# Optimal Mortgage Refinancing with Inattention*

David Berger† Konstantin Milbradt‡ Fabrice Tourre§
Joseph Vavra¶

September 2024

**Abstract**

We build a model of optimal fixed-rate mortgage refinancing with fixed costs and inattention and derive a new sufficient statistic that can be used to measure inattention frictions from simple moments of the rate gap distribution. In the model, borrowers pay attention to rates sporadically so they often fail to refinance even when it is profitable. When paying attention, borrowers optimally choose to refinance earlier than under a perfect attention benchmark. Our model can rationalize almost all errors of "omission" (refinancing too slowly) and a large fraction of the errors of "commission" (refinancing too quickly) previously documented in the data.

**Keywords**:

**JEL codes**:

*We would like to thank John Driscoll for helpful comments. Fabrice Tourre acknowledges financial support from the Bert and Sandra Wasserman endowment.

†Duke University and NBER; david.berger@duke.edu

‡Northwestern University and NBER; milbradt@northwestern.edu

§Baruch College; fabrice.tourre@baruch.cuny.edu

¶University of Chicago and NBER; joseph.vavra@chicagobooth.edu

# 1 Introduction

A large body of evidence suggests that inattention leads many borrowers to miss out on mortgage refinancing opportunities that could save them significant amounts of money. Without a benchmark model of optimal refinancing, it is difficult to know how sub-optimal this behavior is. In this paper, we characterize the impact of inattention on fixed-rate mortgage borrowers' optimal refinancing decisions and show that inattention is important for explaining refinancing patterns observed in the data. In our model, borrowers pay attention to rates sporadically, which means they often fail to refinance even when their "rate gap" — the difference between their current rate and the available rate on a new mortgage — is large. However, when they do pay attention, they optimally choose to refinance for smaller rate gaps than in a model with perfect attention. Thus, inattention can explain both errors of "omission" (refinancing too slowly) and errors of "commission" (refinancing too quickly) documented in the data.[1]

We begin with an analysis of borrowers' optimal decisions when mortgage rates follow an Itô process. Borrowers have a fixed-rate prepayable mortgage contract that can be refinanced at any time, but their cost-minimization objective is hindered by two different frictions: (i) whenever they refinance, they have to pay a fixed cost, and (ii) they only pay attention to the market sporadically. Borrowers optimally refinance when they pay attention and their rate gap is above a threshold. We show that greater inattention systematically lowers this threshold. While we prove this relationship holds under a general class of mortgage rate processes, assessing its magnitude requires further assumptions.

For most of the paper, we assume the mortgage rate follows a Brownian motion, which allows us to derive analytical solutions for optimal refinancing. With this assumption, our model mirrors Agarwal, Driscoll and Laibson (2013) (thereafter, "ADL"), except our borrowers are inattentive and make decisions at discrete points in time. In this Brownian motion case, we calculate the rate gap threshold for optimal refinancing. When borrowers are infinitely attentive, we recover the ADL threshold formula. However, this threshold declines significantly with inattention, and this has important consequences for empirical estimates of refinancing mistakes.

This key result remains under two important extensions. First, although our analytical results assume mortgage rates follow a Brownian motion, similar conclusions hold when rates are mean-reverting. Second, we micro-found inattention by considering a

[1] Agarwal, Rosen and Yao (2016) find that borrowers refinance at rate gaps on average 37 bps too low relative to the theoretically optimal value (errors of commission), and that refinancing decisions occur on average 3 months too late — with 17% of borrowers more than 6 months late — relative to the theoretically optimal refinancing time (errors of omission).

perfectly attentive borrower who incurs a cost each time they check the current mortgage rate. This rational inattention framework yields an optimal threshold that increases with the effective attention rate, consistent with our baseline model where inattention is exogenous.

We then analytically derive our baseline model's implications for the ergodic distribution of rate gaps and rate gaps upon refinancing. We show that the difference in the average of these distributions only depends on 3 parameters: (i) the moving rate, (ii) the attention rate, and (iii) mortgage rate volatility. This yields a sufficient statistic approach to recover the implied attention rate in the data.

Finally, we bring our model to the data to assess the impact of inattention on optimal refinancing. Our preferred estimate for the attention intensity, which best rationalizes our micro-data on rate gaps and prepayments, is 23.4% per year. Using this estimate, we compute borrowers' rate gap thresholds with inattention and quantify the frequency of errors of commission and omission. Compared to the ADL framework, which assumes perfect attention, our model nearly eliminates errors of omission and reduces errors of commission by 32-43%.

# 2 Literature

Our paper builds on Agarwal, Driscoll and Laibson (2013), which derived the optimal rate gap threshold for refinancing long-term fixed-rate mortgages. We extend their model by incorporating inattention, inspired by Calvo (1983), and demonstrate that this friction lowers the optimal refinancing threshold. Additionally, we explore various mortgage rate processes and analyze how inattention affects the distribution of rate gaps.[2] Our focus on inattention builds in part on Andersen et al. (2020). Using Danish data, they estimate a model of mortgage refinancing featuring both state-dependent (arising from fixed costs) and time-dependent inaction (arising from inattention), but their analysis is primarily statistical rather than building from optimizing households.

Our modeling environment focuses on the same frictions as Berger et al. (2024) but very different implications. Our paper analytically characterizes refinancing decisions for *individual* households and resulting implications for empirical rate gaps. Berger et al. (2024) instead studies implications of *heterogeneity* for mortgage market *equilibrium*.

[2] An older version of their paper also included some overlapping conclusions related to inattention. We thank John Driscoll for alerting us to and sharing those results, which do not appear in the published paper. Our results also differ in some important ways from this unpublished material, as discussed in the main text.

Many empirical investigations of mortgage refinancing decisions document some type of mistakes by borrowers (e.g. (Keys, Pope and Pope, 2016)). Agarwal, Rosen and Yao (2016) separate these mistakes into (i) *errors of commission* (refinancing at the "wrong" rate gap) and (ii) *errors of omissions* (not immediately refinancing once the rate gap threshold is reached). We show that our framework can rationalize both types of errors.

The optimal prepayment strategy has also been widely studied in the context of valuing mortgage-backed securities ("MBS"). Stanton (1995) is most relevant to our work, as it uses aggregate MBS prepayment data to estimate a model with heterogeneous fixed refinancing costs and homogeneous inattention. The study finds that borrowers face varying fixed costs and show significant time-dependent inaction. In contrast, our study derives an analytic formula for the optimal refinancing strategy and uses mortgage-level micro-data to estimate inattention.

Our paper also connects to inattention in option exercise beyond mortgage refinancing. For instance, Kadan, Liu and Yang (2009) show that when executives are inattentive, the optimal exercise barrier of their stock options rises with attention so that agents should similarly exercise their options at lower intrinsic values when given the opportunity.

# 3 Mortgage refinancing with inattention: general case

We consider a model of mortgage refinancing decisions similar to that in Berger et al. (2024). We study fixed-rate mortgages that can be refinanced at any time — a contract chosen by the majority of borrowers in the US.

## 3.1 Setup

Time $t$ is continuous. A risk-neutral, long-lived borrower with discount rate $\rho$ has financed the purchase of a house with a long-term fixed-rate prepayable mortgage with coupon $c_t$ and constant unit balance. Let $m_t$ be the prevailing mortgage rate, i.e., the rate that can be locked in when refinancing at time $t$. We assume that $m_t$ is a diffusion with drift $\mu(m_t)$, volatility $\sigma(m_t)$ and infinitesimal generator $\mathcal{L}$.[3] Two separate frictions limit the borrower's refinancing ability. First, the borrower is inattentive and makes decisions

[3] $\mathcal{L}$ is defined over functions $f$ of class $\mathcal{C}^2$ via $\mathcal{L}f(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)$. We assume that $\mu$ and $\sigma$ are Lipschitz-continuous, that $\sigma$ is strictly positive, and that $m_t$ evolves on a bounded domain $[\underline{m}, \bar{m}]$, where $\underline{m}$ (resp. $\bar{m}$) can be arbitrarily small (resp. large).

only at discrete times, modeled as i.i.d. Poisson events occurring with intensity $\lambda$ — the *attention rate*. Second, the borrower bears closing costs $\psi$ when refinancing. Last, the borrower moves from one house to another at intensity $\nu$; when doing so, the mortgage coupon gets reset to the prevailing mortgage rate.[4] For simplicity we rule out the ability for the borrower to either default or cash-out refinance.[5]

In our model, inattention represents the various frictions borrowers face when making refinancing decisions, leading to time-dependent inaction. While it can be interpreted literally, it may also arise from costs associated with obtaining mortgage quotes, as noted in Section 5.2, or from financial frictions, such as the need for employment to obtain or refinance a mortgage in the US.[6]

## 3.2 The role of inattention

Denote $V(m, c)$ the valuation of all future mortgage liabilities for a borrower paying a coupon $c$, when the prevailing mortgage rate is $m$. The borrower solves

$$V(m,c) := \inf_{a\in\mathcal{A}} \mathbb{E}_{m,c}\left[\int_0^{+\infty} e^{-\rho t}\left(c_t^{(a)}dt + a_t\psi dN_t^{(\lambda)}\right)\right], \tag{1}$$
$$\text{s.t.} \qquad dc_t^{(a)} = \left(m_t - c_{t-}^{(a)}\right)\left(a_t dN_t^{(\lambda)} + dN_t^{(\nu)}\right),$$

where $\mathcal{A}$ is a set of progressively measurable binary actions $a = \{a_t\}_{t\geq 0}$ such that $a_t \in \{0,1\}$ at all times, $N_t^{(\lambda)}$ and $N_t^{(\nu)}$ are counting processes with jump intensity $\lambda$ and $\nu$, $c_t^{(a)}$ is the coupon rate on the mortgage for a borrower following strategy $a$, and the subscript on the expectation indicates the initial value of the state variables. At the random points in time when the borrower pays attention, the choice $a_t = 1$ represents a decision to refinance, while $a_t = 0$ means that the borrower chooses to keep their existing mortgage. In Appendix A.1, we establish the following:

**Proposition 1** *$V$ is continuous and strictly increasing in $c$. Borrowers refinance whenever they pay attention and the coupon is $\theta(m)$ above the prevailing mortgage rate. The threshold $\theta(m)$*

[4] One could assume moving-related fixed costs without changing our conclusions. The moving rate $\nu$ can also encapsulate exponential amortization; in that case, $\nu$ is interpreted as the sum of the moving and amortization rates.

[5] We ignore defaults since we view our setting as best suited to study the agency mortgage market, in which default rates are negligible, relative to prepayment rates. Since borrowers are risk-neutral, they have no liquidity-related motives to cash-out refinance; refinancing decisions are only driven by interest cost reduction considerations.

[6] It is straightforward to show that borrowers facing unemployment risk act *as if* they have greater inattention w.r.t. refinancing decisions.

*satisfies the indifference condition*

$$V(m,m)+\psi = V(m, m+\theta(m)). \tag{2}$$

According to (2), whenever the current mortgage coupon is $\theta(m)$ above the current mortgage rate $m$, the borrower is indifferent between (i) refinancing the mortgage, paying the fixed cost $\psi$ and resetting the coupon to the current rate $m$, and (ii) staying put. The Hamilton–Jacobi–Bellman (HJB) equation associated with problem (1) is

$$(\rho+\nu+\lambda)\,V(m,c) = c + \nu V(m,m) + \lambda \min\left\{V(m,c), V(m,m)+\psi\right\} + \mathcal{L}V(m,c). \tag{3}$$

If there exists a function $V$ twice differentiable in $m$ that satisfies (3), then $V$ must be the solution to problem (1).[7] In our next proposition (proven in Appendix A.2), we study the extent to which the threshold $\theta(m)$ varies with the degree of borrower inattention.

**Proposition 2** *Assume there exists a function $V$ twice differentiable in $m$ and once differentiable in the parameter $\lambda$ that satisfies (3). The threshold $\theta(m)$ is an increasing function of $\lambda$.*

Inattention leads borrowers to refinance at smaller rate gaps when they pay attention. Because they check rates infrequently, inattentive borrowers "pull the trigger" earlier after observing rates since they might not get another chance soon.

# 4 Mortgage rates as a random walk

Proposition 2 establishes a general result that relies only on weak assumptions on the statistical properties of the mortgage rate. In this section, we characterize analytically the threshold $\theta(m)$ in the special case where mortgage rates follow a random walk.

**Proposition 3** *Assume that $m_t$ is a Brownian motion with volatility $\sigma$. Introduce the constants $\eta_0, \eta_\lambda$ and $\epsilon_\lambda$, equal to:*

$$\eta_0 := \frac{\sqrt{2\,(\rho+\nu)}}{\sigma} \qquad \eta_\lambda := \frac{\sqrt{2\,(\rho+\nu+\lambda)}}{\sigma} \qquad \epsilon_\lambda := \frac{(\rho+\nu)\,(\eta_0+\eta_\lambda)}{\lambda}.$$

[7] This so-called verification theorem is standard and is thus omitted here. For more details, see Fleming and Soner (2006).

*The borrower valuation satisfies* $V(m,c) = \frac{c}{\rho} + v(z)$*, with* $z := c - m$ *and*

$$v(z) = \begin{cases} k_- e^{\eta_0 (z-\theta)} + \frac{\nu}{\rho+\nu}\left[v(0) - \frac{z}{\rho}\right] & \text{if } z \le \theta \\ k_+ e^{-\eta_\lambda (z-\theta)} + \frac{\nu}{\rho+\nu+\lambda}\left[v(0) - \frac{z}{\rho}\right] + \frac{\lambda}{\rho+\nu+\lambda}\left[v(0) + \psi - \frac{z}{\rho}\right] & \text{if } z \ge \theta. \end{cases} \tag{4}$$

*The constants of integration* $k_-, k_+$ *are given in Appendix A.3. The (state-independent) rate gap threshold* $\theta > 0$ *is given by*

$$\theta = (\rho+\nu)\,\psi + \frac{1}{\eta_0 + \epsilon_\lambda} + \frac{1}{\eta_0} W\left(\frac{-\eta_0}{\eta_0+\epsilon_\lambda}\exp\left\{\frac{-\eta_0}{\eta_0+\epsilon_\lambda}\left[1 + (\rho+\nu)\,(\eta_0+\epsilon_\lambda)\,\psi\right]\right\}\right), \tag{5}$$

*where W is the Lambert-W function.* $\theta$ *increases with the attention rate* $\lambda$*, and asymptotically:*

$$\lim_{\lambda\to 0} \theta = (\rho+\nu)\psi \tag{6}$$

$$\lim_{\lambda\to+\infty} \theta = (\rho+\nu)\psi + \frac{1}{\eta_0} + \frac{1}{\eta_0} W\left(-\exp\left\{-\left[1+(\rho+\nu)\eta_0\psi\right]\right\}\right). \tag{7}$$

*A Taylor expansion of the implicit equation underlying (5) around* $\theta = 0$ *yields an approximation*[8] *and lower bound* $\hat{\theta}$ *of the threshold* $\theta$ *with formula:*

$$\hat{\theta} = \sqrt{\frac{2}{\eta_0}\left(1 + \frac{\epsilon_\lambda}{\eta_0}\right)(\rho+\nu)\,\psi + \left(\frac{\epsilon_\lambda}{\eta_0^2}\right)^2} - \frac{\epsilon_\lambda}{\eta_0^2}. \tag{8}$$

Our proof in Appendix A.3 relies on the observation that the value function can be decomposed into (i) the present value of all future interest payments $c/\rho$ (based on the current mortgage coupon) plus (ii) the value of a refinancing option which, given the unit root behavior of mortgage rates, only depends on the rate gap $z$.

Proposition 3 generalizes the results of Agarwal, Driscoll and Laibson (2013) to the case where borrowers are inattentive. Formula (6) suggests that a completely inattentive borrower should only refinance when its rate gap is above the flow value of the fixed cost $(\rho+\nu)\psi$. At the other extreme, if the borrower is infinitely attentive, the threshold reduces to the ADL formula (7). Importantly, a decrease in $\lambda$ reduces the threshold — a special case of the more general result of Proposition 2. Figure 1a illustrates how the threshold $\theta$ given in (5) varies with attention $\lambda$, and also the degree of precision of our approximation formula (8).

Given our estimated borrower attention (see Section 7) and parameter values consis-

[8] Sensitivity analysis around our baseline parameters suggest that this approximation is quite accurate.

tent with Figure 1a, the rate gap threshold with inattention is 54-65% smaller than that in the perfect attention benchmark, depending on the assumed mortgage rate volatility.[9]

Our analysis does not dispute the existence of refinancing mistakes, but partially explains patterns that might otherwise be seen as errors. Agarwal, Rosen and Yao (2016) and Fuster et al. (2019) suggest that borrowers refinance at rate gaps smaller than the ADL threshold, labeled as "errors of commission". However, when accounting for inattention, borrowers' rate gap thresholds drop, making their refinancing decisions appear rational rather than mistaken. Similarly, both articles note that borrowers delay refinancing even after reaching the optimal threshold. Inattention leads borrowers to delay refinancing by random amounts of time and thus naturally rationalizes these "errors of omission". In Section 7, we explore how our model explains these errors. Alternative models, featuring inattentive but also "naïve" households — maybe they ignore the option value, or maybe they ignore the presence of fixed costs when refinancing[10] — might also lower the rate gap threshold and slow prepayment rates, but struggle to match micro-data on rate gaps and prepayments, as discussed in Section 7.

# 5 Extensions

In this section, we investigate two extensions of our benchmark model. In the first, we show that our results are quantitatively insensitive to our assumption that mortgage rates follow a random walk. In the second, we micro-found inattention by introducing observation costs, and show that the key insight of Section 4 remains.

## 5.1 Mean-reverting mortgage rates

Section 4 relies on the assumption that the mortgage rate $m_t$ is a random walk — a strongly debated empirical question.[11] When rates are mean-reverting, refinancing decisions are state dependent, i.e., $\theta = \theta(m)$. To what extent does mean reversion alter borrowers' thresholds? To address this question we analyze mortgage rates that follow

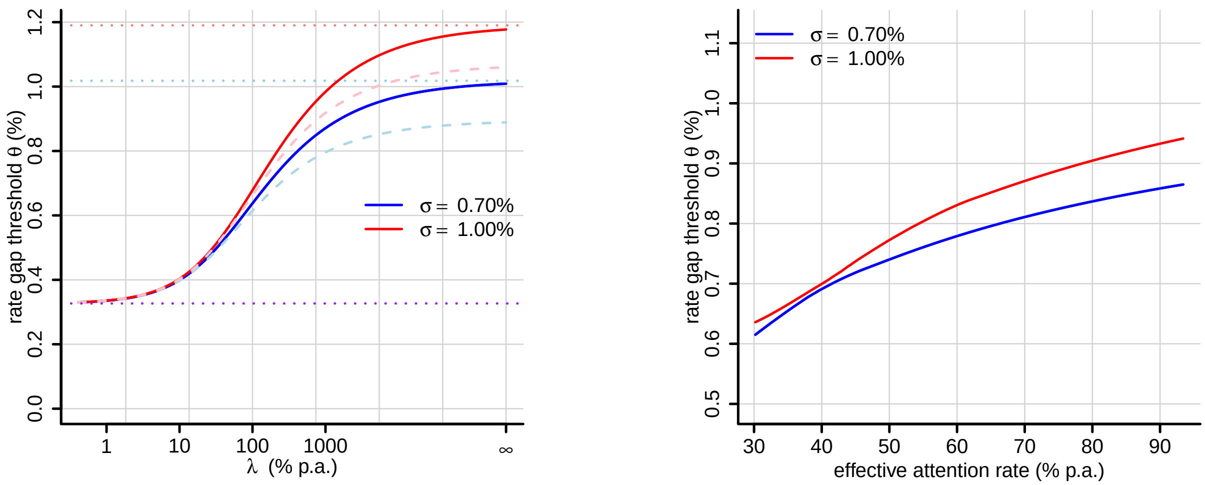
[9] Using our preferred value $\sigma = 0.70\%$, Section 7 estimates an attention rate $\lambda = 23.4\%$ p.a., which leads to a threshold $\theta = 0.46\%$, which is 54% smaller than the relevant ADL threshold; assuming instead $\sigma = 1.00\%$ (used by Agarwal, Driscoll and Laibson (2013) and others), Section 7 estimates an attention rate $\lambda = 11.8\%$ p.a., which leads to a threshold $\theta = 0.41\%$, which is 65% smaller than the relevant ADL threshold.

[10] In the first case, borrowers would use $\theta = (\rho + \nu)\psi$; in the second case, they would use $\theta = 0$.

[11] Stock and Watson (1988) conclude that various short term interest rates appear to contain a unit root; Perron (1989) cannot reject the unit root hypothesis at a 10% level, but can reject at lower levels; more recently, Bierens (1997) conducts various tests, some of them rejecting, and others failing to reject the unit root hypothesis.

**Figure 1:** Rate gap threshold $\theta$

**(a):** Benchmark model **(b):** Rational inattention model




Left-hand side: solid (dashed) lines show the threshold $\theta$ (approximation $\hat{\theta}$) as a function of attention $\lambda$ in the benchmark model of Section 4. The horizontal dotted purple line shows the limit of $\theta$ when $\lambda \to 0$ and the dotted red and blue lines show the limit as $\lambda \to +\infty$. Right-hand side: threshold $\theta$ as a function of the effective attention rate, defined as $1/\mathcal{T}$, with $\mathcal{T}$ the ergodic average observation delay, when varying the observation costs $\phi$, in the rational inattention model of Section 5.2. Figures computed assuming $\rho = 5\%$, $\nu = 11.33\%$, $\psi = 2\%$ and for two different mortgage rate volatilities $\sigma = 0.70\%$ and $\sigma = 1.00\%$.

an Ornstein-Uhlenbeck process:

$$dm_t = -\chi(m_t - \bar{m})dt + \sigma dB_t \tag{9}$$

This specification nests the special case of the Brownian motion studied in Section 4, by setting $\chi = 0$.

We numerically solve our model using baseline parameters, varying the mortgage rate's half-life from 1 to 20 years. The optimal threshold $\theta(m)$ increases with $m$, falling below or above the benchmark value (5) when rates are low or high, respectively. The ergodic average deviates by no more than 2 bps from the benchmark value (5), with a standard deviation of 3-9 bps (resp. 2-5 bps) when $\sigma = 0.70\%$ (resp. $\sigma = 1.00\%$) for the range of half-lives considered.

Thus, our conclusions are similar when using stationary mortgage rates instead of assuming a random walk like in our baseline model, especially for persistent processes as is the case in the data.[12] We maintain the simpler random walk assumption for the

[12]Using the sample period consistent with that used in Section 7, we estimate a half-life of 3.9 years for

remainder of the paper.

### 5.2 Rational inattention

Next, consider an environment where borrowers make continuous decisions but incur a cost $\phi$ to check the current mortgage rate and a cost $\psi$ to refinance. Each time they observe the market, they decide whether to refinance and when to check the market again. The refinancing option value depends on the rate gap $z$, and the optimal strategy remains a cutoff — refinance when $z \geq \theta$, with an optimally chosen threshold $\theta$. Unlike our benchmark model, borrowers now choose their next observation time $\mathbb{T}(z)$ based on the current rate gap, balancing the cost of attention with the potential future benefit of refinancing. This model extension delivers rational inattention, as in Abel, Eberly and Panageas (2007).

Appendix A.5 details the mathematical notation for this model extension. We numerically solve our model to determine the optimal refinancing rate gap $\theta$ and the ergodic average attention delay $\mathcal{T}$ across different observation costs $\phi$. Both $\theta$ and $\mathcal{T}$ increase with observation costs, creating an upward sloping relationship between $\theta$ and the effective attention rate $1/\mathcal{T}$, as Figure 1b illustrates, thus echoing our insight from previous sections. This model supports our preferred estimated attention rate of 23.4% (Section 7) if observation costs are 0.50% of mortgage balances.[13]

## 6 Distributional implications

We now analyze the benchmark model's implications from Section 4 for the distribution of rate gaps and rate gaps at refinancing. These distributions are measurable in the data and have been central to the household finance literature, which assesses the magnitude and frequency of refinancing mistakes. Under the assumptions of Proposition 3, although the mortgage rate is non-stationary, the rate gap $z_t$ admits an ergodic density $f(z)$:

$$f(z)dz := \lim_{t \to +\infty} \frac{1}{t} \int_0^t \mathbb{1}_{\{z_t \in [z, z+dz]\}} dt. \tag{10}$$

The density $f$ can be characterized analytically, as the next proposition shows.

30-year mortgage rates.

[13] While high, as before this implicitly reflects pure observation costs of obtaining a quote but also various financial frictions that limit borrowers' refinancing opportunities.

**Proposition 4** *Assume that $m_t$ is a Brownian motion with volatility $\sigma$. Introduce the constants $\chi_0$ and $\chi_\lambda$ equal to:*

$$\chi_0 := \frac{\sqrt{2\nu}}{\sigma} \qquad \chi_\lambda := \frac{\sqrt{2\left(\nu+\lambda\right)}}{\sigma}.$$

*The rate gap $z_t$ admits an ergodic density $f$ that is an asymmetric Laplace distribution:*

$$f(z) = \frac{1}{\frac{1}{\chi_0}+\frac{1}{\chi_\lambda}} \begin{cases} \exp\left(\chi_0(z-\theta)\right) & \text{if } z \leq \theta \\ \exp\left(-\chi_\lambda(z-\theta)\right) & \text{if } z \geq \theta \end{cases} \tag{11}$$

*Instead, the rate gap upon prepayment admits an ergodic density $\hat{f}$ equal to*

$$\hat{f}(z) = \frac{1}{\left(\frac{1}{\chi_0}+\frac{1}{\chi_\lambda}\right)\nu + \frac{\lambda}{\chi_\lambda}} \begin{cases} \nu\exp\left(\chi_0(z-\theta)\right) & \text{if } z \leq \theta \\ (\nu+\lambda)\exp\left(-\chi_\lambda(z-\theta)\right) & \text{if } z \geq \theta \end{cases} \tag{12}$$

*The following ergodic statistics admit analytic expressions*

$$\text{average prepayment rate} = \lim_{t\to+\infty}\frac{1}{t}\left(N_t^{(\nu)} + \int_0^t a_s dN_s^{(\lambda)}\right) = \nu + \left(\frac{\chi_0}{\chi_0+\chi_\lambda}\right)\lambda \tag{13}$$

$$\text{average rate gap} = \lim_{t\to+\infty}\frac{1}{t}\int_0^t z_s ds = \theta + \frac{1}{\chi_\lambda} - \frac{1}{\chi_0} \tag{14}$$

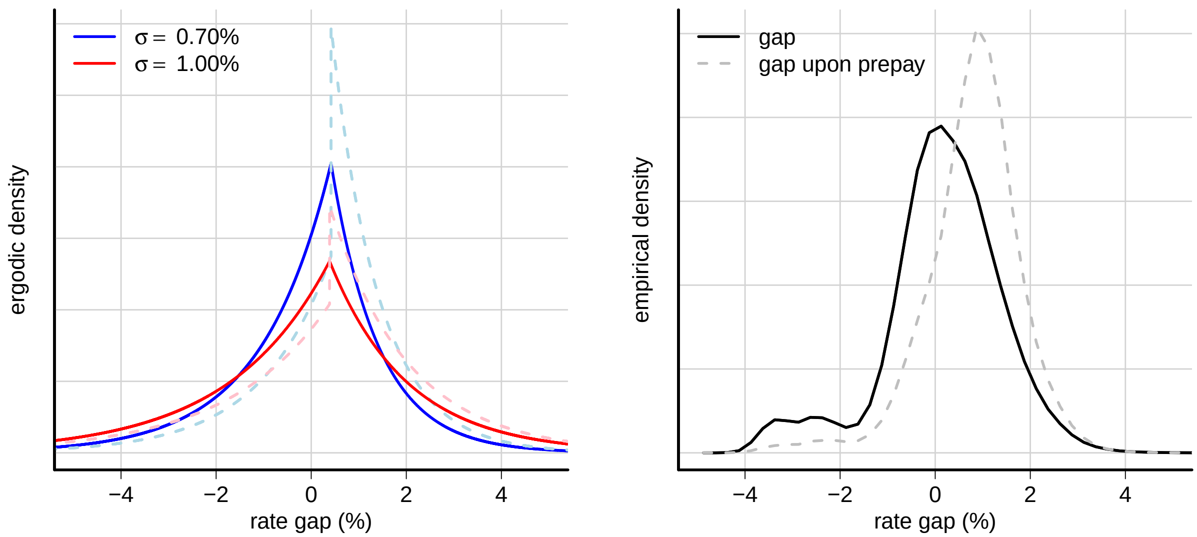
$$\text{average rate gap upon prepay} = \lim_{t\to+\infty}\frac{\int_0^t z_{s-}\left(dN_s^{(\nu)} + a_s dN_s^{(\lambda)}\right)}{N_t^{(\nu)} + \int_0^t a_s dN_s^{(\lambda)}} = \theta \tag{15}$$

See proof in Appendix A.4. Figure 2 compares model-implied densities with empirical data from a sample of 30-year fixed-rate mortgages from Fannie Mae's Single-Family Loan Performance ("SFLP") data. Formula (13) indicates the long-term average prepayment rate falls between $\nu$ and $\nu+\lambda$, while formula (14) suggests that the average rate gap is below the threshold $\theta$ — both expected outcomes. However, formula (15) reveals a surprising result: despite borrowers' strategy to wait until the rate gap exceeds $\theta$ before refinancing, the ergodic average rate gap upon prepayment equals $\theta$.[14] This occurs because when the rate gap is below $\theta$, borrowers prepay at a low intensity $\nu$, lowering the overall average. This result also implies that the average rate gap upon prepayment increases with the attention rate.

The statistics (14) and (15) are measurable in the data. Since they both increase one-

[14] Using a similar approach, one can show that the average rate gap upon refinancing (rather than prepayment, i.e. excluding exogenous move-related prepayments) is equal to $\theta + 1/\chi_\lambda$, thus strictly above the threshold $\theta$.

**Figure 2:** Rate gap density: theory vs data.




Theoretical (left-hand side) and empirical (right-hand side) rate gap distribution $f(z)$ (solid lines) and rate gap upon prepayment distribution $\hat{f}(z)$ (in dash lines) for two mortgage rate volatilities. The left-hand side assumes that $m_t$ is a Brownian motion, and is computed using parameters $\rho = 5\%$, $\nu = 11.33\%$, $\psi = 2\%$ and attention rates estimated in Section 7.

for-one with $\theta$, their difference is independent of this threshold and only depends on (i) the moving rate $\nu$, (ii) attention rate $\lambda$, and (iii) mortgage rate volatility $\sigma$:

$$(\textit{avg rate gap upon prepay}) - (\textit{avg rate gap}) = \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\sqrt{\nu}} - \frac{1}{\sqrt{\nu+\lambda}}\right) \tag{16}$$

Thus, conditional on two moments measurable in the data and assumptions on $\nu$ and $\sigma$, one can invert (16) to recover the effective attention rate $\lambda$ of a population of interest, irrespective of the values of the subjective discount rate $\rho$ or the closing costs $\psi$. We use this approach in Section 7 to estimate the average attention rate in the population of agency mortgage borrowers using SFLP data.[15] Moments (13)-(15), as well as the model-implied prepayment hazard, could also be used in a GMM estimation of model parameters, as well as in tests of over-identifying restrictions. More heuristically, our model could be rejected if we were to observe borrower search behavior and still detect errors of both omission and commission amongst borrowers actively searching. We leave these alternative approaches for future work.

[15] It is important to emphasize that the distributions $f$ and $\hat{f}$ are *long run* rate gap distributions, which would be obtained if we follow one borrower for an arbitrarily long time. In other words, $f$ does *not* characterize the cross-sectional distribution of rate gaps at a particular point in time, but rather the *average* cross-sectional distribution of rate gaps over an arbitrarily long time-horizon.

# 7 Empirical study

Many studies (Keys, Pope and Pope, 2016; Agarwal, Rosen and Yao, 2016; Fuster et al., 2019) use the ADL formula to assess the extent of borrowers' refinancing mistakes. Based on parameter $\psi, \nu, \sigma$ and $\rho$, they determine the optimal refinancing threshold. Borrowers with rate gaps above this threshold are considered to have made "errors of omission", while those refinancing below it are seen as making "errors of comission". Our model can rationalize both type of errors: inattentive borrowers make decisions at discrete *points* in time, often resulting in observed rate gaps above the threshold $\theta$; and inattention lowers $\theta$, thereby reducing the number of errors of commission that an econometrician would measure in the data.

To illustrate our point, we use a 2% random sample of 30-year fixed rate mortgages originated between January 1999 and September 2023, for which we have origination and performance data from the SFLP dataset. For each mortgage-month observation, we compute the mortgage rate available to borrowers refinancing at such time.[16] This allows us to estimate rate gaps for each mortgage over its life, controlling for observables that drive mortgage rates at origination. The right-hand side of Figure 2 shows the empirical distribution of rate gaps and rate gaps upon prepayment in our data. We then use the summary statistic (16) to determine the attention rate in our borrower population, assuming a mortgage rate volatility of either $\sigma = 0.70\%$ (the relevant empirical value for our sample period) or $\sigma = 1.00\%$ (the value used by Agarwal, Driscoll and Laibson (2013)), and a moving/amortization rate $\nu = 11.33\%$.[17]

Given an average rate gap of 0.15% and average rate gap upon prepayment of 0.78%,[18] inverting (16) leads to an attention rate of $\lambda = 23.4\%$ per year (when $\sigma = 0.70\%$) or $\lambda = 11.8\%$ per year (when $\sigma = 1.00\%$).[19] These attention rates are slightly lower than

[16] With our origination data, we regress mortgage coupons onto (i) original LTV, (ii) original FICO, (iii) original principal balance, (iv) dummy for loan purpose, and (v) origination-month fixed effects. For each mortgage-month, we then use our linear model to estimate the rate at which the borrower could refinance, taking into account the then-current mortgage balance and LTV, computed using HPI at the 3-digit zip obtained from the FHFA.

[17] Focusing on the prepayment rate for households with negative rate gaps allows us to estimate an average moving rate of 8% p.a., to which we add the amortization rate of 1/30, leading to our parameter choice $\nu = 11.3\%$.

[18] This latter statistic rules out alternative models of inattentive borrowers who are also "naïve". As mentioned in Section 4, these alternatives deliver an ergodic average rate gap upon prepayment of either $(\rho + \nu)\psi$ or 0, which will be too low — for reasonable values of $\rho, \nu, \psi$ — relative to the data.

[19] A rate of 23.4% per year means that *on average*, households pay attention to mortgage rates once every $1/0.234 = 4.3$ years. These relatively low attention estimates are partly due to the trend decline in rates from the early 80s to 2022, causing the empirical average rate gap to be biased upwards relative to the theoretical ergodic value, thereby potentially biasing downwards our attention rate estimate.

those in Andersen et al. (2020),[20] possibly suggesting that Danes face fewer financial frictions when refinancing.[21] Abel, Eberly and Panageas (2013) provide calibrations for rational inattention to the stock market, where the time between observations ranges from 0.097 to 0.190 years, building on the empirical evidence in Alvarez, Guiso and Lippi (2012). This indicates more frequent observations than in mortgage refinancing, likely due to higher stock market volatility and the fact that mortgage rates depend on individual household characteristics, requiring more time-consuming quotes.

For each mortgage-month observation $(i, t)$ and for both mortgage rate volatilities considered, we then compute the optimal rate gap threshold $\theta_{\infty,it}$ under the perfect attention benchmark, and the corresponding threshold $\theta_{it}$ under our estimated attention rates.[22] These calculations allow us to quantify the frequency and magnitude of errors of omission and of commission imputed by both models, which we show in Table 1.

Comparing the model with inattention to the model with perfect attention in Table 1, several takeaways emerge. For brevity, we focus on the results with $\sigma = 0.70\%$, but conclusions are similar for the case $\sigma = 1.00\%$. First, as expected, the average optimal threshold $\theta$ is significantly lower in the calibrations with inattention than those with perfect attention. Second, 2.97% of mortgage-month observations with a rate gap above the ADL threshold $\theta_\infty$ end up prepaying, supporting the analysis from previous academic work suggesting that borrowers make frequent errors of omission, *under the assumption that they are perfectly attentive*. Once we factor in their inattention, 2.77% of mortgage-month observations with a rate gap above the threshold $\theta$ end up prepaying, vs. a theoretical prepayment rate of 2.85%; in other words, inattention rationalizes almost all errors of "omission".[23] Third, about 310k prepayments occur at rate gaps that are *too low* relative to the ADL threshold $\theta_\infty$ — what the literature labels errors of commissions. Taking into account borrower inattention reduces the optimal threshold, so that 32% of these errors of commission in the perfect attention model can instead be rationalized by the model with inattention.[24] Thus, our model of mortgage refinancing with inatten-

[20] In their model 3, in which households are homogeneous in attention, asleep probability is estimated to be 92% per quarter, corresponding to an attention rate of $-(\ln .92)/.25 = 33\%$ per year.

[21] For instance, in the US a borrower is required to be employed when refinancing, which is not the case in Denmark.

[22] To compute the optimal rate gaps $\theta_{\infty,it}$ and $\theta_{it}$, we use an approach consistent with Agarwal, Driscoll and Laibson (2013): we assume refinancing costs of \$2,000 plus 1% of outstanding loan balance, a tax rate of 28% — since mortgage interest is assumed to be tax deductible — and a subjective discount rate $\rho = 5\%$. We keep $\nu = 11.33\%$ and in order to compute $\theta_{it}$, we use the estimated attention rate consistent with our mortgage volatility assumption. The $i, t$ dependence of the threshold stems from the time-varying balance of the mortgage and the assumed fixed costs that do not scale with loan balance.

[23] The theoretical prepayment rate for these observations is $1 - \exp\left(-(\lambda + \nu)dt\right)$, with $dt = 1/12$ years.

[24] SFLP data does not differentiate prepayments occurring due to refinancing, vs. moves. If we assume that all prepayments occurring at $z_{it} < 0$ are actually moves and if we label as errors of omissions only

**Table 1:** Errors of omission and of commission

| | Perfect Attention | With Inattention | Perfect Attention | With Inattention |
|---|---|---|---|---|
| mortgage rate (annual) volatility $\sigma$ | 0.70% | 0.70% | 1.00% | 1.00% |
| average threshold $\theta_{it}$ | 1.09% | 0.75% | 1.27% | 0.71% |
| errors of omission | | | | |
| mortgage-month obs (M) with $z_{it} > \theta_{it}$ | 7.07 | 10.86 | 5.43 | 11.47 |
| nb of prepayments (M) when $z_{it} > \theta_{it}$ | 0.21 | 0.3 | 0.16 | 0.31 |
| empirical monthly prepay rate when $z_{it} > \theta_{it}$ | 2.97 | 2.77 | 3.01 | 2.72 |
| theoretical monthly prepay rate when $z_{it} > \theta_{it}$ | 100 | 2.85 | 100 | 1.91 |
| errors of commission | | | | |
| mortgage-month obs (M) with $z_{it} < \theta_{it}$ | 28.35 | 24.56 | 29.99 | 23.95 |
| nb of prepayments (M) when $z_{it} < \theta_{it}$ | 0.31 | 0.21 | 0.35 | 0.2 |
| empirical monthly prepay rate when $z_{it} < \theta_{it}$ | 1.08 | 0.87 | 1.17 | 0.85 |
| theoretical monthly prepay rate when $z_{it} < \theta_{it}$ | 0.94 | 0.94 | 0.94 | 0.94 |
| mortgage-month obs (M) with $0 < z_{it} < \theta_{it}$ | 13.67 | 9.88 | 15.31 | 9.27 |
| nb of prepayments (M) when $0 < z_{it} < \theta_{it}$ | 0.21 | 0.12 | 0.26 | 0.11 |
| empirical monthly prepay rate when $0 < z_{it} < \theta_{it}$ | 1.57 | 1.24 | 1.7 | 1.2 |
| mortgage-month obs (M) with $z_{it} < 0$ | 14.68 | 14.68 | 14.68 | 14.68 |
| nb of prepayments (M) when $z_{it} < 0$ | 0.09 | 0.09 | 0.09 | 0.09 |
| empirical monthly prepay rate when $z_{it} < 0$ | 0.62 | 0.62 | 0.62 | 0.62 |

Random sample of 780k mortgages issued between January 1999 and September 2023 (mortgage rate volatility over that sample period is estimated to be equal to 0.70% per year). Panel consists of 35.42M mortgage-month observations. Prepayment rates are expressed in % per month.

tion considerably reduces the two types of mistakes identified in the previous literature, at the same time emphasizing the significant role played by inattention frictions in the data.

# 8 Conclusion

Fixed-rate mortgage refinancing decisions are hindered by various frictions — both behavioral, as well as financial. These frictions interact in non-trivial ways, such that more inattention leads borrowers to refinance earlier, when they have the opportunity to do so. The inattention friction can help rationalize why borrowers both (i) refinance too "early" (compared to a full attention benchmark) and (ii) wait too long to refinance, when it is

those prepayments occurring when $0 < z_{it} < \theta$, then 210k prepayments occur at rate gaps that are *too low* relative to the ADL threshold, amongst which 43% could be rationalized with our borrower inattention model.

profitable for them to do so. Our model of borrower behavior also has consequences for the distribution of rate gaps, which can be informative of the degree of state- and time-dependent frictions faced by borrowers. Finally, our analysis is primarily positive, but we note that our framework is more general than the ADL setting and so might also be useful for more normative analysis of refinancing behavior.

# References

**Abel, Andrew B, Janice C Eberly, and Stavros Panageas.** 2007. "Optimal inattention to the stock market." *American economic review*, 97(2): 244–249.

**Abel, Andrew B, Janice C Eberly, and Stavros Panageas.** 2013. "Optimal inattention to the stock market with information costs and transactions costs." *Econometrica*, 81(4): 1455–1481.

**Agarwal, Sumit, John C Driscoll, and David I Laibson.** 2013. "Optimal mortgage refinancing: a closed-form solution." *Journal of Money, Credit and Banking*, 45(4): 591–622.

**Agarwal, Sumit, Richard J Rosen, and Vincent Yao.** 2016. "Why do borrowers make mortgage refinancing mistakes?" *Management Science*, 62(12): 3494–3509.

**Alvarez, Fernando, Luigi Guiso, and Francesco Lippi.** 2012. "Durable consumption and asset management with transaction and observation costs." *American Economic Review*, 102(5): 2272–2300.

**Andersen, Steffen, John Y Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai.** 2020. "Sources of inaction in household finance: Evidence from the Danish mortgage market." *American Economic Review*, 110(10): 3184–3230.

**Berger, David, Konstantin Milbradt, Fabrice Tourre, and Joe Vavra.** 2024. "Refinancing Frictions, Mortgage Pricing and Redistribution." working paper.

**Bierens, Herman J.** 1997. "Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the US price level and interest rate." *Journal of Econometrics*, 81(1): 29–64.

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

**Fleming, Wendell H, and Halil Mete Soner.** 2006. *Controlled Markov processes and viscosity solutions.* Vol. 25, Springer Science & Business Media.

**Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery.** 2019. "The role of technology in mortgage lending." *The Review of Financial Studies*, 32(5): 1854–1899.

**Kadan, Ohad, Hong Liu, and Jun Yang.** 2009. "Inattention, forced exercise, and the valuation of executive stock options."

**Keys, Benjamin J, Devin G Pope, and Jaren C Pope.** 2016. "Failure to refinance." *Journal of Financial Economics*, 122(3): 482–499.

**Perron, Pierre.** 1989. "The great crash, the oil price shock, and the unit root hypothesis." *Econometrica: journal of the Econometric Society*, 1361–1401.

**Stanton, Richard.** 1995. "Rational Prepayment and the Valuation of Mortgage-Backed Securities." *The Review of Financial Studies*, 8(3): 677–708.

**Stock, James H, and Mark W Watson.** 1988. "Testing for common trends." *Journal of the American statistical Association*, 83(404): 1097–1107.

**Stokey, Nancy L, and Robert E Lucas Jr.** 1989. *Recursive methods in economic dynamics.* Harvard University Press.

# A Appendix

## A.1 Properties of $V$

**Proof of Proposition 1.** The recursive formulation of problem (1) is

$$V(m,c) = \mathbb{E}_m\left[\int_0^{\tau_\nu\wedge\tau_\lambda} e^{-\rho t}c\,dt + e^{-\rho\tau_\nu}\mathbb{1}_{\{\tau_\nu<\tau_\lambda\}}V\left(m_{\tau_\nu}, m_{\tau_\nu}\right)\right.$$
$$\left. + e^{-\rho\tau_\lambda}\mathbb{1}_{\{\tau_\lambda<\tau_\nu\}}\min\left(V\left(m_{\tau_\lambda}, m_{\tau_\lambda}\right)+\psi; V\left(m_{\tau_\lambda}, c\right)\right)\right], \quad \text{(A.1)}$$

where $\tau_\lambda$ (resp. $\tau_\nu$) is an exponentially distributed time with parameter $\lambda$ (resp. $\nu$). Since these times are independent of $m_t$, (A.1) simplifies to

$$V(m,c) = \frac{c}{\rho+\lambda+\nu} + \frac{\lambda}{\rho+\lambda+\nu}\mathbb{E}_m\left[\min\left(V\left(m_\tau, m_\tau\right)+\psi; V\left(m_\tau, c\right)\right)\right]$$
$$+ \frac{\nu}{\rho+\lambda+\nu}\mathbb{E}_m\left[V\left(m_\tau, m_\tau\right)\right], \quad \text{(A.2)}$$

with $\tau$ an exponentially distributed time with parameter $\nu+\lambda$. Since $m_t$ evolves on $[\underline{m}, \bar{m}]$, the set of feasible states $(m,c)$ is a compact set $S := [\underline{m}, \bar{m}] \times [\underline{m}, \bar{m}]$ and the assumptions of theorem 9.2 in Stokey and Lucas Jr (1989) are satisfied, so that the solution to the sequence problem (1) solves Bellman equation (A.2).

Consider the mapping $\mathbb{T}$, operating on the space of continuous bounded functions $\mathcal{C}(S)$ as follows:

$$\mathbb{T}(f)(m,c) := \frac{c}{\rho+\lambda+\nu} + \frac{\lambda}{\rho+\lambda+\nu}\mathbb{E}_m\left[\min\left(f\left(m_\tau, m_\tau\right)+\psi; f\left(m_\tau, c\right)\right)\right]$$
$$+ \frac{\nu}{\rho+\lambda+\nu}\mathbb{E}_m\left[f\left(m_\tau, m_\tau\right)\right]. \quad \text{(A.3)}$$

Since $m_t$ is a diffusion, the transition kernel of $m_\tau$ conditional on the starting point $m_0$ is continuous in $m_0$ and the Feller property holds for any stopping time $\tau$, so in particular it will hold for an exponentially distributed time $\tau$. And since $S$ is compact, $\mathbb{T}$ maps any function $f \in \mathcal{C}(S)$ into a function $\mathbb{T}f \in \mathcal{C}(S)$. $\mathbb{T}$ satisfies Blackwell' sufficiency conditions so is a contraction of modulus $\frac{\lambda+\nu}{\rho+\lambda+\nu}$, and $\mathcal{C}(S)$ is a complete metric space (when equipped with the sup norm), so the contraction mapping theorem shows that $\mathbb{T}$ admits a unique fixed point $V$. Since the period return function $c \to c/(\rho+\lambda+\nu)$ in (A.3) is strictly increasing in $c$, a corollary to the contraction mapping theorem guarantees that $V$ is strictly increasing. The extension of $V$ to $[\underline{m}, \bar{m}] \times [\underline{m}, +\infty)$ is well defined and has limit $+\infty$ as $c \to +\infty$.

Since $V$ is continuous, strictly increasing in $c$ with limit $+\infty$ as $c \to +\infty$, for all $m$ there exists a threshold $\theta(m)$ that satisfies (2). ∎

## A.2 Effect of inattention on threshold

**Proof of Proposition 2.** Consider equation (3) satisfied by $V$, and differentiate w.r.t. $\lambda$:

$$(\rho+\nu+\lambda)\,\partial_\lambda V(m,c) = \min\left(V(m,m)+\psi-V(m,c),0\right)+\nu\partial_\lambda V(m,m) \\ +\lambda\left[\mathbb{1}_{\{c-m\geq\theta(m)\}}\partial_\lambda V(m,m)+\mathbb{1}_{\{c-m<\theta(m)\}}\partial_\lambda V(m,c)\right]+\mathcal{L}\partial_\lambda V(m,c)$$

Noting $\tau$ a Poisson time with arrival rate $\nu+\lambda\mathbb{1}_{\{c-m_t\geq\theta(m_t)\}}$, then

$$\partial_\lambda V(m,c) = \mathbb{E}_m\left[\int_0^\tau e^{-\rho t}\min\left(V(m_t,m_t)+\psi-V(m_t,c),0\right)dt \right. \\ \left. +e^{-\rho\tau}\partial_\lambda V(m_\tau,m_\tau)\right] \tag{A.4}$$

$\partial_\lambda V(m,c) < 0$ since a borrower with a higher attention rate must have a lower present value of future mortgage costs. The source term $\min\left(V(m_t,m_t)+\psi-V(m_t,c),0\right)$ in (A.4) is decreasing in $c$, which means that $\partial_\lambda V(m,c)$ must also be decreasing in $c$. Consider then (2) and differentiate w.r.t. $\lambda$:

$$\partial_\lambda V(m,m) = \partial_\lambda V(m,m+\theta(m)) + \partial_\lambda\theta(m)\partial_c V(m,m+\theta(m))$$

Thus, we have

$$\partial_\lambda\theta(m) = \frac{\partial_\lambda V(m,m)-\partial_\lambda V(m,m+\theta(m))}{\partial_c V(m,m+\theta(m))} \tag{A.5}$$

$V$ is increasing in $c$, which means that the denominator in (A.5) is positive. Similarly, we have established above that $\partial_\lambda V(m,c)$ is decreasing in $c$, meaning that we must have $\partial_\lambda V(m,m) > \partial_\lambda V(m,m+\theta(m))$. Thus $\theta(m)$ is increasing in $\lambda$. ■

## A.3 Special case: $m_t$ as a Brownian motion

**Proof of Proposition 3.** Assume that

$$m_t = m_0 + \sigma B_t, \tag{A.6}$$

with $B_t$ a Brownian motion. $V$, solution of (1), satisfies $V(m,c) = \frac{c}{\rho}+v(z)$, where $z = c-m$ is the rate gap and

$$v(z) := \inf_{a\in\mathcal{A}} \mathbb{E}_z\left[\int_0^{+\infty} e^{-\rho t}\left[\left(\psi-\frac{z_{t-}^{(a)}}{\rho}\right)a_t dN_t^{(\lambda)} - \frac{z_{t-}^{(a)}}{\rho}dN_t^{(\nu)}\right]\right] \tag{A.7}$$

$$dz_t^{(a)} = -\sigma dB_t - z_{t-}^{(a)}\left(a_t dN_t^{(\lambda)} + dN_t^{(\nu)}\right). \tag{A.8}$$

The option value $v$ satisfies

$$(\rho+\nu+\lambda)v(z) = \frac{\sigma^2}{2}v''(z)+\lambda\min\left(v(z),v(0)+\psi-\frac{z}{\rho}\right)+\nu\left(v(0)-\frac{z}{\rho}\right) \tag{A.9}$$

Noting $\theta$ the rate gap above which the borrower finds it optimal to refinance when given the opportunity to do so, we must have

$$(\rho+\nu)v(z) = \frac{\sigma^2}{2}v''(z) + \nu\left(v(0) - \frac{z}{\rho}\right) \qquad z \leq \theta \quad \text{(A.10)}$$

$$(\rho+\nu+\lambda)v(z) = \frac{\sigma^2}{2}v''(z) + \nu\left(v(0) - \frac{z}{\rho}\right) + \lambda\left(v(0) + \psi - \frac{z}{\rho}\right) \qquad z \geq \theta \quad \text{(A.11)}$$

Since $v(z) = O(z)$ as $z \to +\infty$ and as $z \to -\infty$, the value function $v$ must satisfy (4). Since $\theta > 0$, it must be the case that

$$v(0) = \frac{\rho+\nu}{\rho}k_- e^{-\eta_0\theta}$$

$v$ must be continuously differentiable at $z = \theta$. This yields a system of 2 linear equations in the 2 unknown $k_-, k_+$, with solution (using our formula for $v(0)$):

$$k_- = \frac{\lambda}{(\rho+\nu+\lambda)(\eta_0+\eta_\lambda) - \lambda\eta_\lambda e^{-\eta_0\theta}}\left[\eta_\lambda\left(\psi - \frac{\theta}{\rho+\nu}\right) - \frac{1}{\rho+\nu}\right]$$

$$k_+ = \frac{-\lambda}{(\rho+\nu+\lambda)(\eta_0+\eta_\lambda) - \lambda\eta_\lambda e^{-\eta_0\theta}}\left[\left(1 - \frac{\lambda e^{-\eta_0\theta}}{\rho+\lambda+\nu}\right)\left(\frac{1}{\rho+\nu}\right) + \eta_0\left(\psi - \frac{\theta}{\rho+\nu}\right)\right]$$

At $z = \theta$, the borrower is indifferent between (a) continuing with the current mortgage, or (b) paying the fixed cost and refinancing. Thus:

$$v(\theta) = v(0) + \psi - \frac{\theta}{\rho} \Rightarrow k_- = \frac{\rho}{\rho+\nu}v(0) + \psi - \frac{\theta}{\rho+\nu}$$

But since we know $v(0)$ as a function of $k_-$, this yields

$$k_- = k_- e^{-\eta_0\theta} + \psi - \frac{\theta}{\rho+\nu}$$

Using our formula for $k_-$, this yields the implicit equation

$$e^{-\eta_0\theta} + (\eta_0 + \epsilon_\lambda)\theta = 1 + (\rho+\nu)\psi(\eta_0+\epsilon_\lambda), \qquad \text{(A.12)}$$

where $\epsilon_\lambda$ is defined in Proposition 3. (A.12) is solved by

$$\theta = (\rho+\nu)\psi + \frac{1}{\eta_0+\epsilon_\lambda} + \frac{1}{\eta_0}W\left(\frac{-\eta_0}{\eta_0+\epsilon_\lambda}\exp\left\{\frac{-\eta_0}{\eta_0+\epsilon_\lambda}\left[1 + (\rho+\nu)(\eta_0+\epsilon_\lambda)\psi\right]\right\}\right),$$

with $W(\cdot)$ the primary branch of the Lambert-W function. Notice that $\epsilon_\lambda$ is a positive and decreasing function of $\lambda$, converging to zero as $\lambda \to +\infty$. Differentiate (A.12) w.r.t.

$\lambda$ to obtain

$$\frac{\partial \theta}{\partial \lambda} = \frac{\left((\rho+\nu)\psi - \theta\right)\frac{\partial \epsilon_\lambda}{\partial \lambda}}{\epsilon_\lambda + \eta_0\left(1 - e^{-\eta_0 \theta}\right)} > 0,$$

with the last inequality following from $\frac{\partial \epsilon_\lambda}{\partial \lambda} < 0$ and $(\rho+\nu)\psi - \theta = \frac{e^{-\eta_0\theta}-1}{\eta_0+\epsilon_\lambda} < 0$. Thus $\theta$ increases with $\lambda$. Since $\lim_{\lambda\to\infty}\epsilon_\lambda = 0$,

$$\theta_\infty := \lim_{\lambda\to\infty}\theta = (\rho+\nu)\,\psi + \frac{1}{\eta_0} + \frac{1}{\eta_0}W\left(-\exp\left\{-\left[1+(\rho+\nu)\,\eta_0\psi\right]\right\}\right),$$

which is the ADL formula. Lastly, since $\lim_{\lambda\to 0}\epsilon_\lambda = \infty$ and $W(0) = 0$, we have

$$\theta_0 := \lim_{\lambda\to 0}\theta = (\rho+\nu)\,\psi.$$

We can then perform a Taylor expansion of (A.12) around $\theta = 0$, which allows us to obtain an approximation $\hat{\theta}$ of the value $\theta$:

$$\frac{\eta_0^2}{2}\hat{\theta}^2 + \epsilon_\lambda\hat{\theta} - (\rho+\nu)(\eta_0+\epsilon_\lambda)\psi = 0$$

This allows us to conclude that the approximation $\hat{\theta}$ is equal to

$$\hat{\theta} = \sqrt{\frac{2}{\eta_0}\left(1+\frac{\epsilon_\lambda}{\eta_0}\right)(\rho+\nu)\,\psi + \left(\frac{\epsilon_\lambda}{\eta_0^2}\right)^2} - \frac{\epsilon_\lambda}{\eta_0^2}$$

Finally, it is straightforward (but tedious) to verify that the optimal rate gap threshold $\theta$ is identical to that derived above if one were to assume a fixed cost upon moving. ■

## A.4 Ergodic density when $m_t$ is a Brownian motion

**Proof of Proposition 4.** The rate gap follows dynamics described by (A.8). This stochastic process admits an ergodic density $f$, which satisfies the Kolmogorov-Forward equation, for $z \neq \theta$:

$$0 = \frac{\sigma^2}{2}f''(z) - \left(\nu + \lambda \mathbb{1}_{z\geq\theta}\right)f(z) \tag{A.13}$$

Moreover, $f$ is continuous at $z = \theta$, it vanishes at $z \to \pm\infty$, and it integrates to 1. Using the constants $\chi_0, \chi_\lambda$ defined in Proposition 4, we can integrate (A.13) with the above boundary conditions to derive (11). The ergodic average rate gap $\mathbb{E}[z_t]$ is then equal to

$$\mathbb{E}[z_t] = \int_{-\infty}^{+\infty} z f(z)dz = \theta + \frac{\chi_0 - \chi_\lambda}{\chi_0\chi_\lambda}$$

Similarly, the ergodic average prepayment rate equals

$$\lim_{t\to+\infty}\frac{1}{t}\left(N_t^{(\nu)}+\int_0^t a_t dN_t^{(\lambda)}\right)=\int_{-\infty}^{+\infty}\left(\nu+\lambda\mathbb{1}_{z\geq\theta}\right)f(z)dz=\nu+\left(\frac{\chi_0}{\chi_0+\chi_\lambda}\right)\lambda$$

Lastly, the ergodic average rate gap observed at the time of prepayment equals

$$\lim_{t\to+\infty}\frac{\int_0^t z_{t-}\left(dN_t^{(\nu)}+a_t dN_t^{(\lambda)}\right)}{N_t^{(\nu)}+\int_0^t a_t dN_t^{(\lambda)}}=\frac{\int_{-\infty}^{+\infty} z\left(\nu+\lambda\mathbb{1}_{z\geq\theta}\right)f(z)dz}{\int_{-\infty}^{+\infty}\left(\nu+\lambda\mathbb{1}_{z\geq\theta}\right)f(z)dz}=\theta$$

■

## A.5 Rational inattention

Assume that $m_t$ satisfies (A.6). Let $V$ be the valuation for a borrower at the time she observes the current market rate and is not refinancing:

$$V(m,c):=\phi+\min_{T\geq 0}\mathbb{E}_m\left[\int_0^{\min(\tau_\nu,T)} e^{-\rho t}c\,dt+e^{-\rho\tau_\nu}\mathbb{1}_{\{\tau_\nu\leq T\}}V(m_{\tau_\nu},m_{\tau_\nu})\right.$$
$$\left.+e^{-\rho T}\mathbb{1}_{\{\tau_\nu>T\}}\min\left(V(m_T,c);V(m_T,m_T)+\psi\right)\right],$$

We guess that $V(m,c)=\phi+\frac{c}{\rho}+v(z)$. After some algebra, the option value $v(z)$ solves

$$v(z)=\min_{T\geq 0}\frac{\nu}{\rho+\nu}\left(1-e^{-(\rho+\nu)T}\right)\left(v(0)+\phi-\frac{z}{\rho}\right)$$
$$+e^{-(\rho+\nu)T}\left[\phi+\mathbb{E}\left[\min\left(v(z-\sigma B_T);v(0)+\psi-\frac{z-\sigma B_T}{\rho}\right)\right]\right]$$

The optimal rate gap threshold $\theta$ satisfies

$$\frac{\theta}{\rho}+v(\theta)=v(0)+\psi.$$

The optimal wait time $\mathbb{T}(z)$ satisfies the first order condition

$$(\rho+\nu)v(z)-\nu\left(v(0)+\phi-\frac{z}{\rho}\right)=e^{-(\rho+\nu)\mathbb{T}(z)}\left.\frac{d\mathbb{E}\left[\min\left(v(z-\sigma B_T);v(0)+\psi-\frac{z-\sigma B_T}{\rho}\right)\right]}{dT}\right|_{T=\mathbb{T}(z)},$$

in other words the flow cost of waiting (the left-hand side above) has to equal the marginal increase in the present value of expected option payoff from waiting an incremental unit of time. The state of a given borrower is $(T,z)$, with $T$ the time until next observation and $z$ the rate gap. $(T,z)$ is a Markov process that satisfies

$$dT_t=-dt+\mathbb{T}(z_{t-})\mathbb{1}_{\{T_{t-}=0\}}$$

$$dz_t = -\sigma dB_t - z_{t-} \left( dN_t^{(\nu)} + \mathbb{1}_{\{T_{t-}=0\}} \mathbb{1}_{\{z_{t-} \geq \theta\}} \right)$$

$(T, z)$ admits an ergodic density that can be computed via simulation, allowing us to derive various moments — in particular the ergodic average attention delay, defined via

$$\mathcal{T} := \lim_{t \to +\infty} \frac{1}{t} \sum_{0 \leq s \leq t, T_{s-}=0} \mathbb{T}(z_s).$$