geq 1$ to change the conditioning from $x_{\tau_i+h}|\mathcal{F}_{\tau_i}$ to $x_h|\hat{x}$ and write the problem recursively. To get (4), we show that *every element* inside the infinite sum is equal to zero. For this purpose, recall the relationship between ergodic moments and expected duration derived in Auxiliary Theorem 2, $\mathcal{M}_m[x] = \mathbb{E}^{\hat{x}}\left[\int_0^\tau x_t(\delta|\omega)^m\right]/\mathbb{E}^{\hat{x}}[\tau]$, and thus we are left with the simple expression in the fourth line (we also relabel τ_1 as τ):

$$\begin{aligned}
\mathcal{A}_m(\delta) &=^{(1)} \mathbb{E}\left[\int_0^{\tau_1}\left(x_t(\delta|\omega)^m - \mathcal{M}_m[x]\right)dt + \sum_{i=1}^\infty \int_{\tau_i}^{\tau_{i+1}}\left(x_t(\delta|\omega)^m - \mathcal{M}_m[x]\right)dt\right] \\
&=^{(2)} \mathbb{E}\left[\int_0^{\tau_1}\left(x_t(\delta|\omega)^m - \mathcal{M}_m[x]\right)dt + \sum_{i=1}^\infty \mathbb{E}\left[\int_{\tau_i}^{\tau_{i+1}}\left(x_t(\delta|\omega)^m - \mathcal{M}_m[x]\right)dt\Big|\mathcal{F}_{\tau_i}\right]\right] \\
&=^{(3)} \mathbb{E}\left[\int_0^{\tau_1}\left(x_t(\delta|\omega)^m - \mathcal{M}_m[x]\right)dt\right] + \mathbb{E}\left[\sum_{i=1}^\infty \underbrace{\mathbb{E}\left[\int_0^\tau x_t(\delta|\omega)^m\,dt\Big|\hat{x}\right] - \mathcal{M}_m[x]\mathbb{E}[\tau|\hat{x}]}_{=0}\right] \\
&=^{(4)} \mathbb{E}\left[\int_0^\tau \left(x_t(\delta|\omega)^m\right)dt\right] - \mathcal{M}_m[x]\mathbb{E}[\tau].
\end{aligned}$$

As a final step, define the following value function conditional on a particular initial condition x:

$$v^m(x) \equiv \mathbb{E}^x\left[\int_0^\tau x_t(\omega)^m dt\right] - \mathcal{M}_m[x]\mathbb{E}^x[\tau], \tag{A.83}$$

and notice that $\mathcal{A}_m(\delta)$ is equal to the average of $v^m(x)$ across all initial conditions after the perturbation, given by the shift in the ergodic distribution ($F_0(x) = F(x-\delta)$):

$$\mathcal{A}_m(\delta) = \int v^m(x)dF(x-\delta). \tag{A.84}$$

2. State's support. Since Brownian motions are continuous in t, and initial conditions are identical across agents (by the assumption of homogeneous resets), the ergodic set is connected. Thus, the support of x is given by an interval $[\underline{x}, \overline{x}]$.

3. Taylor approximation to $\mathcal{A}_m(\delta)$ and decomposition into two terms. We do a first order Taylor approximation of $\mathcal{A}_m(\delta)$ around zero: $\mathcal{A}_m(\delta) = \mathcal{A}_m(0) + \mathcal{A}_m'(0)\delta$. Since $\mathcal{A}_m(0) = 0$ by definition, we have that: $\mathcal{A}_m(\delta) = \delta\mathcal{A}_m'(0)$, which we now characterize. Start from the representation in (A.84), expressed in terms of the marginal density of x:

$$\mathcal{A}_m(\delta) = \int v^m(x)f(x-\delta)dx. \tag{A.85}$$

The derivative with respect to δ, at $\delta = 0$, is given by:

$$\begin{aligned}\mathcal{A}_m'(0) &= \frac{\partial}{\partial \delta}\int v^m(x) f(x-\delta)dx\Big|_{\delta=0} = -\int v^m(x) f'(x)dx = -\, v^m(x)f(x)|_{\underline{x}}^{\overline{x}} + \int \frac{d}{dx} v^m(x) f(x) dx \\ &= \int \frac{d}{dx} v^m(x) f(x) dx\end{aligned}$$

where in the third equality we do integration by parts, and in the fourth equality we use the result that there is no mass at the endpoints (or $Pr^{x=\underline{x}}[\tau = 0] = Pr^{x=\overline{x}}[\tau = 0] = 1$). The previous expression says that the effect of the perturbation is equivalent to the changes in the stopping time problem of one agent when her initial conditions change (derivative of v^m with respect to x), averaged across all the possible initial conditions (the steady state distribution). In turn, as we show next, changes in the stopping time problem are reflected by alterations in the state paths and by shifts in duration.

From v^m's definition in (A.83), take its derivative with respect to initial conditions and substitute it back into $\mathcal{A}_m'(0)$

$$\mathcal{A}_m'(0) = \int \frac{\partial}{\partial x} \mathbb{E}^x\left[\int_0^\tau x_t(\omega)^m dt\right] dF(x) - \mathcal{M}_m[x]\int \frac{\partial \mathbb{E}^S[\tau]}{\partial x} dF(x)$$

Lastly, by adding and subtracting the term $\int \mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt\right] dF(x)$, we re-express $\mathcal{A}_m'(0)$ as the sum of three terms Γ_m, Θ_m, and Θ_0 defined in the brackets.

$$\mathcal{A}_m'(0) = \underbrace{\int \mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt\right] dF(x)}_{\mathcal{B}_m} - \mathcal{M}_m[x] \underbrace{\int \frac{\partial \mathbb{E}^x[\tau]}{\partial x} dF(x)}_{\Theta_0} + \underbrace{\int \left(\frac{\partial}{\partial x}\mathbb{E}^x\left[\int_0^\tau x_t^m dt\right] - \mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt\right]\right) dF(x)}_{\mathcal{C}_m}. \quad \text{(A.86)}$$

Now we further characterize each of these terms. Note that for various extensions, the proof up to this point is exactly the same. The results change from this point onwards as we make use of the particular stochastic process for the uncontrolled state.

4. Characterize $\mathcal{B}_m$. Since $x_t = x + \nu t + \sigma W_t$, for all $t \le \tau$ we have that

$$\mathcal{B}_m \equiv \int \mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt\right] dF(x) = \int \mathbb{E}^x\left[\int_0^\tau m x_t^{m-1} dt\right] dF(x).$$

Applying Itō's Lemma to x_t^m we have $dx_t^m = \nu m x_t^{m-1}dt + \sigma m x_t^{m-1} dW_t + \frac{\sigma^2}{2} m(m-1)x_t^{m-2}dt$, and integrating both sides from 0 to τ and taking expectations with initial condition x we get

$$\mathbb{E}^x\left[x_\tau^m\right] - x^m = m\sigma \underbrace{\mathbb{E}^x\left[\int_0^\tau x_t^{m-1} dW_t\right]}_{=0 \text{ by OST}} + \nu \mathbb{E}^x\left[\int_0^\tau m x_t^{m-1} dt\right] + \frac{\sigma^2 m}{2}\mathbb{E}^x\left[\int_0^\tau (m-1) x_t^{m-2} dt\right].$$

Given that $\int_0^t x_s^m dW_t$ is a martingale with zero initial condition and it is well-defined by assumption, we apply the Optional Sampling Theorem (OST) to conclude that $\mathbb{E}^x\left[\int_0^\tau x_t^{m-1} dW_t\right] = 0$. Solve for $\mathbb{E}^x\left[\int_0^\tau m x_t^{m-1} dt\right]$:

$$\mathbb{E}^x\left[\int_0^\tau m x_t^{m-1} dt\right] = \frac{\mathbb{E}^x\left[x_\tau^m\right] - x_t^m}{\nu} - \frac{\sigma^2 m}{2\nu}\mathbb{E}^x\left[\int_0^\tau (m-1)x_t^{m-2} dt\right].$$

Integrating both sides across all initial conditions, defining $\varphi_m^\Gamma \equiv \frac{1}{\nu}\left(\mathbb{E}^x\left[x_\tau^m\right] - x_t^m\right)$ and $\Gamma_m \equiv \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau \varphi_m^\Gamma(x_t)dt\right]}{\mathbb{E}^{\hat{x}}[\tau]}$, and recognizing $\mathcal{B}_m$ and $\mathcal{B}_{m-1}$ we get:

$$\mathcal{B}_m = \Gamma_m - \frac{\sigma^2 m}{2\nu}\mathcal{B}_{m-1}, \qquad \Gamma_0 = 0, \quad \text{(A.87)}$$

where we used the Auxiliary Theorem 2, exchanging the ergodic distribution for the local occupancy measure.

5. Characterize $\mathcal{C}_m$. With similar steps as in the previous point, we characterize $\mathcal{C}_m$ as follows.

$$\mathcal{C}_m \equiv \int \underbrace{\frac{\partial}{\partial x}\mathbb{E}^x\left[\int_0^\tau x_t^m dt\right]}_{A} - \underbrace{\mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt\right]}_{B} dF(x).$$

First we get an expression for the term A. Applying Itō's Lemma to x_t^{m+1} we have $dx_t^{m+1} = (m+1)\nu x_t^m dt + \sigma(m+1)x_t^m dW_t + \frac{\sigma^2}{2} m(m+1)x_t^{m-1}dt$. Integrating both sides from 0 to τ, taking expectations with initial condition x, using the OST, and rearranging we get: $\mathbb{E}^x\left[\int_0^\tau x_t^m dt\right] = \frac{1}{\nu(m+1)}\left(\mathbb{E}^x\left[x_\tau^{m+1}\right] - x_t^{m+1}\right) - \frac{\sigma^2 m}{2\nu}\mathbb{E}^x\left[\int_0^\tau x_t^{m-1} dt\right]$, and its derivative with respect to initial condition x:

$$A \equiv \frac{\partial}{\partial x}\mathbb{E}^x\left[\int_0^\tau x_t^m dt\right] = \frac{1}{\nu}\left(\frac{\partial \mathbb{E}^x\left[x_\tau^{m+1}\right]/(m+1)}{\partial x} - x_t^m\right) - \frac{\sigma^2 m}{2\nu}\frac{\partial \mathbb{E}^x\left[\int_0^\tau x_t^{m-1}dt\right]}{\partial x}$$

Now, for the term B, recall from the characterization of Γ_m that

$$B \equiv \mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^m}{\partial x}dt\right] = \frac{1}{\nu}\left(\mathbb{E}^x\left[x_\tau^m\right] - x_t^m\right) - \frac{\sigma^2 m}{2\nu}\mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^{m-1}}{\partial x}dt\right].$$

Subtract the equations for A and B and simplify to obtain:

$$\frac{\partial \mathbb{E}^x\left[\int_0^\tau x_t^m dt\right]}{\partial x} - \mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^m}{\partial x}dt\right] = \underbrace{\frac{1}{\nu}\left(\frac{\partial \mathbb{E}^x\left[x_\tau^{m+1}\right]/(m+1)}{\partial x} - \mathbb{E}^x\left[x_\tau^m\right]\right)}_{\varphi_m^\Theta(x_t)} - \frac{\sigma^2 m}{2\nu}\left\{\frac{\partial \mathbb{E}^x\left[\int_0^\tau x_t^{m-1}dt\right]}{\partial x} - \mathbb{E}^x\left[\int_0^\tau \frac{\partial x_t^{m-1}}{\partial x}dt\right]\right\}$$

Integrating with the ergodic distribution and using the definition of Θ_m in (A.81) and recognizing $\mathcal{C}_m$ and $\mathcal{C}_{m-1}$ we get:

$$\mathcal{C}_m = \Theta_m - \frac{\sigma^2 m}{2\nu}\mathcal{C}_{m-1}, \qquad \mathcal{C}_{-1} = 0. \tag{A.88}$$

Define $\mathcal{Z}_m \equiv \mathcal{B}_m + \mathcal{C}_m$, which implies $\mathcal{Z}_m = \Gamma_m + \Theta_m - \frac{\sigma^2 m}{2\nu}\mathcal{Z}_{m-1}$. Combine the results in (A.87), (A.87) and (A.88) to obtain: $\mathcal{A}'_m(0) = (\mathcal{Z}_m - \mathcal{M}_m[x]\Theta_0)$.

6. Characterize Θ_0. We corroborate that the expression $\int \frac{\partial \mathbb{E}^x[\tau]}{\partial x}dF(x)$ is equal to Θ_0. By the OST, we have $\mathbb{E}^x[x_\tau] - x = \nu\mathbb{E}^x[\tau]$. Thus $\frac{\partial \mathbb{E}^x[\tau]}{\partial x} = \frac{1}{\nu}\left[\frac{\partial \mathbb{E}^x[x_\tau]}{\partial x} - 1\right]$. Substituting and using Auxiliary Theorem 2 we recover the expression for Θ_0 in the definition of Θ_m:

$$\Theta_0 \equiv \int \frac{\partial \mathbb{E}^x[\tau]}{\partial x}dF(x) = \int \frac{1}{\nu}\left[\frac{\partial \mathbb{E}^x[x_\tau]}{\partial x} - 1\right]dF(x)$$

□

Lemma 9. ***[Representation for intensive margin]*** *The intensive margin Γ_m defined as*

$$\Gamma_m \equiv \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau \varphi^\Gamma(x_t)dt\right]}{\mathbb{E}^{\hat{x}}\left[\tau\right]}, \qquad \text{with} \qquad \varphi_m^\Gamma(x_t) = \frac{1}{\nu}\left(\mathbb{E}^x\left[x_\tau^m\right] - x_t^m\right).$$

can be represented as a function of steady state moments as:

$$\Gamma_m = m\mathcal{M}_{m-1,1}[x,a] + \frac{\mathbb{1}_{\{m\geq 2\}}\sigma^2 m(m-1)}{2\nu}\mathcal{M}_{m-2,1}[x,a].$$

Proof. Start in (1) from the definition of Γ_m and $\varphi_m^\Gamma(S)$, then (2) exchange the time integral with the expectation conditional on adjustment $\mathbb{E}^{\hat{x}}[\cdot]$, which introduces an indicator $\mathbb{1}_{\{t\leq\tau\}}$. Use the law of iterated expectations in (3) to condition on the set $\{t \leq \tau\}$.

$$\begin{aligned}
\nu\mathbb{E}^{\hat{x}}[\tau]\Gamma_m &=^{(1)} \mathbb{E}^{\hat{x}}\left[\int_0^\tau \mathbb{E}^{x_t}\left[x_\tau^m\right] - x_t^m dt\right] =^{(2)} \int_0^\infty \mathbb{E}^{\hat{x}}\left[\left(\mathbb{E}^{x_t}\left[x_\tau^m\right] - x_t^m\right)\mathbb{1}_{\{t\leq\tau\}}\right]dt \\
&=^{(3)} \int_0^\infty \mathbb{E}^{\hat{x}}\left[\mathbb{E}\left\{\left(\mathbb{E}^{x_t}\left[x_\tau^m\right] - x_t^m\right)\mathbb{1}_{\{t\leq\tau\}}\Big| t\leq\tau\right\}\right]dt =^{(4)} \int_0^\infty \mathbb{E}^{\hat{x}}\left[\mathbb{E}\left\{\mathbb{E}^{x_t}\left[x_\tau^m\right] - x_t^m\Big| t\leq\tau\right\}\right]dt \\
&=^{(5)} \int_0^\infty \mathbb{E}^{\hat{x}}\left[\mathbb{E}\left[x_\tau^m - x_t^m\Big| t\leq\tau\right]\right]dt =^{(6)} \int_0^\infty \mathbb{E}^{\hat{x}}\left[\mathbb{E}\left\{\left(x_\tau^m - x_t^m\right)\mathbb{1}_{\{t\leq\tau\}}\Big| t\leq\tau\right\}\right]dt \\
&=^{(7)} \mathbb{E}^{\hat{x}}\left[\int_0^\tau \left(x_\tau^m - x_t^m\right)dt\right] =^{(8)} \mathbb{E}^{\hat{x}}\left[x_\tau^m\int_0^\tau dt\right] - \mathbb{E}^{\hat{x}}\left[\int_0^\tau x_t^m dt\right]
\end{aligned} \tag{A.89}$$

We now characterize $\mathbb{E}^{\hat{x}}\left[x_\tau^{m+1}\int_0^\tau dt\right]$. Applying Ito's lemma followed by the OST to $Y_t^m \equiv x_t^m\int_0^t ds$

$$dY_t^m = x_t^m dt + \mathbb{1}_{\{m\geq 1\}}\nu m x_t^{m-1}\int_0^t ds dt + \mathbb{1}_{\{m\geq 1\}}\sigma m x_t^{m-1}\int_0^t ds dW_t + \mathbb{1}_{\{m\geq 2\}}\frac{m(m-1)\sigma^2}{2}x_t^{m-2}\int_0^t ds dt \tag{A.90}$$

$$\mathbb{E}^{\hat{x}}\left[Y_\tau^m\right] = \mathbb{E}^{\hat{x}}\left[x_\tau^m\int_0^\tau dt\right] = \underbrace{\mathbb{E}^{\hat{x}}\left[\int_0^\tau x_t^m dt\right]}_{\mathcal{M}_m[x]\mathbb{E}^{\hat{x}}[\tau]} + \mathbb{1}_{\{m\geq 1\}}\nu m\underbrace{\frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau x_t^{m-1}tdt\right]}{\mathbb{E}^{\hat{x}}[\tau]}}_{\mathcal{M}_{m-1,1}[x,a]} + \mathbb{1}_{\{m\geq 2\}}\frac{m(m-1)\sigma^2}{2}\underbrace{\mathbb{E}^{\hat{x}}\left[\int_0^\tau x_t^{m-2}tdt\right]}_{\mathcal{M}_{m-2,1}[x,a]} \tag{A.91}$$

Using equations (A.89) and (A.91), we have that

$$\Gamma_m = \mathbb{1}_{\{m\geq 1\}} m \mathcal{M}_{m-1,1}[x,a] + \frac{\mathbb{1}_{\{m\geq 2\}}\sigma^2 m(m-1)}{2\nu}\mathcal{M}_{m-2,1}[x,a]$$

□

Lemma 10. ***[Representation for extensive margin]*** *Assume the moments of the adjustment size can be written as:*

$$g_m(x) = \mathbb{E}^{\hat{x}+x}\left[(\hat{x}-\Delta x)^m\right] - \mathbb{E}^{\hat{x}}\left[(\hat{x}-\Delta x + x)^m\right] \tag{A.92}$$

Then the extensive margin given by $\Theta_m \equiv \frac{\mathbb{E}^{\hat{x}}\left[\int_0^\tau \varphi_m^\Theta(x_t)dt\right]}{\mathbb{E}^{\hat{x}}[\tau]}$ *with* $\varphi_m^\Theta(x) \equiv \frac{1}{\nu}\left(\frac{\partial \mathbb{E}^x\left[x_\tau^{m+1}/(m+1)\right]}{\partial x} - \mathbb{E}^x\left[x_\tau^m\right]\right)$ *can be represented as a function of steady state moments as follows:*

$$\Theta_m = \sum_{j=0}^{\infty}\theta_{m,j}\mathcal{M}_j[x] \quad \textit{with} \quad \theta_{m,j} = \sum_{k\geq j}^{\infty}\frac{1}{\nu}\frac{\left[\frac{d^{k+1}g_{m+1}(x)}{dx^{k+1}}/m+1 - \frac{d^k g_m(x)}{dx^k}\right]\Big|_{x=0}\hat{x}^{k-j}}{k!j!}. \tag{A.93}$$

- *If* $\tau|x_t \sim \tau$, $g_m(x) = \theta(m,j) = 0$ *for all* m, i.

Proof. Using a change of variable in assumption (A.92), we have that :

$$\mathbb{E}^y\left[x_\tau^m\right] = g_m(y-\hat{x}) + \mathbb{E}^{\hat{x}}\left[(y-\hat{x}+x_\tau)^m\right]. \tag{A.94}$$

Using the previous equation we have that

$$\begin{aligned}
\Theta_m &= \frac{\mathbb{E}^{\hat{S}}\left[\int_0^\tau\left(\frac{\partial \mathbb{E}^S[x_\tau^{m+1}/(m+1)]}{\partial x} - \mathbb{E}^S[x_\tau^m]\right)dt\right]}{\nu\mathbb{E}^{\hat{S}}[\tau]} && \text{(A.95)}\\
&= \frac{1}{\nu}\frac{\mathbb{E}^{\hat{S}}\left[\int_0^\tau\left[\left[\frac{dg_{m+1}(y-\hat{x})}{dy}/(m+1) + \mathbb{E}^{\hat{x}}\left[(y-\hat{x}+x_\tau)^{m+1}\right]\right] - \left[g_m(y-\hat{x}) + \mathbb{E}^{\hat{x}}\left[(y-\hat{x}+x_\tau)^m\right]\right]\right]dt\right]}{\mathbb{E}^{\hat{S}}[\tau]}\\
&= \frac{1}{\nu}\frac{\mathbb{E}^{\hat{S}}\left[\int_0^\tau\left[\frac{dg_{m+1}(y-\hat{x})}{dy}/(m+1) - g_m(y-\hat{x})\right]dt\right]}{\mathbb{E}^{\hat{S}}[\tau]}\\
&= \frac{1}{\nu}\frac{\mathbb{E}^{\hat{S}}\left[\int_0^\tau\left[\sum_{j=0}^\infty \frac{d^j}{j!dx^j}\left[\frac{dg_{m+1}(x)}{dx}/(m+1) - g_m(x)\right]_{x=0}(y-\hat{x})^j\right]dt\right]}{\mathbb{E}^{\hat{S}}[\tau]}\\
&= \frac{1}{\nu}\sum_{j=0}^{\infty}\sum_{z=0}^{j}\frac{\left[\frac{d^{j+1}g_{m+1}(0)}{dx^{j+1}}/m+1 - \frac{d^j g_m(0)}{dx^j}\right]\hat{x}^z}{j!}\begin{pmatrix} j\\ z\end{pmatrix}\frac{\mathbb{E}^{\hat{S}}\left[\int_0^\tau x_t^{j-z}dt\right]}{\mathbb{E}^{\hat{S}}[\tau]}\\
&= \frac{1}{\nu}\sum_{j=0}^{\infty}\sum_{z=0}^{j}\frac{\left[\frac{d^{j+1}g_{m+1}(0)}{dx^{j+1}}/m+1 - \frac{d^j g_m(0)}{dx^j}\right]\hat{x}^z}{j!}\begin{pmatrix} j\\ z\end{pmatrix}\mathcal{M}_{j-z}[x]\\
&= \sum_{j=0}^{\infty}\sum_{z=0}^{j}\underbrace{\frac{1}{\nu}\frac{\left[\frac{d^{j+1}g_{m+1}(0)}{dx^{j+1}}/m+1 - \frac{d^j g_m(0)}{dx^j}\right]\hat{x}^z}{z!(j-z)!}}_{=\mathcal{H}_{m,j,z}}\mathcal{M}_{j-z}[x]\\
&= \sum_{h=0}^{\infty}\theta_{m,h}\mathcal{M}_h[x], \ \textit{with}\ \theta_{m,h} = \sum_{k\geq h}^{\infty}\mathcal{H}_{m,k,k-h}. && \text{(A.96)}
\end{aligned}$$

If $\tau|x_t \sim \tau$, then we have that

$$\begin{aligned}
g_m(x) &= \mathbb{E}^{\hat{x}+x}\left[(\hat{x}-\Delta x)^m\right] - \mathbb{E}^{\hat{x}}\left[(\hat{x}-\Delta x+x)^m\right]\\
&= \mathbb{E}\left[(\hat{x}+x+\nu\tau^{x+\hat{x}}+\sigma W_{\tau^{x+\hat{x}}})^m\right] - \mathbb{E}\left[(\hat{x}+x+\nu\tau^{\hat{x}}+\sigma W_{\tau^{\hat{x}}})^m\right]\\
&= \mathbb{E}\left[(\hat{x}+x+\nu\tau+\sigma W_\tau)^m\right] - \mathbb{E}\left[(\hat{x}+x+\nu\tau+\sigma W_\tau)^m\right]\\
&= 0 && \text{(A.97)}
\end{aligned}$$

□