CAPITAL TAXATION AND ASSET PRICE VOLATILITY*

PAU BELDA†

Latest version here.

Abstract

Recent decades have seen higher wealth volatility due to asset price fluctuations. This paper argues that declining capital taxes are a key factor. In a model where investors learn about prices, lower taxes amplify the pass-through from expectations to prices, fueling a price-expectations spiral that generates volatility. Survey evidence supports this higher pass-through when taxes were cut. The estimated model replicates many U.S. stock market moments and suggests tax cuts increased volatility by 35%. Capital gains tax cuts had twice the impact on volatility compared to dividend tax cuts, emerging as particularly relevant for financial instability.

JEL codes: D83, D84, E44, G12, G14, H20, H31

*I am especially grateful to Albert Marcet for disciplining my thoughts throughout the research process. In addition, I have benefited from comments from Luis E. Rojas, Jordi Caballé, Nicola Gennaioli, Ricardo Reis, Thomas Piketty, Dino Palazzo, Jaume Ventura, Joan Alegre, Ramon Ruiz, Francisco Palomino, Ram S. Yamarthy and different audiences in many presentations at conferences and seminars. The initial stages of the project were motivated by passionate discussions about the causes of the Great Depression with José Luís Peña-Mir. This research received funding from the European Research Council under the European Union's Horizon2020, research and innovation program GA project number 788547 (APMPAL-HET). The views expressed in this paper are only personal and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System. First version: July 2020. This version: August 2024.

†Federal Reserve Board. Address: 20th Constitution Ave. N.W., 20551 Washington, DC. PO box: K-6403. E-mail: pau.beldaitortosa@gmail.com.

1. Introduction

Wealth is back (Piketty and Zucman, 2014), but it has returned more volatile. While the aggregate wealth-income ratio has surged, this rise has been marked by substantial fluctuations tied to asset price booms and busts. The Dotcom bubble, for instance, created and subsequently destroyed wealth equivalent to about 50% of the GDP at that time, with the figure climbing to around 100% during the housing bubble. The larger wealth volatility starkly contrasts with the greater stability observed in many macroeconomic indicators since the mid-1980s. This disconnect challenges conventional asset pricing theories, which posit a strong link between asset valuations and the underlying real conditions. I argue that the reduction in capital taxes, especially on capital gains, is a key factor behind this disconnect.

This paper presents three main results. First, the variance of asset valuations is inversely related to capital taxes. This insight holds under different expectation models, including Rational Expectations (RE), but is amplified when agents extrapolate from prices. Second, using a structurally estimated model that matches U.S. stock market data, I find that capital tax cuts led to significant increases not only in mean valuations but also in the volatility of these valuations. These results suggest that tax cuts have first-order consequences not only for corporations' values or wealth inequality, as argued by McGrattan and Prescott (2005) and Piketty (2014), but also for financial instability. Third, not all capital taxes are the same: Capital Gains Tax (CGT) cuts have twice the impact on volatility than Dividends Tax (DT) cuts. This result suggests a special role for CGT in curbing excessive asset price volatility, contrasting with the inconclusive evidence on the popular Financial Transactions Tax.[1]

I start by studying how a constant tax can affect price volatility. In a general asset pricing model à la Lucas (1978) with taxes on dividends and capital gains, the tax rate times the probability of these payments appear as a wedge between payoff expectations and prices. This wedge affects the pass-through from expectations to asset values. For instance, if taxes are higher, an equivalent increase in optimism leads to a smaller change in prices than otherwise. Given that shifts in investor expectations are the most proximate driver of price movements, controlling the pass-through from expectations to prices is critical in shaping market volatility.

The mechanism works for any given model of expectations, including RE, but it is dynamically reinforced when investors extrapolate from prices, as has been documented in surveys (for instance,

[1]The literature tends to suggest that while such a tax may curb speculative trading, its impact on price volatility is unclear. See, for instance, Buss and Dumas (2019) and Dávila (2023) for theoretical analysis and Umlauf (1993) and Cappelletti, Guazzarotti, and Tommasino (2017) for empirical evidence.

Greenwood and Shleifer, 2014; Cassella and Gulen, 2018). Extrapolation is microfounded following the *Internal Rationality* approach of Adam and Marcet (2011), whereby imperfect information pushes fully rational agents to learn about equilibrium prices from available data. Bayesian learning connects prices back to expectations, giving rise to a price-expectations spiral: more optimism leads to higher prices, which in turn feed back into further optimism. In these self-fulfilling booms and busts, higher taxes act as a brake, reducing the pass-through at each iteration of the price-expectations loop. Thus, with extrapolation, taxes stabilize prices because they dampen the price-expectations spiral that causes market instability.

A key innovation in the model is the incorporation of a CGT paid *on realization* (that is, paid only when the asset is sold), a salient feature of the tax code that has often been omitted in asset pricing models. The *on realization* clause opens the door to a behavioral reaction in which investors could defer taxation by strategically retiming their portfolio rebalancing decisions, potentially undoing the stabilizing effect outlined above. This behavioral response, known as the *lock-in effect*, was invoked by Stiglitz (1983) to conclude that a higher CGT would destabilize—rather than stabilize—markets. I show that the relative strength of the effect depends on how much investors adjust their gains realization decisions when the tax rate changes, the so-called *tax elasticity of realization*. It turns out that only when investors are very reactive, with an elasticity lower than -1, the *lock-in effect* dominates. In a structural estimation of the model, the elasticity is found to be -0.51, aligned with recent estimates by Agersnap and Zidar (2021). The result suggests that, while the lock-in effect is relevant, it is not dominant, and the stabilizing effect of taxes prevails.

According to these theoretical insights, the observed fall in capital taxes since the 1980s would have increased the pass-through from beliefs to asset prices, boosting self-fulfilling booms and busts in asset prices that are behind the rise in wealth volatility. Indeed, using a long-run series of measured expectations, I find that the pass-through increased significantly after the 1980s.[2] In addition, the mechanism is consistent with the evidence that the sensitivity of investment decisions to beliefs is higher for tax-advantaged investors (Giglio et al., 2021). Thus, the theory relates lower taxes to higher volatility through a mechanism grounded in microdata.

However, how much can tax cuts actually explain? To answer this question, I estimate the model using the Simulated Method of Moments (SMM) and stock market data since 1946. The model shows quantitative consistency with both conventional asset pricing statistics—such as excess volatility, return predictability, and the equity premium—and survey expectations. Importantly,

[2] An alternative measure that does not require surveys is an IRF that captures the responses of valuations to equivalent shocks before and after the 1980s. The one-year-forward IRF coefficient increased by 75% after the 1980s, reflecting a greater sensitivity of valuations to shocks. See Section 3.

I estimate the model using two subsamples, before and after the 1980s, to highlight significant changes in price dynamics that many asset pricing models miss.[3] When fed with empirical time series of tax cuts, the model also replicates the changes since the 1980s: higher valuations, increased volatility, lower predictability, and a higher pass-through from expectations to prices.

Given the model's remarkable quantitative performance, I use it to estimate the average effect of tax cuts. I simulate scenarios without tax cuts and calculate a measure conceptually similar to an average treatment effect. The baseline estimate indicates that 40% of the increase in the variance of valuations was due to tax cuts. In an extensive sensitivity analysis that explores parameter uncertainty and alternative modeling assumptions (for instance, tax foresight as suggested by Kueng, 2014), the effect often falls between 30-40%. For mean valuations, the average effect is around 25-35%. Thus, tax cuts appear to be crucial for both mean and variance.

While the previous results apply to both CGT and DT, these types of taxes have significantly different effects. Specifically, I find that CGT cuts have twice the impact on volatility compared to equivalent DT cuts. This difference arises because, although both CGT and DT affect the pass-through from beliefs to prices, only the CGT level influences price growth in equilibrium. A constant DT cancels out in this expression and, consequently, does not appear in the Bayesian updating equation. As a result, CGT has a more profound effect on market volatility through both the direct pass-through and the dynamics of belief updating.

An interesting implication is that, through the lens of the model, tax cuts triggered a reduction in the informational efficiency of the stock market. In the model with learning, price dynamics are not fully efficient as they reflect investor sentiment movements partly detached from fundamentals. Lower taxes would have exacerbated this issue, making prices more sensitive to subjective beliefs. Using the estimated model, I run a formal test of changes in market efficiency and find that efficiency fell by approximately 35%, with tax cuts being responsible for about 40% of this fall.

This paper contributes to several branches of literature. The first examines the effect of CGT on asset price volatility. While early debates were dominated by the view that CGT was destabilizing due to the lock-in effect (for instance, Somers, 1948; Gemmill, 1956; Somers, 1960; Stiglitz, 1983), more recent work by Dai, Shackelford, and Zhang (2013) found that volatility increased after tax cuts, using a difference-in-difference approach in a cross-section of assets around the 1978 and 1997 CGT reductions. While they suggested a reduction in risk-sharing with the government as a

[3] As Chen, Winkler, and Wasyk (2019) pointed out, models with habits à la Campbell and Cochrane (1999) or long-run risks Bansal and Yaron (2004) struggle to match asset price dynamics mostly after the 1980s. Extensions like Lettau, Ludvigson, and Wachter (2008) can explain part of the stock price surge due to lower macroeconomic risk, but not the larger price swings.

potential mechanism, I find this risk-sharing effect to be relatively unimportant in my simulations.[4] Instead, the idea of tax cuts enabling price-expectations spirals—which finds precedent in Haugen and Heins (1969) and Haugen and Wichern (1973)—offers an alternative explanation consistent with Dai, Shackelford, and Zhang (2013) statistical findings, aggregate data, and surveys.

This paper also contributes to the literature on the effects of taxes on mean asset prices (see Sialm, 2009 for a review). While the relevance *on realization* regime of CGT is acknowledged (see, for instance, Dai et al., 2008), modeling it in general equilibrium setups has proven challenging. For instance, Sialm (2006) included stochastic taxes in a Lucas tree, but considered a tax on equity values rather than on capital gains. In addition, much of the literature evaluating the consequences of tax cuts on asset pricing (as McGrattan and Prescott, 2005; Brun and González, 2017; McGrattan, 2023) or investment (for instance, Gourio and Miao, 2010; Anagnostopoulos, Cárceles-Poveda, and Lin, 2012) has abstracted from this nuance, potentially biasing their estimates. Conversely, the portfolio choice literature has incorporated more sophisticated modeling of capital gains taxation, but taking asset prices as given (for instance, Dammon, Spatt, and Zhang, 2001; Ehling, Tompaidis, and Yang, 2019). In contrast, following the approach in Gavin, Kydland, and Pakko (2007) and Gavin, Keen, and Kydland (2015), I incorporate the lock-in effect in a quantitative model.

In addition, this article contributes to the literature on excess volatility. Specifically, it relates to papers highlighting the role of subjective beliefs in explaining such volatility (for instance, Barberis et al., 2018; Jin and Sui, 2022; Bordalo et al., 2024). The closest works to this study are Adam, Marcet, and Nicolini (2016) and Adam, Marcet, and Beutel (2017), who emphasized the role of learning about prices rather than about fundamentals as a key element in quantitatively accounting for both market data and surveys. Building on their approach, this paper explores the interaction between beliefs and capital taxes, offering new theoretical insights that better account for the changes in asset pricing dynamics since the 1980s. In addition, the literature on Rational Bubbles (for instance, Martin and Ventura, 2018), has pointed to lower interest rates as a possible driver of bubbles, a hypothesis also suggested in Adam (2020). This paper proposes a complementary force—tax cuts—and offers a detailed quantification of their impact on asset price dynamics.

More broadly, this article contributes to a re-emerging literature on the taxation of capital gains. Recent work by Agersnap and Zidar (2021) has found a lower tax elasticity of realization, which Sarin et al. (2022) has used to suggest a higher revenue potential. Aguiar, Moll, and Scheuer (2024) have argued that the *on realization* clause proves optimal whenever price fluctuations are

[4] Additionally, in reality, investors can only claim compensation for losses up to $\$3,000$, a tiny amount, while the rest is transferred as a tax credit to compensate for future tax liabilities. Thus, the tax credit is transferred to future possibly good states, being then a dubious risk-sharing device. I thank Andrea Prestipino for pointing this out.

not entirely driven by cashflow changes. While capital taxation has been typically thought in the efficiency-equity plane (see Stantcheva, 2020, for a review), and financial stability concerns have been addressed through proposals for a Financial Transaction Tax (for instance, Keynes, 1936; Tobin, 1978; Stiglitz, 1989), this article suggests that existing capital income taxes, especially on capital gains, can be effective tools in preventing asset price bubbles.

The rest of the article proceeds as follows. Section 2 presents our theoretical model linking capital taxes to asset price volatility. Section 3 documents empirical facts on capital taxes, asset prices, and investor expectations. Section 4 details our quantitative model, estimation strategy, and model performance. Section 5 quantifies the effects of tax cuts and examines implications for market efficiency. Section 6 concludes.

2. Theory

This section analyzes the CGT's impact on PD ratio volatility using an asset pricing model. Section 2.1. presents a general consumption-based model with a tax on realized capital gains. Section 2.2. explores the relationship between taxes and volatility in a temporary equilibrium. Section 2.3. derives additional results when beliefs are endogenous. Section 2.4. introduces the lock-in effect. Section 2.5. considers a dividend tax.

2.1. The model

In this section, a general consumption-based asset pricing model is set up. Its basic layer is an endowment exchange economy model as in Lucas (1978), generalized to allow for the presence of habits, different fundamental growth processes, a general probability measure as in Adam and Marcet (2011) and a tax on realized capital gains.

The economy is populated by a unit mass of infinitely living investors. This stochastic exchange economy has a single risky asset S in the form of a contract that each period pays dividends D_t, consisting of a perishable consumption good. D_t is a random variable that evolves according to some exogenous stochastic process specified later. When the time starts, each investor is endowed with one unit of stock ($S^i_{-1} = 1$). There are competitive markets for goods and stocks. Goods are used as the numeraire and stocks are tradable at a real price P. Short-selling is not allowed, and there is an upper bound on the amount of stock holdings, that is, $0 \leq S^i_t \leq \bar{S}$. Lower and upper bounds on S^i_t are assumed for convenience: Economically, the lower bound rules out short-selling strategies aimed at avoiding taxes; Mathematically, these bounds ensure that the feasibility set is

compact.

Every period, investors face some risk of being hit by a catastrophic liquidity shock. $z_t^i \sim Bernoulli(\pi)$ is a random variable indicating that possibility, with π being the probability of that event occurring. If the event materializes ($z_t^i = 1$), investor i has to sell all her stock holdings. There is a tax $\tau \in [0, 1)$ on realized capital gains.[5] Capital gains and losses are treated symmetrically. Tax revenues are transferred back to each individual according to her contribution, represented by T_t^i.

Investors know the stochastic processes followed by D_t and z_t^i. Besides, they take T_t^i and P_t as non-control variables. However, other investors beliefs are unknown for agent i, that is, investors homogeneity is not common knowledge and beliefs coordination is not taken as given as in models with Rational Expectations. The underlying probability space is given by $(\Omega^i, \mathcal{B}^i, \mathcal{P}^i)$ where Ω is the state space with $\omega = \{D_t, z_t^i, T_t^i, P_t\}_{t=0}^{\infty}$ as a typical element, $\mathcal{B}$ denotes the σ-algebra of Borel subsets of Ω and $\mathcal{P}$ is a subjective probability measure over $(\Omega, \mathcal{B})$.

Investors' program. Each investor faces a consumption-savings problem: She chooses sequences of consumption, stock holdings and stock purchases $\{C_t^i, S_t^i, X_t^i\}_{t=0}^{\infty}$ by solving the following optimization program:

$$\max_{\{C_t^i, S_t^i, X_t^i\}_{t=0}^{\infty}} \mathbb{E}_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t U(C_t^i, H_t^i) \tag{1}$$

subject to the budget constraint

$$C_t^i + P_t X_t^i \leq D_t S_{t-1}^i + z_t^i \Big(S_{t-1}^i P_t - \tau(G_{t-1}^i + (P_t - P_{t-1}) S_{t-1}^i) \Big) + T_t^i \tag{2}$$

the stock holdings law of motion

$$S_t^i = (1 - z_t^i) S_{t-1}^i + X_t^i \tag{3}$$

the unrealized capital gains G_t^i law of motion

$$G_t^i = (G_{t-1}^i + (P_t - P_{t-1}) S_{t-1}^i)(1 - z_t^i) \tag{4}$$

and the stock holdings bounds specified above, given the initial individual stock holdings.

$\delta \in (0, 1)$ is a discount factor and H_t is a time-varying habit. It is assumed to depend on past individual and aggregate consumption $H_t^i = H(C_{t-1}, C_{t-1}^i)$, embedding models with both

[5] Note realization is exogenous, driven by the liquidity shock. In Section 2.4, I endogeneize π by including accumulation costs. The introduction of this liquidity shock is equivalent to assuming that investors expect that a fixed proportion of the capital gains will be realized as in Sialm (2009).

internal and external habit formation (for example, Abel, 1990; Constantinides, 1990; Campbell and Cochrane, 1999). U is assumed to be continuous and concave on both arguments.

The degree of generality of the utility function, the dividends stochastic process and the subjective probability measure it is enough to encompass a wide range of asset pricing models, including popular elements such as habits, long-run risk or subjective beliefs. The next section derives a result in this setup for a given vector of beliefs and later I will specialize U, $\{D_t\}$ and $\mathcal{P}$ to obtain results for model-consistent beliefs.

2.2. A general result in a temporary equilibrium

In this section, I show that the variance of the Price-Dividend ratio is inversely related to τ under certain conditions. Toward that end, consider the following Euler equation, which must be satisfied in equilibrium

$$P_t = \mathbb{E}_t^{\mathcal{P}^i}\left[\delta\frac{\lambda_{t+1}}{\lambda_t}\Big(D_{t+1} + P_{t+1} - z_{t+1}^i\tau(P_{t+1} - P_t)\Big)\right] \tag{5}$$

where λ_t is the Lagrange multiplier in the period t budget constraint, potentially depending on current and past consumption. Assume that the agents know the probability of the liquidity shock and let $\beta_t^D \equiv \mathbb{E}_t^{\mathcal{P}^i}\left[\delta\frac{\lambda_{t+1}}{\lambda_t}\frac{D_{t+1}}{D_t}\right]$, $\beta_t^P \equiv \mathbb{E}_t^{\mathcal{P}^i}\left[\delta\frac{\lambda_{t+1}}{\lambda_t}\frac{P_{t+1}}{P_t}\right]$ and $\beta_t^M \equiv \mathbb{E}_t^{\mathcal{P}^i}\left[\delta\frac{\lambda_{t+1}}{\lambda_t}\right]$. Manipulating the previous expression and assuming $\mathbb{C}ov\Big(\frac{\lambda_{t+1}}{\lambda_t}P_{t+1}, z_{t+1}^i\Big) = 0$ the following PD ratio can be obtained:[6]

$$\frac{P_t}{D_t} = \frac{\beta_t^D}{1 - (1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M} \tag{6}$$

This expression offers a structural mapping, $p : \mathbb{R}_+^3 \to \mathbb{R}_+$, from a vector of beliefs $\boldsymbol{\beta_t}$ into the PD ratio. For it to be positive, $\beta_t^M < \beta_t^P + \frac{1-\beta_t^P}{\pi\tau}$, which is assumed throughout.

Lemma 1. *Let $\mathbb{V}(P_t/D_t)$ be the variance of the PD ratio, $\boldsymbol{\Sigma}$ be the beliefs covariance matrix, and $\boldsymbol{\omega} \equiv \nabla p(b)$ be the sensitivity of prices to beliefs at the approximation point $b \in \mathbb{R}_+^3$. Then, up to a first-order approximation, the variance of the price-to-dividend ratio can be expressed as:*

$$\mathbb{V}\Big[\frac{P_t}{D_t}\Big] = \boldsymbol{\omega}\boldsymbol{\Sigma}\boldsymbol{\omega}^T \tag{7}$$

Proof. Appendix B.1

[6] In the next subsection I show that $\mathbb{C}ov\Big(\frac{\lambda_{t+1}}{\lambda_t}P_{t+1}, z_{t+1}^i\Big) = 0$ holds in models involving both Rational Expectations (as Bernoulli trials and the fundamentals growth shocks are both exogenous and i.i.d.) and Learning (as agents think of prices as an AR(1) process).

Lemma 1 highlights that the variance of asset valuations is influenced by two objects: the volatility of beliefs and the elasticity of prices to that volatility. Consequently, taxes can only affect the variance if they impact at least one of these two components. For now, keep $\boldsymbol{\beta}_t$ as given. Let $\mathcal{Q} \equiv \frac{(1-\beta_t^P+\pi\tau(\beta_t^P-\beta_t^M)+2(1-\pi\tau)(\beta_t^P-\beta_t^M))(1-\pi\tau)}{(1-\beta_t^P+\pi\tau(\beta_t^P-\beta_t^M)-2\pi\tau(\beta_t^P-\beta_t^M))\pi\tau}$ at the approximation point. The following lemma focuses on the effects of taxes on ω.

Lemma 2. *Let i)* Σ *be such that* $\Sigma_{12} = \Sigma_{21} \geq 0$, $\Sigma_{13} = \Sigma_{31} \leq 0$, $\Sigma_{23} = \Sigma_{32} \leq 0$ *and* $\Sigma_{33} < \mathcal{Q}\Sigma_{22}$*; ii)* $\beta_t^M \leq \beta_t^P \leq 1$*; and iii)* $\tau \leq 1/(2\pi)$*. Then,*

$$\frac{d\boldsymbol{\omega}}{d\tau}\boldsymbol{\Sigma}\boldsymbol{\omega}^T < 0 \tag{8}$$

Proof. Appendix B.2

Lemma 2 shows that under some restrictions on the variance-covariance expectations matrix, on beliefs, and on taxes, the effect of taxes on the expression is negative. Now, I discuss that these are mild restrictions. First, $\Sigma_{12} = \Sigma_{21} \geq 0$ assumes a positive covariance between expected price and dividend growth, which is standard in dividends-based asset pricing. The conditions $\Sigma_{13} = \Sigma_{31} \leq 0$ and $\Sigma_{23} = \Sigma_{32} \leq 0$ assume a non-positive covariance between expected risk-adjusted payoffs and the expected discount factor. In models with constant discounting, this covariance is zero; in models with time-varying risk, there is typically a expected counter-cyclical premium for bearing risk, so this covariance is typically negative (e.g., models with habits). $\Sigma_{33} < \mathcal{Q}\Sigma_{22}$ requires that beliefs about future risk are not too volatile with respect to risk-adjusted capital gains. The scaling constant $\mathcal{Q}$ is larger than 1, so the condition is rather permissive for expected risk variations.

Condition ii) puts an upper bound on beliefs, which is analogous to the one required in Adam, Marcet, and Nicolini (2016). Besides, it assumes that the expected risk is not greater than the expected risk-adjusted price growth, which is typically the case in a boom. Otherwise, investors expect a subsidy rather than a tax, so τ boosts returns rather than diminishing them. Finally, condition iii) assumes an upper bound on taxes. Since $\pi \in [0, 1]$, the upper bound is in the interval $[1, \infty]$, which is always satisfied given $\tau \in (0, 1)$.

These two lemmas are useful to prove the following proposition:

Proposition 1. *Under the assumptions of Lemma 1 and Lemma 2, the variance of the Price-*

Dividend ratio is inversely related to the tax level

$$\frac{d}{d\tau}\mathbb{V}\Big[\frac{P_t}{D_t}\Big] < 0 \tag{9}$$

Proof. By *Lemma 1*, $\mathbb{V}\Big[\frac{P_t}{D_t}\Big] = \boldsymbol{\omega}\boldsymbol{\Sigma}\boldsymbol{\omega}^T$. The total derivative of $\mathbb{V}\Big[\frac{P_t}{D_t}\Big]$ with respect to τ is given by

$$\frac{d}{d\tau}\mathbb{V}\Big[\frac{P_t}{D_t}\Big] = \frac{d\boldsymbol{\omega}}{d\tau}\boldsymbol{\Sigma}\boldsymbol{\omega}^T + \boldsymbol{\omega}\boldsymbol{\Sigma}\frac{d\boldsymbol{\omega}^T}{d\tau} \tag{10}$$

By *Lemma 2*, $\frac{d\boldsymbol{\omega}}{d\tau}\boldsymbol{\Sigma}\boldsymbol{\omega}^T < 0$ and $\boldsymbol{\omega}\boldsymbol{\Sigma}\frac{d\boldsymbol{\omega}^T}{d\tau} < 0$.

The result holds under a general consumption-based asset pricing setup that requires only mild restrictions on beliefs. Given the wealth of expectations models, this result offers insight into the direct effect of a CGT on price stability across asset pricing theory. Nonetheless, taxes might affect beliefs in different ways, potentially amending *Proposition 1.* In the next section, I study this indirect effect in particular setups.

2.3. Special cases with endogenous beliefs

In this section, I specialize some of the previous assumptions to derive additional results when beliefs are internal to the model. Assume $U(C_t^i, H_t^i) = C_t^i$ such that $\lambda_t = 1\ \forall t$. Let $d_t \equiv \ln\frac{D_t}{D_{t-1}}$ and consider the following stochastic process:

$$d_t = (1-\rho)\mu + \rho d_{t-1} + \varepsilon_t^d \tag{11}$$

whereby the log dividend growth contains a drift, an autoregressive component whose weight is determined by $0 < \rho < 1$ and an innovation $\varepsilon_t^d \sim i.i.\mathcal{N}(0, \sigma_d^2)$. Assume that the agents' information set includes this process as well as the probability of liquidity shocks. Then, the only thing missing to fully characterize $\mathcal{P}$ is a pricing function which, as a general equilibrium object, depends on agents' information about other market participants. For that, consider two models: full information and Rational Expectations (RE); and imperfect information and Learning.

2.3.1. Rational Expectations Equilibrium

First, assume that rational agents have all the information about market participants. In particular, $\mathcal{P}^i = \mathcal{P}^j$ for all pairs of i, j is common knowledge. Then, the standard procedure to solve an expectational difference equation -forward iteration on prices, the use of the Law of Iterated

Expectations and a standard transversality condition- delivers the following expression:

$$P_t = \mathbb{E}_t\left[\sum_{j=1}^{\infty}\left(\frac{\delta}{1-\delta\mathbb{E}_t(z_{t+1}\tau)}\right)^j D_{t+j}\prod_{l=1}^{j-1}(1-z_{t+l}\tau)\right] \tag{12}$$

The current price is equal to the familiar discounted sum of future dividends slightly modified by a two-fold τ influence. First, τ appears in the denominator of the discount factor, amplifying it. Second, it dampens how longer-term dividends are expected to affect current prices (note the upper index $j-1$ in the product operator). Given the information that agents have about fundamentals, the expression can be further simplified to obtain

$$\frac{P_t^{RE}}{D_t} = \sum_{j=1}^{\infty} f_0(\tau, j) f_1(d_t, j) \tag{13}$$

with $f_0(\tau, j) = \left(\frac{\delta}{1-\delta\pi\tau}\right)^j (1-\pi\tau)^{j-1}$ and

$$f_1(d_t, j) = exp\left\{\mu j + \frac{\sigma_d^2}{2(1-\rho)^2}\left(j - 2\frac{\rho}{1-\rho}(1-\rho^j) + \rho^2\frac{1-\rho^{2j}}{1-\rho^2}\right) + \frac{\rho}{1-\rho}(1-\rho^j)(d_t-\mu)\right\}$$

which converges iff $\left(\frac{\delta}{1-\delta\pi\tau}\right)(1-\pi\tau)exp\left\{\mu + \frac{\sigma_d^2}{2(1-\rho)^2}\right\} < 1$.[7] Given the autoregressive nature of growth, the PD ratio at time t is fully determined by the current growth rate d_t, which encapsulates expectations about short-term dividends and future prices. When $\rho = 0$, dividend growth becomes an i.i.d. process and the equilibrium PD ratio boils down to the following constant

$$\frac{P_t^{RE2}}{D_t} = \frac{\delta exp\{\mu + \sigma_d^2/2\}}{1-(1-\pi\tau)\delta exp\{\mu+\sigma_d^2/2\} - \delta\pi\tau} \tag{14}$$

provided that $|\frac{\delta(1-\pi\tau)}{1-\delta\pi\tau}exp\{\mu+\sigma_d^2/2\}| < 1$ so that the series convergence. This expression reproduces the structure of the general pricing equation (6), being a special case for $\beta_t^M = 1$ and $\beta_t^D = \beta_t^P = \delta exp\{\mu + \sigma_d^2/2\}$. Thus, investors expect capital gains to mimic dividend growth and the PD ratio becomes a constant.

[7] The derivation of the equilibrium PD ratio follows the procedure outlined by Burnside (1998). The inclusion of taxes is an easy extension given that liquidity shocks and dividend growth shocks are independent and then, $\mathbb{C}ov_t\left(\prod_{l=1}^{j}\frac{D_{t+l}}{D_{t+l-1}}, \prod_{l=1}^{j-1}(1-z_{t+l}\tau)\right) = 0$.

2.3.2. Internally Rational Equilibrium

Now consider the more general case where agents have limited market knowledge, ignoring $\mathcal{P}^i = \mathcal{P}^j$ for all i, j pairs. In this case, the Law of Iterated Expectations (LIE) does not apply as the probability measure of the marginal agent in future periods $\mathcal{P}^j$ is unknown for i, being potentially different from $\mathcal{P}^i$. Therefore, the above-described standard procedure cannot be applied (see Adam and Marcet (2011)). Instead, the expression characterizing equilibrium prices is

$$P_t = \frac{\delta}{1 - \delta \mathbb{E}_t^{\mathcal{P}^i}(z_{t+1}\tau)} \mathbb{E}_t^{\mathcal{P}^i}\left[D_{t+1} + P_{t+1}(1 - z_{t+1}^i \tau)\right] \tag{15}$$

showing that the current asset price is equal to the expected one-period ahead discounted payoffs. This can be easily rewritten in terms of the PD ratio such that

$$\frac{P_t}{D_t} = \frac{\delta \mathbb{E}_t^{\mathcal{P}}\left[\frac{D_{t+1}}{D_t}\right]}{1 - \delta \mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}(1 - z_{t+1}\tau)\right] - \delta\tau \mathbb{E}_t^{\mathcal{P}}(z_{t+1})} \tag{16}$$

that is, the ratio depends on beliefs about dividends and price growth and the probability of the liquidity shock. Although agents know $\mathbb{E}_t^{\mathcal{P}}\left[\frac{D_{t+1}}{D_t}\right] = exp\{(1-\rho)\mu + \rho d_t + \sigma_d^2/2\} \equiv \bar{\beta}_t{}^D$ and $\mathbb{E}_t^{\mathcal{P}}(z_{t+1}) = \pi$, expectations about future prices (and its interaction with future taxes) have not been determined yet. Since the LIE cannot be applied, beliefs about future prices must be characterized elsewhere and then plugged in expression (16) to pin down P_t. In other words, under imperfect knowledge, rational agents cannot deduce equilibrium prices from their optimality conditions and need a model of prices to form beliefs. The proposed model is[8]

$$\ln \frac{P_t}{P_{t-1}} = \ln b_t + \varepsilon_t^P \tag{17}$$

$$\ln b_t = \ln b_{t-1} + \nu_t^b \tag{18}$$

with b being an unobserved permanent component and $\varepsilon_t^P \sim i.i.\mathcal{N}(0, \sigma_P^2)$ and $\nu_t^b \sim i.i.\mathcal{N}(0, \sigma_b^2)$. Agents use a Kalman filter to estimate b_t. Let $\ln b_t | I^{t-1} \sim \mathcal{N}(\ln \beta_{t-1}, \sigma^2)$ be the prior belief based on information up to $t-1$ and $\sigma^2 = \frac{\sigma_b^2 + \sqrt{(\sigma_b^2)^2 + 4\sigma_b^2 \sigma_P^2}}{2}$ the steady state Kalman filter uncertainty. Optimal filtering implies that the posterior expectations when period t information is revealed are

[8] This is only one among many possible options. It is the researchers' choice, in the same vein as the form of the utility function or the particular stochastic process for fundamentals. The reason for choosing this subjective model is twofold. Theoretically, it is close to RE, sharing a similar autoregressive structure. Empirically, it is consistent with the extrapolation and sluggishness of investors' beliefs detected in surveys.

given by $\ln b_t | I^t \sim \mathcal{N}(\ln\beta_t, \sigma^2)$ with[9]

$$\ln\beta_t = \ln\beta_{t-1} + g\Big(\ln\frac{P_{t-1}}{P_{t-2}} - \ln\beta_{t-1}\Big) \tag{19}$$

and

$$g \equiv \frac{\sigma^2}{\sigma^2 + \sigma_P^2} \tag{20}$$

Hence, subjective price growth expectations are given by

$$\mathbb{E}_t^{\mathcal{P}}\Big[\frac{P_{t+1}}{P_t}\Big] = \bar{\rho}\beta_t \tag{21}$$

with $\bar{\rho} \equiv exp\Big\{\frac{\sigma^2+\sigma_P^2+\sigma_b^2}{2}\Big\}$. The last object to be determined is $\mathbb{E}_t^{\mathcal{P}}\Big[z_{t+1}\frac{P_{t+1}}{P_t}\Big]$. Given the subjective model of prices,

$$\mathbb{E}_t^{\mathcal{P}}\Big[z_{t+1}\frac{P_{t+1}}{P_t}\Big] = \mathbb{E}_t^{\mathcal{P}}\Big[z_{t+1}(\bar{\rho}\beta_t + u_t + \vartheta_{t+1} + \varepsilon_{t+1}^P)\Big] = \mathbb{E}_t^{\mathcal{P}}\Big[z_{t+1}\Big]\bar{\rho}\beta_t = \pi\bar{\rho}\beta_t \tag{22}$$

Where $u_t \sim i.i.\mathcal{N}(0, \sigma^2)$ is the Kalman forecast error. Plugging the subjective price expectations into the pricing function (15), the equilibrium PD ratio with imperfect information and learning reads as

$$\frac{P_t^L}{D_t} = \frac{\delta\bar{\beta}_t{}^D}{1 - (1-\pi\tau)\delta\bar{\rho}\beta_t - \delta\pi\tau} \tag{23}$$

assuming that $1 > (1-\pi\tau)\delta\bar{\rho}\beta_t + \delta\pi\tau$ holds. In terms of the general pricing formula, $\beta_t^M = 1$ and $\beta_t^D = \delta\bar{\beta}_t{}^D$ as with RE, but $\beta_t^P = \delta\bar{\rho}\beta_t$. Price beliefs might differ from fundamental beliefs, being an additional source of instability.

2.3.3. Volatility and taxes with endogenous beliefs

In this section, I revisit *Proposition 1* within the contexts of Rational Expectations and Learning. First, I show that *Proposition 1* holds under both setups. Then, I study the effect of taxes on the expectations variance-covariance matrix.

Lemma 3. *Let $\bar{\rho}\beta_t > 1$. Then, the sensitivity of prices to beliefs is decreasing on the tax level*

[9] As it is standard in the Learning literature, equation (19) contains lagged price growth. This assumption imparts recursivity in the model and rules out equilibria multiplicity. It can be rationalized by assuming that agents observe in period t information about the lagged transitory component ε_{t-1}^P. See Adam, Marcet, and Beutel (2017) for a discussion.

for both Rational Expectations and Learning

$$\frac{d}{d\tau}\omega < 0 \tag{24}$$

Proof. Appendix B.3.

Taking Σ as given, in both models, taxes reduce the pass-through from beliefs to price fluctuations. Then, the logic of *Proposition 1* applies, even with milder restrictions due to the additional model structure. Additionally, there is the possibility of taxes affecting Σ. In this vein, *Lemma 4* proves that taxes influence Σ only when agents learn from prices:

Lemma 4. *Assume perfect knowledge of the stochastic process governing dividend growth and the liquidity shock. On top of that, consider the following cases:*

Case 1 [Rational Expectations]: Assume rationality and full information such that the equilibrium PD ratio is given by equation (13). Then,

$$\Sigma^{RE} = \mathbb{V}(d_t) = \frac{\sigma_d^2}{1-\rho^2} \tag{25}$$

and $\frac{d}{d\tau}\Sigma^{RE} = 0$.

Case 2 [Learning]: Assume rationality, imperfect information and Learning about prices such that the equilibrium PD ratio is given by equation (23). Let ν *and* υ *be constants that depend on* τ *and assume* $g \leq \bar{g}$.[10] *Then,*

$$\Sigma^{L} = \begin{bmatrix} \mathbb{V}(ln\beta_t) & \mathbb{C}(ln\beta_t, d_t) \\ \mathbb{C}(ln\beta_t, d_t) & \mathbb{V}(d_t) \end{bmatrix} = \frac{\sigma_d^2}{1-\rho^2} \begin{bmatrix} \nu & \upsilon \\ \upsilon & 1 \end{bmatrix} \tag{26}$$

with $\mathbb{C}(ln\beta_t, d_t) > 0$, $\frac{d\nu}{d\tau} < 0$, $\frac{d\upsilon}{d\tau} < 0$.

Proof. Appendix B.3.

Σ^L comprises two components: i) the fundamental risk represented by Σ^{RE}; and ii) a matrix illustrating how this risk is propagated through fluctuations in expectations. Under RE, beliefs about future payoffs are determined solely by d_t, making fundamental risk the only source of beliefs' volatility. However, under Learning, d_t fully determines beliefs about short-term dividend

[10] The constants ν, υ and $\bar{g}$ are defined in Appendix B.3.

growth but not about capital gains, which are also influenced by their past values. This leads to an amplification of fundamental risk through the autoregression of price beliefs, resulting in additional volatility.

Taxes impact the extent of fundamental risk amplification through subjective beliefs about capital gains. Lower taxes amplify the effect of fundamental volatility and reduce the covariance between capital gains beliefs and dividend growth. This decoupling increases fluctuations in the price-dividend (PD) ratio. Consequently, *Proposition 1* can be supplemented with an additional channel through which taxes affect valuation volatility.

Proposition 2. *When there is price extrapolation, the negative effect of taxes on the variance of the Price-Dividend ratio is amplified through the negative effect of taxes on the covariance matrix of expectations:*

$$\frac{d}{d\tau}\mathbb{V}\Big[\frac{P_t}{D_t}\Big] = \frac{d\boldsymbol{\omega}}{d\tau}\boldsymbol{\Sigma}\boldsymbol{\omega}^T + \boldsymbol{\omega}\boldsymbol{\Sigma}\frac{d\boldsymbol{\omega}^T}{d\tau} + \underbrace{\omega\Big(\frac{d}{d\tau}\Sigma\Big)\omega^T}_{amplification} << 0 \tag{27}$$

This effect is absent under Rational Expectations.

Proof. By Lemma 3, $\frac{d\boldsymbol{\omega}}{d\tau}$ are negative for both Rational Expectations and Learning. Besides, by Lemma 4, $\frac{d}{d\tau}\Sigma^{RE} = 0$ and $\frac{d}{d\tau}\Sigma^{L} \leq 0$.

Altogether, when agents learn from prices, taxes reduce the variance of the PD ratio by reducing both ω^L and Σ^L. In other words, differently from RE, under Learning taxes also have an indirect effect on the variance of the PD ratio via capital gains beliefs fluctuations. These beliefs amplify the fundamental risk, with the degree of amplification inversely related to the tax level. The direct (through ω^L) and indirect (through Σ^L) tax effects reinforce each other: lower taxes increase the pass-through from beliefs to prices; more volatile prices translate into more fluctuating beliefs. Hence, under Learning, the tax level becomes a more powerful regulator of Price-Dividend volatility.

2.4. The lock-in effect

In this section, I model the behavioral response of investors to taxes, which might counteract the previous results. In particular, I let investors decide when the accumulated capital gains are realized and taxes are paid. In this case, higher taxes might lead to investors holding their assets longer than otherwise, which is known as the *lock-in effect*. This effect is a primary concern in the literature, forming the key argument for why a CGT would boost rather than dampen price

volatility (for instance, Somers, 1948; Somers, 1960; Stiglitz, 1983). Up to an approximation, this effect is dominated by the stabilizing effect outlined in Section 2.3. if the *tax elasticity of realitzation* is not too negative.

Consider the same exact model of Section 2.1., but replace the liquidity shock z_t by a choice of the timing of realization. Following Gavin, Kydland, and Pakko (2007) and Gavin, Keen, and Kydland (2015), each investor manages a stock of unrealized capital gains G_t facing accumulation costs. G follows this law of motion

$$G_t = G_{t-1} + (P_t - P_{t-1})S_{t-1} - \hat{G}_t + AC_t \tag{28}$$

with $\hat{G}_t$ are the realized capital gains, and AC_t stands for adjustment costs. It is assumed $AC_t = \hat{G}_t - G_{t-1} - \phi(\bar{\pi}_t)G_{t-1}$ for $\bar{\pi}_t \equiv \hat{G}_t/G_{t-1}$, with $\phi'(\cdot) > 0$, $\phi''(\cdot) < 0$. It follows that when the realization of capital gains is deferred, the cost function penalizes investors with extra unrealized capital gains, increasing the future tax liability of households. With these adjustment costs, investors face an additional trade-off: realize $\hat{G}_t$ capital gains out of the stock of unrealized and pay taxes $\tau\hat{G}_t$ today or delay the realization and face an extra tax liability in the future. Thus, the unrealized gains accumulation function captures the idea that gains are not deferred indefinitely, but often retimed to be realized in the medium- to long-run, as suggested in Sarin et al. (2022).

The investor's problem is as before except for two additional controls, realized $\hat{G}_t$ and unrealized G_t gains, the unrealized capital gains law of motion (28) and the budget constraint

$$C_t + P_t X_t \leq D_t S_{t-1} - \tau\hat{G}_t + T_t \tag{29}$$

Let $\phi(\bar{\pi}_t) = (\tau)^{1+\xi} ln(\bar{\pi}_t)$. This cost function has two components. First, the penalty grows with the tax level for standard values of ξ, reflecting that higher taxes make the delay more costly. Second, the penalty decreases with the amount of gains realized; on the contrary, it explodes if agents realize nothing, that is, $\lim_{\hat{G}_t \to 0} -\phi(\bar{\pi}_t) = +\infty$.

Given the assumption of risk neutrality, the first-order conditions boil down to

$$P_t = \delta \mathbb{E}_t^{\mathcal{P}} \Big[D_{t+1} + P_{t+1} - \mu_{t+1}(P_{t+1} - P_t) \Big] \tag{30}$$

$$\bar{\pi}_t = \mu_t \tau^{\xi} \tag{31}$$

$$\mu_t = \tau^{1+\xi}\mathbb{E}_t^{\mathcal{P}}\left[\delta\mu_{t+1}\Big(1 - ln(\bar{\pi}_{t+1})\Big)\right] \tag{32}$$

Equation (30) replaces equation (5). They differ only in one detail: the liquidity shock that determined the payment of taxes z_{t+1} is replaced by the Lagrange multiplier associated with the capital gains accumulation equation (28). The presence of μ points out that the additional burden unrealized capital gains represent diminishes the one-period ahead payoffs. In turn, according to equation (31), the optimal fraction of gains that agents should realize $\bar{\pi}_t$ depends on two terms: positively on the shadow price of unrealized gains μ_t indicating that the more costly non-realizing gains is the more agents realize them; and on the tax level τ. If $\xi < 0$, a higher τ lowers the optimal realization, reflecting the lock-in effect. From this equation is clear that ξ is the tax elasticity of realization

$$\frac{\partial\bar{\pi}_t}{\partial\tau}\frac{\tau}{\bar{\pi}_t} = \mu_t\xi\tau^{\xi-1}\frac{\tau}{\mu_t\tau^{\xi}} = \xi \tag{33}$$

which is a structural parameter of the model that can be directly estimated from the data. Finally, equation (32) is a first-order stochastic difference equation determining the shadow price of unrealized capital gains. Shifting it one-period ahead:

$$\mu_{t+1} = \tau^{1+\xi}\mathbb{E}_{t+1}^{\mathcal{P}}\left[\delta\mu_{t+2}\Big(1 - ln(\bar{\pi}_{t+2})\Big)\right] = \tau^{1+\xi}M_{t+1} \tag{34}$$

Then, operating on the Euler Equation (30)

$$P_t = \delta(1-\tau^D)\mathbb{E}_t^{\mathcal{P}}\left[\frac{D_{t+1}}{D_t}\right]D_t + \delta\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}(1-\tau^{1+\xi}M_{t+1})\right]P_t + \delta\tau^{1+\xi}P_t\mathbb{E}_t^{\mathcal{P}}\Big[M_{t+1}\Big] \tag{35}$$

Let $m_t \equiv \mathbb{E}_t^{\mathcal{P}}\Big[M_{t+1}\Big]$. Using the subjective model of prices,

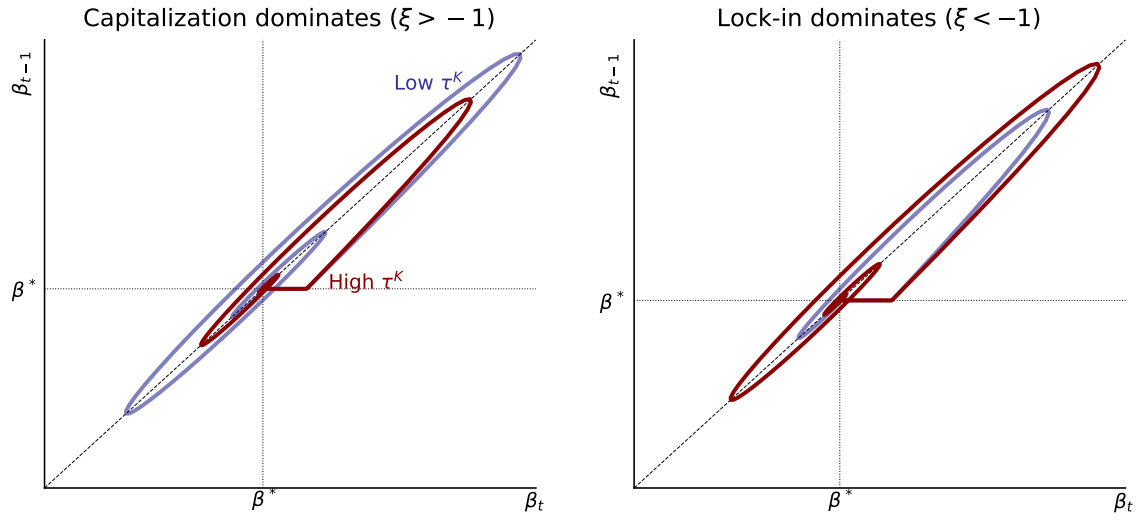
$$\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}M_{t+1}\right] = \mathbb{E}_t^{\mathcal{P}}\left[(\bar{\rho}\beta_t + u_t + \vartheta_{t+1} + \varepsilon_{t+1}^P)M_{t+1}\right] \approx \bar{\rho}\beta_t m_t \tag{36}$$

under the assumption that $\mathbb{C}ov(M_{t+1}, x_t) \approx 0$ for $x_t = (u_t, \vartheta_{t+1}, \varepsilon_{t+1}^P)$ where u_t is the Kalman prediction error. With that approximation, the PD ratio with endogenous capital gains realization satisfies

$$\frac{P_t}{D_t} = \frac{\delta\bar{\beta}_t^{\,D}}{1-(1-\tau^{1+\xi}m_t)\delta\bar{\rho}\beta_t - \delta\tau^{1+\xi}m_t} \tag{37}$$

This formula embeds the one under exogenous realization (23) for $m_t = \pi$ and $\xi = 0$. Consider m_t as given for a moment. In this case, the formula says that the lock-in effect dominates when

Figure I: ***Phase diagram illustrating the effect of taxes on volatility.***



The graph illustrates the dynamics of subjective capital gains expectations (β) when shocked by a one-off dividend shock. The blue line corresponds to a tax level of $\tau = 0.1$; the red line uses $\tau = 0.4$. In the right panel, the stabilization effect dominates, using $\xi = -0.5$; in the left panel, the lock-in effect dominantes, using $\xi = -1.2$.

$\xi < -1$, that is

$$\xi \lessgtr -1 \quad \Rightarrow \quad \frac{\partial P_t/D_t}{\partial \tau} \gtrless 0 \tag{38}$$

However, for equation (37) to be an equilibrium equation, m_t must be characterized. Since m_t is a function of the state variables, including P_t/D_t, now there is no closed form for equilibrium prices and the tax effects on PD volatility must be explored via simulations.

Figure I plots the propagation of a one-off dividend shock via capital gains beliefs under Learning. Note that the response of a capital gains beliefs to a fundamental shock under RE would be none. Hence, the graphs show the emergence of excess volatility from Learning and how the tax level regulates it. Two cases are considered. When $\xi > -1$, labeled "capitalization", lower taxes amplify the propagation of a fundamental shock. In this case, the logic of *Proposition 2.2.* goes through. Nonetheless, if $\xi < -1$, the lock-in effect dominates and lower taxes dampen the propagation, stabilizing beliefs and prices. Altogether, the model delivers both the stabilizing and destabilizing effect depending on the value of ξ.[11]

[11] When numerically computing m_t, it turns out that the influence of P_t/D_t on it is small. Thus, the intuition of the dominance of the lock-in when $\xi < -1$ appeared in expression (38) is approximately correct.

2.5. A tax on dividends

In this section, I analyze the effect of a tax on dividends τ^D on the variance of the Price-Dividend ratio. The main result is that τ^D affects the variance only through the sensitivity of prices to beliefs, without any effect on the beliefs formation process.

I start by writing a modified budget constraint

$$C_t + P_t X_t \leq (1 - \tau^D) D_t S_{t-1} + T_t \tag{39}$$

which mimics the original one (expression 2) except that includes a tax on dividends and disregards the elements related to the taxation of capital gains. Apart from this modification, consider the setup of Section 2.3. with learning about prices. Following the same procedure, the equilibrium PD ratio is given by

$$\frac{P_t}{D_t} = \frac{\delta(1-\tau^D)\mathbb{E}_t^{\mathcal{P}}\left[\frac{D_{t+1}}{D_t}\right]}{1 - \delta\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}\right]} \tag{40}$$

From this expression, *Proposition 1* and *Proposition 2* can be reexamined for τ^D.

Proposition 3. *The variance of the Price-Dividend ratio is inversely related to the dividend tax level*

$$\frac{d}{d\tau^D}\mathbb{V}\left[\frac{P_t}{D_t}\right] < 0 \tag{41}$$

However, even when agents learn from prices, τ^D does not have any influence on the expectations covariance matrix:

$$\frac{d}{d\tau^D}\Sigma^L = 0 \tag{42}$$

Proof. Appendix B.4.

This result emphasizes a distinction between taxes on dividends and capital gains: while both affect the pass-through from beliefs to prices, only taxes on capital gains influence belief fluctuations. Hence, capital gains taxes emerge as particularly influential for asset price stability.

3. Facts

In this section, I document some facts related to capital taxes and asset prices. For concreteness and data availability, I focus on the stock market. The analysis employs a long-run dataset spanning

from 1946 to 2018, partitioned into two equal halves —1946.I to 1982.II and 1982.III to 2018.IV- to highlight the substantially different behavior of stock prices in recent decades.[12] In Section 3.1., I document the decline in effective capital taxes. Then, Section 3.2. collects a comprehensive list of asset pricing statistics commonly used in the literature. Finally, Section 3.3., provides evidence about the pass-through from expectations to prices.

3.1. Facts about personal capital taxes

Taxes on personal capital income (dividends, capital gains and interests) decreased substantially over the last decades (McGrattan and Prescott (2005), Sialm (2009)). The literature has focused on the analysis of effective average marginal rates, which are value-weighted means of the marginal tax rates of investors in the various income brackets once adjusted for several features of the tax code (such as maximum and minimum taxes, partial inclusion of social security, or phaseouts of the standard deduction).[13]

The National Bureau of Economic Research (NBER) TAXSIM program by Feenberg and Coutts (1993) computes these rates. One important factor not accounted for in the TAXSIM model is the change in the market structure from taxable to nontaxable accounts following pro-savings regulatory changes (see McGrattan and Prescott, 2005). To correct for this, I compute the taxable share η as the share of equities in the hands of tax paying units following Rosenthal and Austin (2016). Hence, the effective rates are given by:

$$\boldsymbol{\tau}_t^X = \boldsymbol{\tau}_t^x (1 - \eta_t) \tag{43}$$

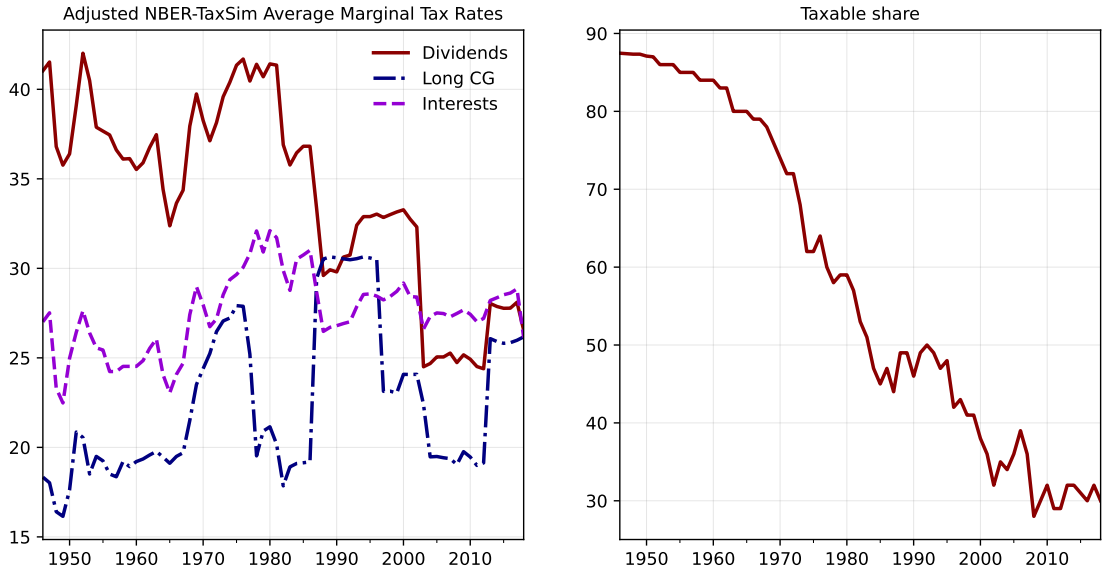
where $\boldsymbol{\tau}_t^x = (\tau_t^d, \tau_t^{lkg}, \tau_t^b)$ are the adjusted NBER rates for dividends, long-term capital gains and interests including federal and state-level rates. Details on the data sources and adjustments are available in Appendix A; computation details on the nontaxable share are available in Appendix D.

As illustrated in figure II, taxes exhibited a substantial variation throughout the period. In particular, since the 1980s the different rates tended to decrease with the introduction of significant reforms such as President Carter 1978' and President Clinton 1997' cuts to the capital gains tax,

[12] The latter half largely aligns with the legalization of stock repurchases in 1982 and marks the commonly accepted onset of the Great Moderation in 1984. In 1982, the Securities and Exchange Commission's Rule 10B-18 provided a safe harbor for stock repurchases under the Rule´s conditions. See `this file` for a summary. With this change, stock buybacks began to grow substantially, becoming the dominant form of corporate payout for the SP&500 firms since 1997.

[13] These rates are provided by the TAXSIM program of the NBER and can be accessed on his website. Before 1960, τ_t^d, τ_t^{skg} and τ_t^{lkg} rates were taken from Sialm (2009). See Appendix A for details.

Figure II: ***Capital taxes in the postwar period.***



The left graph plots the effective average marginal tax rates on dividends τ_t^d (red), long capital gains τ_t^{lkg} (blue) and interest income τ_t^b (violet). The right panel plots the taxable share of equities. Annual series 1946-2018. See Appendix A for data sources and Appendix D for details on the computations.

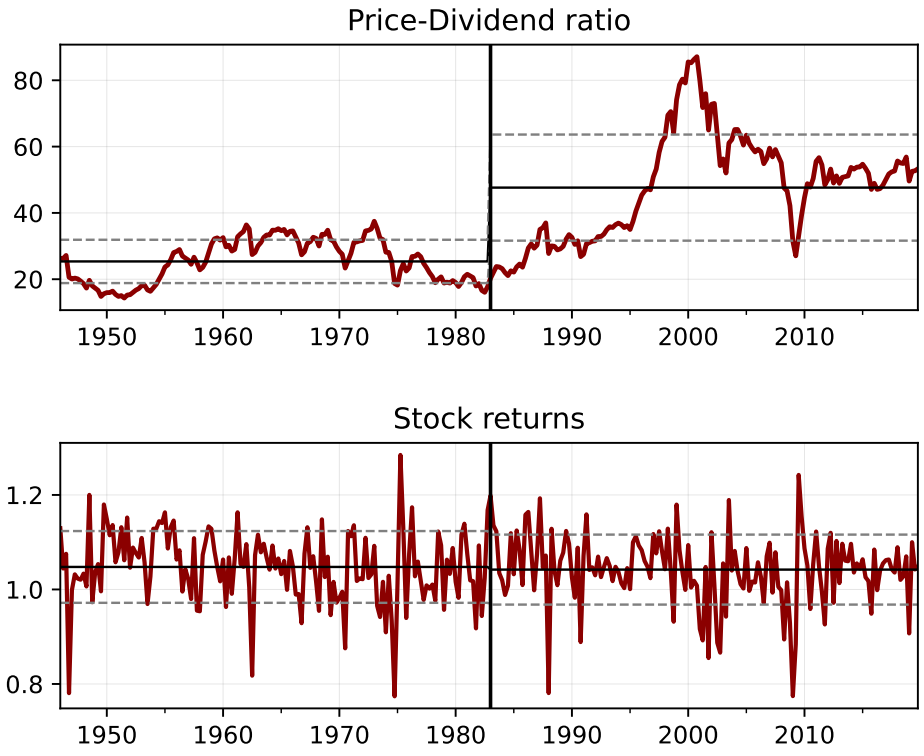
partially counteracted by President Reagan 1986's reform, or sharp dividend tax cuts introduced by President Reagan in 1982 and President Bush in 2001 and 2003. In addition, the taxable share fell remarkably from 87% in 1946 to just 30% in 2018, reflecting a fundamental change in the market structure toward tax-free pension accounts and institutional investors.

3.2. Facts about the stock market

Stock prices exhibited substantially different behavior since the 1980s. Here I summarize the change in many statistics covering the main topics in asset pricing: excess volatility, predictability, valuations levels and the equity premium. The evidence is summarized in the following facts, which are collected in table I.

Fact 1: Lack of a Great Moderation in the stock market. The unconditional variance of the log PD ratio went up by a factor of 2 since 1982. A Campbell and Shiller (1988)'s decomposition reveals that this rise can be attributed to two main factors. First, there has been a shift in the correlation between valuations and future dividend growth. Before 1982, the comovement was weakly negative; subsequently, it became stronger and positive. In other words, returns and dividend growth, which moved in tandem before the 1980s, became decoupled afterward. This shift accounts for 47% of

Figure III: ***Valuations and returns in the postwar period.***



The graph plots the evolution of the PD ratio and real stock returns in the 1946:I-2018:IV period. The continuous lines plot the mean of each subperiod 1946:I-1982:II and 1982:III - 2018:IV. The dotted bands show ± one standard deviation.

the increase in the variance of the PD ratio. The remaining increase is due to a larger covariance between prices and future returns. A related observation is that, despite the reduction in the dividend growth volatility, stock return volatility remained largely stable.

Fact 2: A fall in excess return predictability. The well-known predictive power of the PD ratio to explain excess return went down after 1982. Consider the following projection:

$$xr_{t,t+h} = c_0 + c_1^h \frac{P_t}{D_t} + u_t^{xr} \tag{44}$$

with

$$xr_{t,t+h} = \prod_{j=1}^{h} R_{t+j} - \prod_{j=1}^{h} R_{t+j}^b \tag{45}$$

R_t^b being the gross real return on 90 days T-bills and $h = 20$. It turns out that the magnitude of c_1^h dropped from -0.07 pre-1980s to just -0.02 post-1980s and the R^2 fell from 0.65 to 0.35. Cochrane (2009) showed that the high predictability at longer horizons is related to the persistence of the

PD ratio, so perhaps lower predictability comes from lower persistence. As shown in Table I, this is not the case; persistence is statistically the same across the two subsamples.

Fact 3: A surge in capital gains. The mean PD ratio almost doubled after 1982. This fact has been extensively documented in the literature -see, for example, Shiller (2000) and McGrattan and Prescott (2005)- and is illustrated in figure III. In accounting terms,

$$\frac{P_{t+n}/D_{t+n}}{P_t/D_t} = exp\Big\{\Delta \ln P_{t+n} - \Delta \ln D_{t+n}\Big\} \tag{46}$$

Hence, the accounting reason for a higher PD ratio is that the increase in price growth exceeds the dividend growth. Indeed, table I shows that the average price growth was 7.35% after 1982 from 1.57% before 1982 while dividends grew only 3.18% after from 1.95% before 1982.

Fact 4: A higher equity premium. The historical average of the difference between the real stock return and the real 90-day T-bill increased substantially after the 1980s, moving from an average of 6.32% to 8.92%. The main reason behind the higher equity premium is a boost in stock returns, largely due to the sharp increase in capital gains. The risk-free rate followed an upward trajectory from low to high rates since the 1950s to the early 1980s and a reverse downward path from high to low rates afterward, exhibiting a very similar average magnitude in both periods.

3.3. Facts about the mechanism

A key insight of the model was that low taxes generate more volatility by increasing the pass-through from beliefs to prices. In this section, I measure this pass-through using survey-like data. As finding investors' survey data going back to 1946 is problematic, I follow Adam, Marcet, and Beutel (2017) to extend the UBS Gallup survey for the 1946-2018 period using the following algorithm:

$$\ln \mathbb{E}_t^s\Big(\frac{P_{t+1}}{P_t}\Big) = (1-g)\ln \mathbb{E}_{t-1}^s\Big(\frac{P_t}{P_{t-1}}\Big) + g \ln \frac{P_t}{P_{t-1}} \tag{47}$$

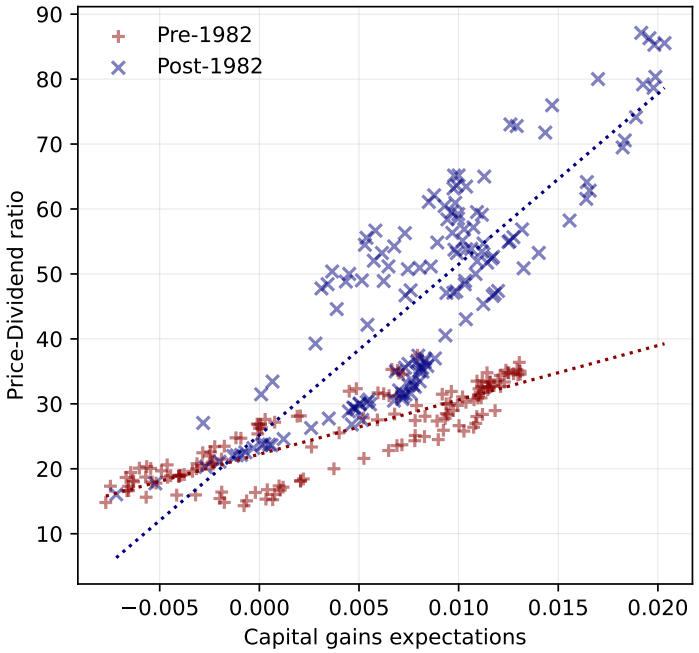
where $\mathbb{E}^s$ stands for survey expectations, P_t is the real S&P500 price, and g is estimated using non-linear least squares to minimize the mean square distance between the survey and the algorithm-implied data points. The point estimate is $\hat{g} = 0.0276$ and the initial condition is set equal to zero growth. This equation is consistent with applying a Kalman filter over an unobserved component model. It turns out to replicate the survey data points very well, exhibiting a correlation of 0.94 in the period that both series overlapped (1998.I to 2007.IV). The approach of generating survey-like data by projecting existing surveys on other variables has also been used in Nagel and Xu (2022).

Table I: ***Stock pricing facts: 1946-1982 versus 1982-2018.***

		1946-1982	1982-2018
	Fact 1: Lack of a Great Moderation		
Volatility of the PD ratio	$\mathbb{V}(\ln P_t/D_t)$	7.53	13.10
		(1.55)	(3.40)
Covariance PD - future dividends	$\mathbb{C}(\ln P_t/D_t, \bar{d}_t)$	-0.80	2.20
		(0.22)	(0.52)
Covariance PD - future returns	$\mathbb{C}(\ln P_t/D_t, \bar{r}_t)$	-8.33	-10.90
		(1.76)	(2.90)
Stock returns volatility	$\sigma(r_t^s)$	7.30	7.27
		(0.68)	(0.82)
Dividend growth volatility	$\sigma(D_t/D_{t-1}-1)$	2.52	1.97
		(0.41)	(0.35)
	Fact 2: A fall in excess-return predictability		
5y Regression coefficient	c^{20}	-0.07	-0.02
		(0.01)	(0.00)
5y coefficient of determination	R^2	0.65	0.35
		(0.11)	(0.11)
Persistence PD ratio	$corr(PD_t, PD_{t-1})$	0.96	0.98
		(0.13)	(0.07)
	Fact 3: A surge in capital gains		
Price-Dividend ratio	$\mathbb{E}(PD_t)$	25.27	46.78
		(1.57)	(3.98)
Dividend growth	$\mathbb{E}(D_t/D_{t-1}-1)$	1.95	3.18
		(0.35)	(0.34)
Stock price growth	$\mathbb{E}(P_t/P_{t-1}-1)$	1.57	7.35
		(1.32)	(1.28)
	Fact 4: A higher equity premium		
Real stock returns	$\mathbb{E}(r_t^s)$	5.96	9.98
		(1.35)	(1.32)
Equity premium	$\mathbb{E}(r_t^s - r_t^b)$	6.32	8.92
		(1.26)	(1.26)
	Fact 5: Higher pass-through from beliefs and shocks to prices		
Correlation of prices and survey expectations	$corr(PD_t, \ln\beta_t)$	0.75	0.72
		(0.07)	(0.11)
Pass-through from beliefs to prices	$\tilde{\omega}$	0.54	1.61
		(0.07)	(0.30)
Response of the PD ratio to a shock	$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	0.25	0.44
		(0.11)	(0.09)

This table reports a list of statistics using the data sources described in Appendix A. The variables are in real terms. Growth rates and returns are annualized. Newey-West standard errors are in parentheses, except for the IRFs that are obtained via bootstrapping with 1000 repetitions. $\bar{d}_t = \sum_{j=1}^{\infty} \hat{\rho}^{j-1} \Delta lnD_{t+j}$ *and* $\bar{r}_t = \sum_{j=1}^{\infty} \hat{\rho}^{j-1} r_{t+j}^s$ *come from a standard Campbell-Shiller decomposition.*

Figure IV: ***The pass-through from expectations to valuations.***



The graph plots the quarterly PD ratio of the S&P500 against the log real expected price growth consistent with the UBS Gallup survey. The periods are 1946:I–1982:II and 1982:III–2018:IV.

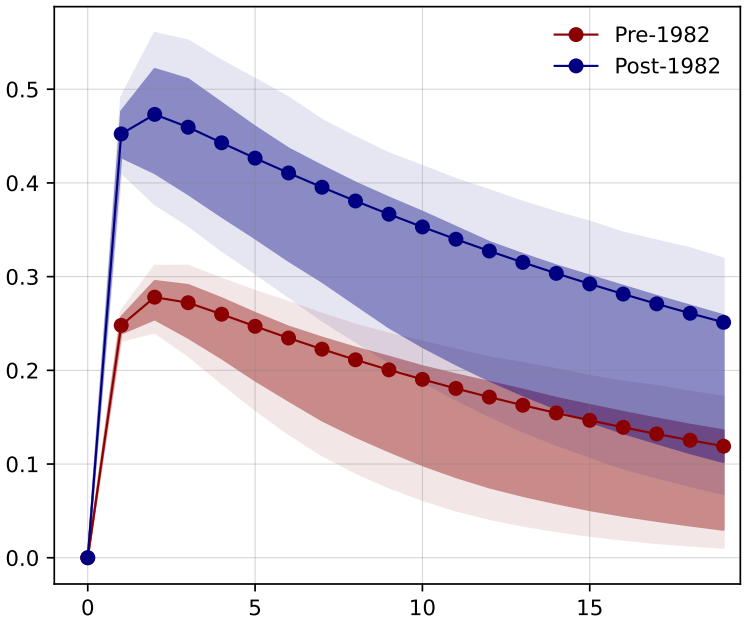
Furthermore, a constant gain rule has proven to be a parsimonious and reasonable way to model other expectations series (see, for example, Malmendier and Nagel, 2016) and is a robust optimal filter (Evans, Honkapohja, and Williams, 2010). Appendix C shows that this series of long-run capital gains beliefs tracks existing long-run survey data very well, such as the Livingston Survey from the Federal Reserve Bank of Philadelphia, the LLM-generated beliefs of Bybee (2023), the long-term earning growth forecasts from IBES or the Bullish-Bearish spread from AAII. In other words, capital gains expectations generated from an adaptive algorithm rationalize not only direct measures of capital gains beliefs, but also measures of investors' sentiment and long-term expected earnings.

With this long series of survey-like expectations, I run the following regression for each subperiod

$$PD_t = \alpha + \tilde{\omega} \ln \mathbb{E}_t^s \left(\frac{P_{t+1}}{P_t} \right) + \varepsilon_t^s \tag{48}$$

Similarly to the finding in Adam, Marcet, and Beutel (2017), figure IV illustrates that $\tilde{\omega}$ went up substantially after the 1980s. Thus, consistent with the model predictions, the pass-through from beliefs to valuations was three times larger on average after the 1980s.

Figure V: ***IRF: PD ratio response to a price growth shock.***



The graph plots the response function of the PD ratio to a price growth shock at a quarterly frequency. Confidence bands at 68% and 95% levels are shown, coming from a bootstrap procedure with 1000 repetitions.

In the model, subjective beliefs are a mechanism that amplifies the propagation of fundamental shocks. Thus, an alternative measure of the pass-through that does not require measuring expectations is the response of valuations to structural shocks. In particular, let the response of the PD ratio to a shock ε_t^P at the horizon h defined as

$$\mathbb{E}(PD_{t+h}|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+h}|\varepsilon_t^P = 0, I_t) \tag{49}$$

where ε_t^P is a price growth shock and I_t is the information set when the shock hits the system. I use a minimalist VAR with price growth and the PD ratio. The VAR is estimated following the approach of Kilian and Vigfusson (2011). As shown in figure V, the response is significantly larger in the second subsample in the short and medium term. For instance, at the one-year horizon, the point estimate for the response is 75% higher in the post-1982 sample.

4.- A Quantitative Model

In this section, I introduce a quantitative version of the model and empirically validate it. Section 4.1. sets up the extensions and introduces a novel application of the Parameterized Expectations Algorithm to solve the model. In Section 4.2., I discuss the parameterization, which includes an estimation via the Simulated Method of Moments and a sensitivity analysis for the main parameters. Section 4.3. presents the estimation results and formal tests of the model ability to match the data statistics. Section 4.4. discusses the differences between a tax system on realization vs. on accrual.

4.1. Model extension and solution

The model is extended in five dimensions. First, the assumption of risk neutrality is replaced by a constant relative risk aversion regulated by the parameter γ. Second, an additional source of exogenous income is introduced to avoid an unrealistically too-high correlation between dividends and consumption at odds with the data. Following Adam, Marcet, and Beutel (2017), agents get a wage endowment W_t each period that follows this process:

$$\ln\Big(1+\frac{W_t}{D_t}\Big) = (1-p)\ln(1+w) + p\ln\Big(1+\frac{W_{t-1}}{D_{t-1}}\Big) + \varepsilon_t^w \tag{50}$$

D_t are aggregate dividends, $1+w$ is the average wage-dividend ratio and $p \in [0,1)$ is its quarterly persistence. Innovations are normally distributed with mean zero and variance σ_w^2.

Third, taxes on interest income τ^B are also included and all capital taxes are assumed to be random variables rather than constants. Fourth, investors know the wage process (equation (50)) but have imperfect knowledge about prices, dividends and taxes. As before, they use subjective models to forecast these variables learning from past data such that

$$\mathbb{E}_t^{\mathcal{P}}\Big[\frac{D_{t+1}}{D_t}\Big] = \beta_t^d exp\{\frac{\sigma^2+\sigma_d^2+\sigma_{\tilde{d}}^2}{2}\} \tag{51}$$

and

$$\mathbb{E}_t^{\mathcal{P}}\Big[\boldsymbol{\tau}_{t+1}\Big] = \hat{\boldsymbol{\tau}}_t \tag{52}$$

where $\boldsymbol{\tau} = [\tau, \tau^D, \tau^B]'$. To avoid the proliferation of free parameters related to the learning process, I restrict the steady-state Kalman gain to be the same across all the variables. Appendix F contains the details of the subjective models of dividends and taxes and the derivations of the subjective conditional expectations.

Fifth, risk-free bonds with a gross rate return R_t^b are introduced such that investors can choose between stocks and bonds to store value. Finally, stock buybacks are included in the form of negative supply shocks

$$\bar{S}_t = 1 - \varepsilon_t^s \tag{53}$$

where $\bar{S}_t$ represents outstanding shares. For the simulations, $\{\varepsilon_t^s\}$ is given by the empirical time series of stock repurchases as a percentage of the total SP500 valuations.

A recursive solution via the Parameterized Expectations Algorithm. A recursive solution boils down to a time-invariant stock demand function $S_t = S(\boldsymbol{X}_t)$, where $\boldsymbol{X}_t$ is the vector of state variables. The main difficulty in deriving such an invariant function is that there are 3 optimality conditions -for stocks, bonds, and unrealized capital gains- that involve unknown subjective conditional expectations. Let $\mathcal{E}(\boldsymbol{X}_t)$ be the vector of conditional expectations. I approximate it using the Parameterized Expectations Algorithm (PEA), originally proposed by Marcet (1988), which replaces $\mathcal{E}(\boldsymbol{X}_t)$ with some parametric functions $\boldsymbol{\psi}(\boldsymbol{X}_t; \chi)$. Building on homotopy, the main idea is to approximate the consumption policy function using a function rooted in economic theory. In particular, in the spirit of Hakansson (1970), the consumption policy is linear in wealth, with the propensity to consume depending on after-tax return expectations. A stock demand function S_t^d can be then derived from the budget constraint and equilibrium prices can be obtained using the equity market clearing condition:

$$S_t^d = S^d\left(\frac{P_t}{D_t}, \cdot\right) = \bar{S}_t \tag{54}$$

Because of the handy approximating functions, this equation can be analytically solved for prices such that

$$\frac{P_t}{D_t} = \frac{\chi_1 \delta (1 - \hat{\tau}_t^D) \beta_t^d}{\bar{S}_t - \chi_1 \delta (1 - \chi_2 \hat{\tau}_t^{1+\xi}) \beta_t} \left(1 + \frac{W_t}{D_t}\right) \tag{55}$$

where χ are parameters of the approximating functions. See Appendix E for all the details about the PEA implementation and accuracy measures.[14]

4.2. Simulated Method of Moments estimation

In this section, I discuss the parameterization of the model. Let θ be the vector of p parameters. I divide it into two subvectors: $\widetilde{\theta}$ contains $(p-k)$ parameters whose value is taken from the data and

[14] As it is common in similar models, the formula has a discontinuity. In this case, the literature often resorts to a projection facility that restricts beliefs to ensure non-negative and non-explosive prices (see, for instance, Adam, Marcet, and Nicolini, 2016). Appendix F contains the details.

$\bar{\theta}$ contains the remaining k parameters, estimated through an extension of the Simulated Method of Moments, following Adam, Marcet, and Nicolini (2016). Hence, a total of k parameters are estimated to match a set of M statistics with $k << M$. In the baseline application, $k = 5$ and $M = 18$. Let $\hat{\theta}$ be the estimate of the vector $\bar{\theta}$. $\hat{\theta}$ minimizes the distance between the model $\widetilde{\mathcal{S}}(\bar{\theta})$and the empirical data statistics $\hat{\mathcal{S}}$, that is,

$$\hat{\theta} = \arg\min_{\bar{\theta}\in\Theta} \left[\hat{\mathcal{S}} - \widetilde{\mathcal{S}}(\bar{\theta})\right]' \hat{W} \left[\hat{\mathcal{S}} - \widetilde{\mathcal{S}}(\bar{\theta})\right] \tag{56}$$

where $\hat{\mathcal{S}}$ and $\widetilde{\mathcal{S}}(\bar{\theta})$ are $M \times 1$ vectors, $\hat{W}$ is a $M \times M$ weighting matrix and Θ is a compact subset of $\mathbb{R}^N$ that contains the true parameter vector θ_0 in its interior. The weighting matrix is the diagonal of the inverted Newey-West variance-covariance matrix of the observed statistics $\hat{\mathcal{S}}$. To compute $\widetilde{\mathcal{S}}(\bar{\theta})$, I follow a Montecarlo procedure as in Adam, Marcet, and Nicolini (2016) for tacking wage uncertainty while I feed the model with the empirical time series for dividend growth, capital taxes and stock buybacks.

Table II contains θ. First, I discuss the parameters in the subvector $\tilde{\theta}$. The risk aversion parameter is set to 1 for reasons related to the model solution algorithm. Similar papers in the literature typically use a value around 2 (for example, Campbell and Cochrane, 1999; Adam, Marcet, and Beutel, 2017), so this lower value, if anything, highlights the model's ability to generate high returns for reasons other than risk aversion. The persistence of the wage-dividend is 0.99, which is the empirical value for the 1946-2018 period. The vector $\bar{\theta}$ is estimated using a total of $M = 18$ statistics (listed in Table III). In unrestricted estimations, $\delta > 1$ which violates standard assumptions that place it in the $(0, 1)$ interval. Hence, I fix $\delta = 0.995$, a value close but below 1, and proceed to estimate the other five parameters.

The estimated parameters in the subvector $\bar{\theta}$ are collected in the bottom part of II. A few comments are pertinent. First, σ_w is the only parameter that is allowed to change from one subperiod to the other; apart from it, all the differences in the model statistics from one period to another are driven by the exogenous empirical time series of dividends, buybacks and taxes together with the model's endogenous mechanisms. The fact that σ_w increases is consistent with the empirical data; the point estimates of these deviations are 0.0307 and 0.0520 respectively. Beyond that, σ_w is used freely to match the data as good as possible, encapsulating other motives leading to higher volatility. Second, the estimated Kalman gain is $\hat{g} = 0.0287$ which is virtually identical to the value obtained using surveys, 0.0276. This fact highlights that survey-like capital gains expectations are capable of explaining salient features of stock pricing data, including the

Table II: ***Baseline parameters.***

Parameters	θ	
Non included in SMM	$\tilde{\theta}$	
CRRA parameter	γ	1
Persistence wage-dividend ratio	p	0.99
Discount factor	δ	0.995
Included in SMM	$\bar{\theta}$	
Average wage-dividend ratio	w	13.86
Standard deviation WD shock (pre-1982)	σ_W	0.0543
Standard deviation WD shock (post-1982)	σ_W	0.0649
Kalman gain	g	0.0287
Tax elasticity of realization	ξ	-0.51

The table reports the value of the parameters for the baseline version of the model with $M = 18$.

significant changes after the 1980s.

Third, the baseline estimate of the tax elasticity of realization is $\hat{\xi} = -0.51$. This number is in the lower bound of the $[-0.5, -0.3]$ interval found by Agersnap and Zidar (2021). This finding seems remarkable as their estimates come from micro-data on gains realization while my estimation aims at explaining stock market volatility from an aggregate perspective. A way of interpreting the similarity between $\hat{\xi}$ and empirical estimates is this: the value for ξ that best serves the model to replicate the data is very close to empirical estimates, suggesting that the model is a good representation of reality. The estimate is somewhat lower in absolute value than the elasticity used by the Joint Committee on Taxation, which is -0.70 (Gravelle, 2020) and the estimate obtained by Dowd, McClelland, and Muthitacharoen (2015), -0.72. This lower elasticity might be related to a wider time window that captures medium-term realizations that were only locked in in the short term (Sarin et al., 2022).

Sensitivity of the parameters to SMM moments. A natural question is how sensitive is $\hat{\theta}$ to the different statistics, an issue sometimes framed as which statistics drive the parameters identification. Andrews, Gentzkow, and Shapiro (2017) proposed a *sensitivity* statistic Λ, that measures the reaction of $\hat{\theta}$ to perturbations of different statistics. An estimator of this parameter is given by

$$\hat{\Lambda} = -(\hat{J}(\hat{\theta})'\hat{W}\hat{J}(\hat{\theta}))^{-1}\hat{J}(\hat{\theta})'\hat{W} \tag{57}$$

where $\hat{J}(\hat{\theta}) = \frac{\partial[\hat{\mathcal{S}} - \tilde{\mathcal{S}}(\hat{\theta})]}{\partial\hat{\theta}}$ is the Jacobian of the moments used in SMM evaluated at the estimated

parameter vector and $\hat{W}$ is the weighting matrix. Based on that measure, I compute the following statistic:

$$\dot{\theta} = \hat{\theta} \pm \hat{\Lambda}\sqrt{\hat{\Sigma}_{\mathcal{S}}} \tag{58}$$

where $\sqrt{\hat{\Sigma}_{\mathcal{S}}}$ is the Newey-West standard error of the data statistics. $\dot{\theta}$ provides bands for the effect on the parameters of a $\pm$ one standard deviation perturbation in each statistic.[15] If statistic j is particularly relevant for parameter p, the interval $\dot{\theta}_k$ would be larger.

Figure VI plots $\dot{\theta}$ for the less conventional parameters ξ and g. Consider the effect of a one standard deviation perturbation on $\tilde{\omega}$ in the pre-1982 period due, for instance, to expectations measurement error. The sensitivity statistic for $\tilde{\omega}$ in the pre-1982 period is equal to 0.0621. Thus, if $\tilde{\omega}$ is one standard deviation overstated, $\hat{\xi}$ would be asymptotically biased upward, with the "right" estimate being -0.5710. By contrast, if $\tilde{\omega}$ is underestimated, the "right" $\hat{\xi}$ would be -0.4468. The interval $[-0.5710, -0.4468]$ delimits sensitivity bands of the elasticity ξ with respect to the statistic $\tilde{\omega}$. Consider this other example: A reader thinks that the return volatility after the 1980s is poorly estimated and fears that this mismeasurement might crucially affect the estimated Kalman gain. According to $\dot{\theta}$, it turns out that the sensitivity of g to return volatility after the 1980s is -0.00004, so it barely has an effect on $\hat{g}$.

In general, ξ is much more sensitive than g to perturbations in the statistics. In the most extreme case, a one-standard deviation perturbation to one of the elements of the Campbell-Shiller decomposition might move $\hat{\xi}$ from -0.37 to -0.65. Besides, the statistics related to the model main mechanism, $\tilde{\omega}$ and the IRF coefficient, are particularly influential for identifying the parameters.

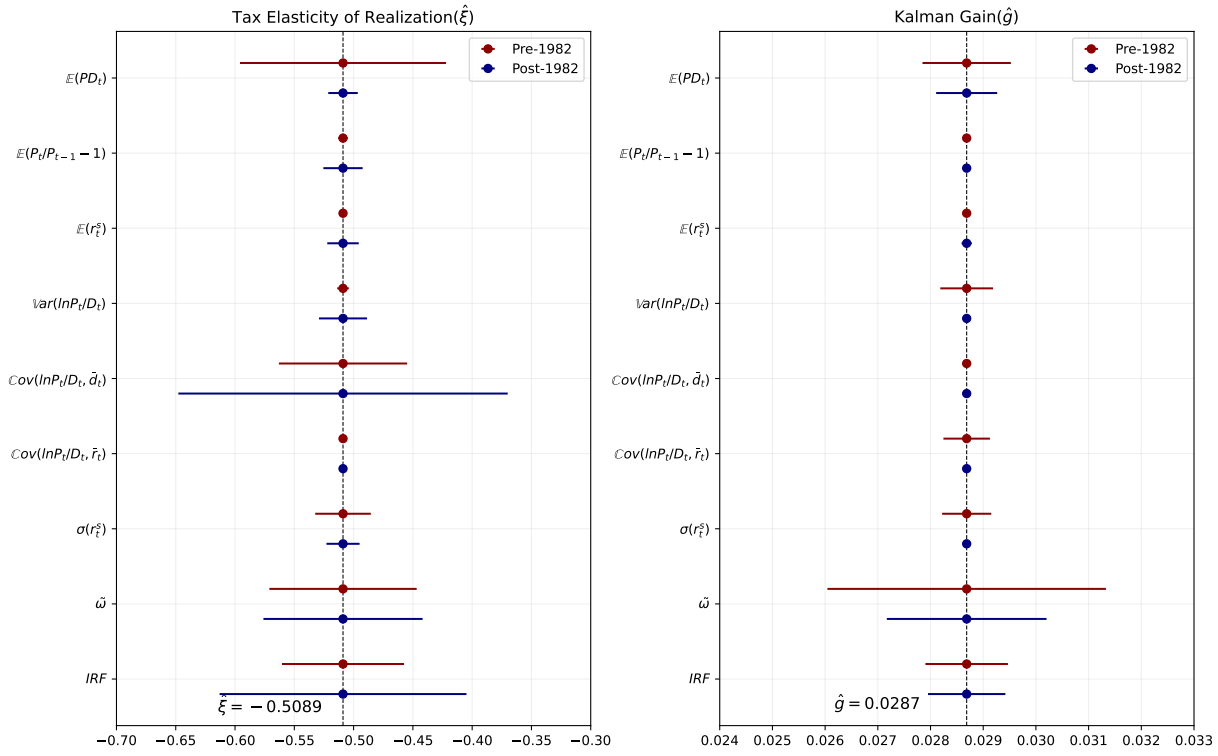
4.3. Model evaluation

This section assesses the model's ability to replicate the given set of facts. This analysis is a two-step process. I start by evaluating the model's performance with the subset of statistics included in the SMM. Then, I examine its performance with the subset of statistics not included in the estimation, which serves as an additional out-of-estimation check. To measure the similarity between model and data statistics, I conduct t-tests for each statistic j with the null hypothesis j with the null $H_0 : \hat{\mathcal{S}}_j = \widetilde{\mathcal{S}}_j(\bar{\theta})$.

Table III presents the estimation results, providing a comprehensive overview of the model's performance. A remarkable observation is the model's ability to pass almost all individual t-

[15] Note Andrews, Gentzkow, and Shapiro (2017) theory works for different local perturbations. I use a one standard deviation as that ensures comparable magnitudes for all the different statistics.

Figure VI: ***Sensitivity of the tax elasticity of realization*** ξ ***and the Kalman gain*** g ***to the SMM statistics.***



This figure reports the statistic $\dot{\theta}$ *as defined in equation (58), reflecting the reaction of each parameter to a* $\pm$ *one standard deviation perturbation of each statistic. The statistics are defined in Table I.*

tests, with t-statistics generally below 2, indicating a robust fit to the data. The model delivers a significant increase in the mean PD ratio, associated with a sharp rise in mean capital gains. Similarly, the model matches both the level and increase in stock returns. Furthermore, the model effectively captures both the variance of the PD ratio and the Campbell-Shiller decomposition. In the first subperiod, the PD ratio closely mirrors the movements of future returns, whereas the dynamics of dividends remain largely uncorrelated. In the second subperiod, the model replicates well the increase in the overall variance of the PD ratio due to the transition from negative to positive covariance between the PD ratio and $\bar{d}_t$, and a more pronounced negative covariance between the PD ratio and $\bar{r}_t$. Additionally, the model successfully estimates the level of stock return volatility, which, despite a mild increase, remains within the empirical confidence bands.

Besides the statistics related to stock pricing, the model does a good job in matching the moments associated with its main mechanism. Thus, it delivers a correct magnitude for the sensitivity of prices to beliefs $\tilde{\omega}$ as well as a notable increase after 1982. The same is true for the IRF coefficient.

Beyond the statistics incorporated into the SMM procedure, the model exhibits several remarkable features. It generates a considerable reduction in the predictability of five-year excess stock returns, specially in terms of c^{20}. However, the magnitudes of c^{20} and R^2 fall out of the confidence intervals, illustrating how -despite the connection between volatility and predictability- quantitatively one does not perfectly imply the other. The serial correlation of the PD ratio is also very high, as in the data, a fact related to the persistence of expectations derived from the random-walk structure in the subjective models of prices, dividends, and taxes. Additionally, capital gains expectations are strongly pro-cyclical with a correlation of about -0.70 as in the data.

Finally, the model also yields an adequate equity premium of around 6% as well as an increase of 2pp in recent decades, while using realistic income processes, a non-negative discount factor and a low risk aversion coefficient, which are the elements of the risk premium puzzle identified by Cochrane (2017). Although macrofinance models typically need either a high risk aversion level or high fundamental volatility to match mean returns, this model resorts to two alternative forces, combining the insights from Adam, Marcet, and Nicolini (2016) and McGrattan and Prescott (2003). In line with the Learning literature, non-fundamental volatility coming from subjective beliefs makes compatible realistic income processes with high and volatile stock returns. However, belief volatility alone cannot account for the entire effect (Adam, Marcet, and Nicolini, 2016, Adam, Marcet, and Beutel, 2017). Thus, the introduction of tax cuts helps to get an increase in the PD ratio level, generating extra capital gains that increase mean returns. Appendix G contains a more detailed analysis of the issue, including a decomposition of stock returns.

Table III: ***Quantitative performance of the model.***

	US data		Model			
	1946-1982	1982-2018	1946-1982		1982-2018	
	$\hat{S}_j$	$\hat{S}_j$	$\widetilde{S}_j(\hat{\theta})$	t-stat	$\widetilde{S}_j(\hat{\theta})$	t-stat
	Included in the SMM estimation					
$\mathbb{E}(PD_t)$	25.48	47.09	27.04	-1.13	38.04	2.20
$\mathbb{E}(P_t/P_{t-1}-1)$	1.57	7.35	3.04	-0.55	5.79	0.58
$\mathbb{E}(r_t^s)$	5.96	9.98	7.59	-0.57	9.36	0.22
$\mathbb{V}(\ln P_t/D_t)$	7.53	13.10	7.85	-0.20	13.21	-0.03
$\mathbb{C}(\ln P_t/D_t, \bar{d}_t)$	-0.80	2.20	-0.43	-1.68	2.03	0.32
$\mathbb{C}(\ln P_t/D_t, \bar{r}_t)$	-8.33	-10.90	-8.28	-0.03	-11.17	0.09
$\sigma(r_t^s)$	7.30	7.27	6.45	1.25	7.63	-0.44
$\tilde{\omega}$	0.54	1.61	0.58	-0.66	1.18	1.46
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P=1, I_t)-\mathbb{E}(PD_{t+4}\|\varepsilon_t^P=0, I_t)$	0.25	0.44	0.30	-0.45	0.48	-0.46
	Non-included in the SMM estimation					
c^{20}	-0.07	-0.02	-0.06	-0.77	-0.03	3.53
R^2	0.65	0.35	0.36	2.61	0.25	0.93
$corr(PD_t, PD_{t-1})$	0.96	0.98	0.97	-0.67	0.97	-0.34
$\mathbb{E}(r_t^s - r_t^b)$	6.32	8.92	6.72	-0.32	8.85	0.05
$corr(PD_t, \beta_t)$	0.75	0.72	0.65	1.39	0.77	-0.49
$\mathbb{E}(D_t/D_{t-1}-1)$	1.95	3.18	1.95	0.00	3.18	0.00
$\sigma(D_t/D_{t-1}-1)$	2.52	1.97	2.52	0.00	1.97	0.00

This table reports the M statistics included in the SMM estimation. The first two columns report the observed statistics for the US data. The next four columns report model-implied statistics and their t-statistics using table II's parameters. Rates of growth and returns are annualized. Variances and covariances have been multiplied by 100.

4.4. A comment about the on realization vs. on accrual system

In this section, I compare the CGT on realization with the simpler approach taken in the literature. The exercise boils down to compare the effective tax rate on capital gains τ_t with the Lagrange multiplier attached to the unrealized gains accumulation equation (28), μ_t. As it is clear from the Euler equation for stocks in a system with accumulation costs -expression (30), an on accrual system imposes $\mu_t = \tau_t$. In other words, the price of unrealized capital gains simply boils down to the effective tax rate. With accumulation costs, this equivalence is generally false, as μ_t needs to satisfy the optimality condition (32), which includes considerations about expected future rates, shadow prices, realizations and the stock of unrealized gains. In fact, the difference equation determining μ_t involves a conditional expectation that depends on the state variables.

How different is $\{\mu_t\}$ from $\{\tau_t\}$? It turns out that $corr(\mu_t, \tau_t) = 0.65$ on average across the $K = 1000$ simulations. This number suggests that while the two are obviously related, their dynamics are not fully equivalent. Thus, a rule of thumb consisting in dividing τ_t by a certain constant would fail to capture certain dynamics of an on realization system. Illustrating this idea, figure VII plots some simulations of $\{\mu_t\}$ and the observed time series for CGT, $\{\tau_t\}$.

5. The effects of tax cuts

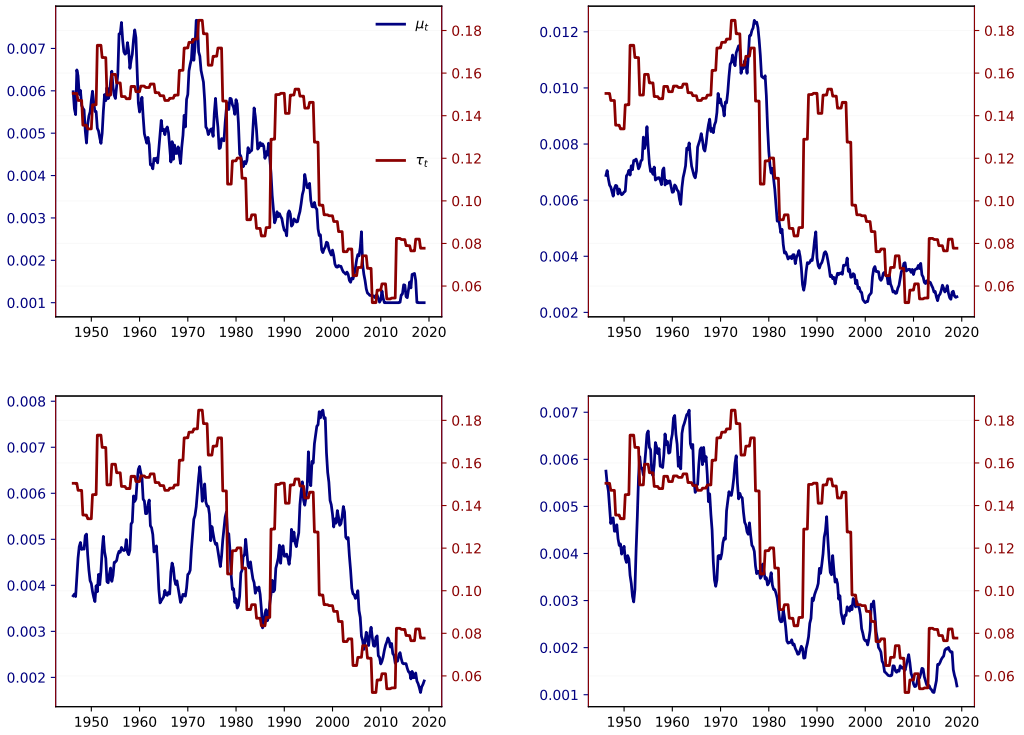
In this section, the effect of tax cuts on the stock market is evaluated. Leveraging on the good quantitative performance of the model, I use it to run experiments that quantify the effect of tax cuts on different statistics. In Section 5.1., I measure the relative average effect of tax cuts. Section 5.2. explores the sensitivity of the relative average effect of tax cuts to key parameters and assumptions. In Section 5.3., I present a calibration exercise. Finally, Section 5.4. studies the consequences of tax cuts for the informational efficiency of stock valuations.

5.1. The relative average effect of tax cuts

In this section, I introduce a measure of the effect of tax cuts on stock market valuations. Leaning on the good quantitative properties of the model, I use counterfactual simulations of scenarios without tax cuts.

I use the following procedure: Let $\boldsymbol{\tau} = [\tau, \tau^d]$ and fix $\boldsymbol{\tau}_t$ at a particular level $\bar{\boldsymbol{\tau}}$ for all the years from 1982 to 2018; Simulate the model under these alternative tax path and collect the statistics $(\widetilde{S}_j(\hat{\theta})^{\text{post}} | \boldsymbol{\tau}_t = \bar{\boldsymbol{\tau}})$; Compute the difference between the baseline model statistic and the statistic for the constant vector of taxes -that is, $\widetilde{S}_j(\hat{\theta})^{\text{post}} - (\widetilde{S}_j(\hat{\theta})^{\text{post}} | \boldsymbol{\tau}_t = \bar{\boldsymbol{\tau}})$. This difference is a treatment

Figure VII: ***Lagrange multiplier vs. effective tax rate on capital gains.***



The blue line plots simulations of $\{\mu_t\}$ *from the model. The red line is the time series of* τ*, as defined in equation (43).*

effect that measures what would have happened after 1982 if taxes would have been kept constant at $\boldsymbol{\tau}$ level; As different statistics are in different units, re-scale that difference by the observed increase in the US data, $\hat{S}_j^{\text{post}} - \hat{S}_j^{\text{pre}}$. Thus, I obtain a Relative Average Treatment Effect (RATE) -that is, the percent change in the statistic due to tax cuts relative to the observed increase:

$$RATE_j = \frac{\widetilde{S}_j(\hat{\theta})^{\text{post}} - (\widetilde{S}_j(\hat{\theta})^{\text{post}} | \boldsymbol{\tau}_t = \bar{\boldsymbol{\tau}})}{\hat{S}_j^{\text{post}} - \hat{S}_j^{\text{pre}}} \times 100 \tag{59}$$

This statistic can be decomposed into two ratios: (i) the effect of tax cuts in the model (called "RATE-m" for "model" in table IV); ii) the model performance:

$$RATE_j = \underbrace{\frac{\widetilde{S}_j(\hat{\theta})^{\text{post}} - (\widetilde{S}_j(\hat{\theta})^{\text{post}} | \boldsymbol{\tau}_t = \bar{\boldsymbol{\tau}})}{\widetilde{S}_j(\hat{\theta})^{\text{post}} - \widetilde{S}_j(\hat{\theta})^{\text{pre}}}}_{\text{RATE-m}} \times \underbrace{\frac{\widetilde{S}_j(\hat{\theta})^{\text{post}} - \widetilde{S}_j(\hat{\theta})^{\text{pre}}}{\hat{S}_j^{\text{post}} - \hat{S}_j^{\text{pre}}}}_{\text{Model fit}} \times 100 \tag{60}$$

This decomposition helps to identify situations where tax cuts are crucial in the model but the model falls short of matching the data or, contrarily, the model overstates the increase in the statistics and tax cuts are not that important in the model.

To run the counterfactual, I selected $\boldsymbol{\tau} = [0.15, 0.24]$. For CGT, 15% represents the mean level from 1946 to 1982 and the highest level recorded between 1982 and 2018. Thus, it represents a reasonable historical benchmark for a scenario without tax cuts. For DT, I choose 24%, which is the value at the end of 1981, before the 1982 cuts. I assume that the constant effective rate is announced and implemented by the government at the same time and investors immediately incorporate that information to their expectations. If, instead, investors do not believe the policy right away and learn about the effective rates, learning happens very fast and results are virtually unchanged.

Table IV shows the RATE of tax cuts in key statistics. According to the model, tax cuts accounted for 40% of increase in volatility, 30% of the higher mean valuations, 25% of the higher pass-through from beliefs to prices, and more than 50% of the increase in the PD ratio response to price shocks. Thus, tax cuts appear as a first-order driver of stock market valuations, including its volatility. In most cases, RATE and RATE-m deliver very similar numbers, reflecting the good performance of the model, but there are two exceptions. First, the effect of tax cuts on $\mathbb{E}(PD_t)$ in the model is much more significant, explaining up to 60% of the increase, but the model only generates about half of the empirical increase. Similarly, in the model, tax cuts explain almost half of the increase in $\tilde{\omega}$, but the model generates just half of the empirical increase.

Columns 4 to 9 of table IV show the effect of tax cuts only on CGT or DT. The contributions of cuts to each tax are fairly similar but there are exceptions. DT cuts seem to be more influential for the increase in mean valuations. However, the RATE measure does not consider that the magnitude of tax cuts on capital gains and dividends was different: While CGT went down from around 15% to 8%, the decline in the DT rate was steeper, starting from a rate around 25% and going down to around 7%. Thus, despite the RATE of both tax types on the variance of the PD ratio is fairly similar, it must be that the RATE of a tax cut of 1pp is notably larger for CGT. To quantify this idea, I divide the RATE by the average tax cut in tax τ over the post-1982 period, what might be called a Unitary Relative Average Treatment Effect (URATE):

$$URATE_j = \frac{RATE_j}{\frac{1}{T}\sum_{t=1982.III}^{T}(\tau - \tau_t)} \tag{61}$$

As expected, the effect of CGT on the variance of the PD ratio is much larger than that of dividends tax cuts, 6.33 vs. 2.75, that is, a 1pp cut in CGT increase volatility by around 6% while a 1pp cut to DT would increase volatility by less than 3%. CGT cuts have also a stronger influence on the IRF coefficient. Moreover, the impact of CGT cuts on the variance of the PD ratio is twice as large as its effect on the mean PD ratio, 6.33 vs 3.19. Contrarily, DT cuts had a similar effect on both the variance and mean of the PD ratio, 2.75 and 2.40 respectively. Considered together, these figures illustrate the particular relevance of CGT for asset price volatility.

5.2. Sensitivity of the Relative Average Treatment Effect

In this section, we explore how different assumptions of the model and the uncertainty about the model parameters affect the Relative Average Treatment Effect.

I start by exploring the effects of variation of the two less conventional parameters, the tax elasticity of realization ξ and the Kalman gain g, on the estimated effects of tax cuts. Based on the sensitivity analysis of Section 3.3. and other estimations in the literature, a grid of values is constructed for both parameters reflecting uncertainty bands. For the tax elasticity of realization, I consider $\xi \in [-0.7, -0.3]$. The largest deviation of $\hat{\xi}$ when the statistics were perturbed was -0.64 to -0.37. In addition, Agersnap and Zidar (2021) found $\xi \in [-0.5, -0.3]$, and the US Congress uses $\xi = -0.70$. The interval I consider allows for all these values. In turn, I use $g \in [0.0236, 0.0336]$, which is a wide range of values around the point estimate 0.0286, that includes the largest deviations from it in the perturbation exercise and includes also other estimations of the parameter such as Adam, Marcet, and Beutel (2017).

Table IV: ***Relative Average Treatment Effect of Tax Cuts.***

	RATE	RATE-m	URATE
	All taxes		
$\mathbb{V}(\ln P_t/D_t)$	39.11	40.59	
$\mathbb{E}(PD_t)$	29.73	58.18	
$\tilde{\omega}$	25.07	45.58	
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	56.72	59.46	
	CGT		
$\mathbb{V}(\ln P_t/D_t)$	25.40	26.36	6.33
$\mathbb{E}(PD_t)$	12.79	25.03	3.19
$\tilde{\omega}$	12.38	22.50	3.08
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	30.93	32.42	7.71
	DT		
$\mathbb{V}(\ln P_t/D_t)$	21.30	22.11	2.75
$\mathbb{E}(PD_t)$	18.57	36.34	2.40
$\tilde{\omega}$	14.45	26.26	1.86
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	30.41	31.89	3.93

This table reports the Relative Average Treatment Effect of capital gains and dividends tax cuts. The column "RATE" refers to the statistic in equation (59). The column "RATE-m" stands for the first component of the decomposition in equation (60). The counterfactual tax levels are set to $\tau = 0.15$ *and* $\tau^d = 0.24$.

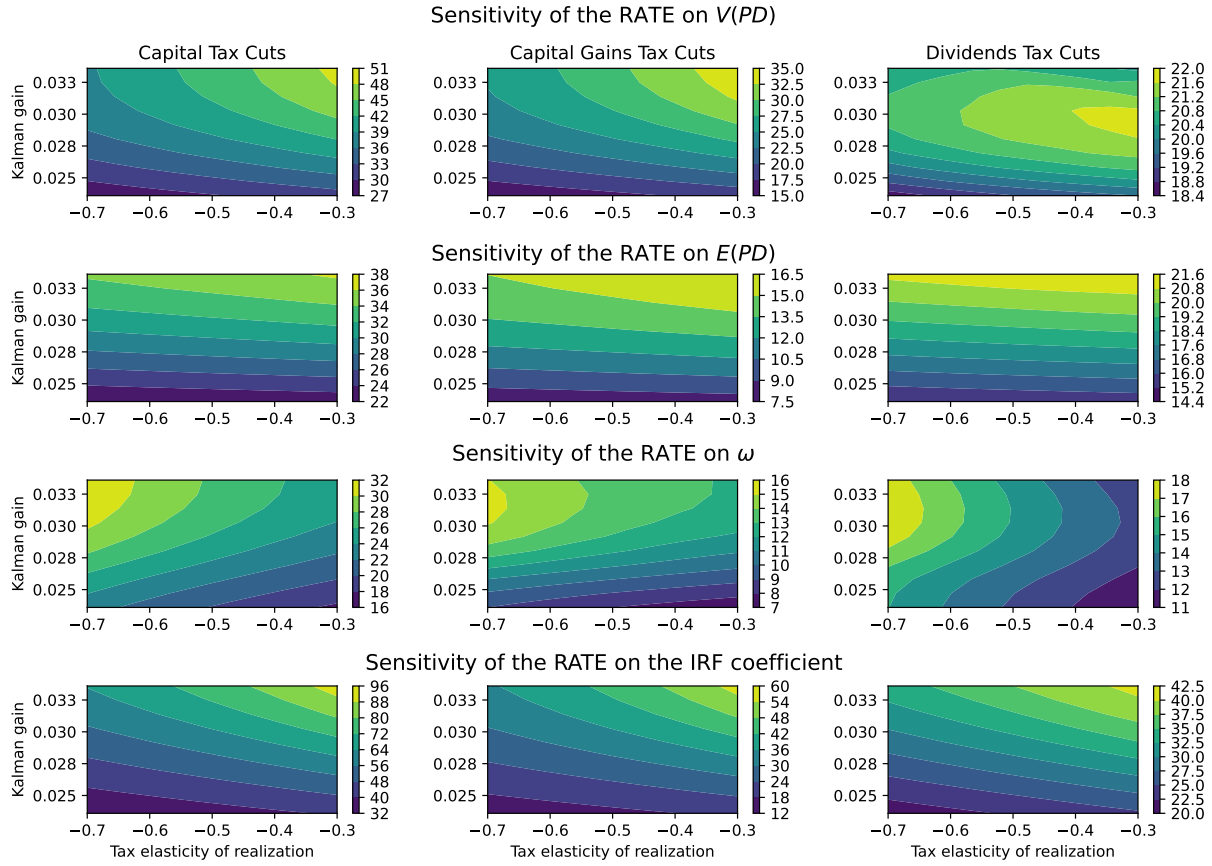
Figure VIII plots the sensitivity of RATE to variations within the uncertainty bands of the parameters. The figures represent a type of confidence bands for the RATE estimates, reflecting the uncertainty of the underlying parameters. I start by discussing the sensitivity of the effects of tax cuts on the variance of the PD ratio, which corresponds to the plots in the first row. As illustrated in the first column, the effect of tax cuts is increasing in both ξ and g. Thus, the RATE of capital tax cuts on the volatility of the PD registers a lower bound of 27.72% if ξ is very negative and g is small and an upper bound of 48.97% if the elasticity is less negative and the gain is higher. Intuitively, the more negative ξ is, the stronger the lock-in effect and the less taxes affect volatility. Similarly, the higher g is, the more agents incorporate new information, the stronger the price-expectations spiral, and the greater the destabilizing effect of lower taxes. Most of this variation in the RATE for $\mathbb{V}(P_t/D_t)$ is due to the uncertainty around the effects of CGT cuts, which oscillate between 15.40% and 34.53% while the bands for the effects of dividends tax cuts are tighter, between 18.58% and 21.77% . The reason is that the effect of CGT is highly sensitive to ξ due to the lock-in effect and g due to the degree of price extrapolation, both factors being second-order for DT.

As illustrated in the second row of figure VIII, the confidence bands for $\mathbb{E}(P_t/D_t)$ are tighter, going from 22.02% to 36.28%. In this case, the confidence intervals for both CGT and DT are more similar than for the variance, illustrating the lock-in effect is particularly relevant for volatility. The effect of tax cuts on $\tilde{\omega}$ increases with g and decreases with ξ, oscillating between 17.66% and 31.62%, and is quite symmetric with the CGT and DT cuts. As for the IRF coefficient, the RATE is increasing in both g and ξ, showing ample bands from 33.21% to 91.74%. These large bands are mostly related to CGT cuts due to the role of the lock-in effect and price extrapolation.

Besides parameter uncertainty, I experiment with alternative assumptions about key elements of the model. In particular, I focus on alternative assumptions about taxes and expectations. To run this exercise, I use the parameterization outlined in table II, such that the marginal effect of the particular assumption is measured. Table V reports the results for the RATE of tax cuts, focusing on the four most important statistics.

I start by exploring the possibility of tax foresight, as there exists evidence that markets have substantial ability to forecast taxes (see, for instance, Kueng, 2014). To address this possibility, I endow investors with perfect foresight of near-term taxes (which are indeed the only relevant ones under Learning), such that $\mathbb{E}_t^{\mathcal{P}}(\tau_{t+1}) = \tau_{t+1}$ and $\mathbb{E}_t^{\mathcal{P}}(\tau_{t+1}^d) = \tau_{t+1}^d$. It turns out the effect of near-term foresight is very modest, delivering RATE very similar to the baseline ones; the RATE on volatility is 36% instead of 39%.

Figure VIII: ***Sensitivity of the Relative Average Treatment Effect of tax cuts.***



Each row shows the RATE on one variable: the PD ratio variance at the top, the mean PD ratio at the top middle, the pass-through $\tilde{\omega}$ at the bottom middle and the IRF coefficient at the bottom. The first column shows the effect of no tax cuts; the second illustrates no CGT cuts; the third, no DT cuts. In the x-axis of each subplot, there is the considered interval of variation for the tax elasticity of realization ξ. In the y-axis of each subplot, there is the interval considered for the Kalman gain g. The colorbars in each subplot show the interval of RATE when varying ξ and g.

Another option is to use statutory tax rates. In general, this alternative implies higher and more stable tax rates, as they are not affected by regulatory changes or by the decline in the taxable share of equity income. Since the model is estimated for lower taxes, the higher and more stable statutory rates reduce the overall volatility and mean valuations generated by the model. As a result, RATEs are smaller; for instance, the RATE on volatility is 33% rather than the baseline of 40%. Yet, internally to the model, RATEs are very similar to the baseline case—as captured by the RATE-m column—illustrating that the endogenous dynamics are robust to different tax measurements.

A third option regarding taxes is the case without rebates. Thus far, I have assumed taxes were lump-sum rebated to investors, ruling out income effects. Without rebates, agents' budgets have richer dynamics in equilibrium, as they include realized capital gains. Despite this, results are virtually identical, perhaps unsurprisingly given the assumption of logarithmic utility.

In terms of expectations, an important element of the quantitative model is that agents learn about prices, dividends and taxes. A more standard assumption is RE, whereby expectations fluctuate given the AR(1) nature of dividend growth and taxes but there is no extrapolation from prices. In this case, tax cuts have a significantly lower effect on volatility, about 25%. Thus, the presence of information frictions and learning accounts for about 3/8 of the effects of tax cuts on volatility. Interestingly, though, RATE-m for volatility is much higher than for Learning, 62.15% vs 40.59%, which highlights that despite the RE model is incapable of generating a sufficient increase in volatility, within the model tax cuts are the main reason of the increase in fluctuations. A related observation is the inability to replicate the increase in the IRF coefficient, which again suggests the interaction of tax cuts with price extrapolation is a quantitatively crucial to explain the data.

On the contrary, the effect of tax cuts on mean valuations is higher, 40% instead of 30%. This result illustrates that replicating better the run-up in valuations -as RE does- is not sufficient to explain the increase in volatility better -in fact, Learning is much better at this. Although not reported in the table, besides volatility, the model with RE has serious problems in terms of too low serial correlation of the PD ratio, too low responses of the PD ratio to shocks or a too low correlation between prices and expectations. Thus, the volatility in beliefs introduced by Learning appears as key to replicate salient features of the data.

A midway is agents learning about prices, but hold Rational Expectations about wages, dividends and taxes. In this case, tax cuts explain a bit less of the increase in volatility, 30.24% instead of 39.11%, and a bit more of the mean valuations run-up, 43.05% instead of 29.73%, the higher pass-through, 32.76% instead of 25.07%, and the IRF coefficient. Thus, learning about div-

idends and taxes seems to dampen the increase in valuations and pass-through a bit, but helps in replicating better the increase in volatility without exaggerating the change in the IRF coefficient.

Finally, I report RATEs for three alternative parameterizations. In these three cases, the model has been reestimated; the estimated parameters values can be found in table VII of Appendix F. First, I impose σ_W, the standard deviation of wage shocks, to be constant in the two subperiods rather than allow it to change. In terms of RATE, results are remarkably similar to the baseline; for instance, the RATE on volatility is 34.35% instead of the baseline of 39.11%. However, the model fit regarding volatility is much worse, explaining only about 1/3 of the rise on it −delivering a RATE-m of 90%. In other words, without the increase in σ_W the model falls short of volatility and virtually all the change in volatility comes from tax cuts. Thus, allowing for a change in σ_W which reflects non-tax cuts causes of higher volatility is crucial for the good performance of the model, but it is relatively unimportant for measuring the effects of tax cuts. Second, I impose $\xi = -0.7$, which is the value used by the US Congress, and reestimate the other parameters. The RATE for volatility is 28%, which is very close to the value obtained in the lower bound of the sensitivity analysis in figure VIII.

Asset pricing models often face criticism for their poor out-of-sample performance. To address this issue, I conduct a highly restricted estimation using only data from the first subsample (that is, $t < 1982 : Q3$), imposing a constant σ_W across periods (specifically, the one estimated for the first subsample) and fixing $\xi = -0.7$. Even under these constraints, the RATE on volatility remains close to 30%. However, the model's fit regarding volatility is significantly worse, explaining only about 1/3 of the observed increase, which is primarily attributable to tax cuts—the RATE-m is 91.38%. This scenario closely resembles the case of a homogeneous σ_W when using the entire sample, suggesting that estimating the model using only one subsample does not substantially alter the results.

5.3. Historical calibration

In this section, I illustrate the effect of tax cuts through a calibration exercise. Instead of using supply shocks to replicate the empirical series of stock repurchases, I now calibrate a series of supply shocks to perfectly match the empirical dynamics of the PD ratio. This approach ensures that the model matches the data precisely using the empirical time series of taxes. Then I simulate the same "no tax cuts" scenario as before and compare the two scenarios.

Figure IX illustrates the results of our calibration exercise. The red line represents the estimated model with calibrated supply shocks, which precisely matches the empirical data. In contrast, the

Table V: ***Relative Average Treatment Effect under alternative assumptions.***

	RATE	RATE-m	RATE	RATE-m
	Tax Foresight		Statutory τ	
$\mathbb{V}(\ln P_t/D_t)$	36.23	37.81	33.15	37.57
$\mathbb{E}(PD_t)$	29.65	57.83	21.39	56.51
$\tilde{\omega}$	23.98	43.53	19.37	45.48
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	45.27	47.43	35.87	55.69
	No tax rebates		Rational Expectations	
$\mathbb{V}(\ln P_t/D_t)$	38.64	40.29	25.95	62.15
$\mathbb{E}(PD_t)$	30.30	58.24	40.01	70.63
$\tilde{\omega}$	25.18	45.89	20.21	97.56
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	57.32	59.81	28.30	43.46
	Only price learning		No increase in σ_W	
$\mathbb{V}(\ln P_t/D_t)$	30.24	38.06	34.35	89.81
$\mathbb{E}(PD_t)$	43.05	72.11	27.25	61.75
$\tilde{\omega}$	32.76	57.25	25.34	53.14
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	66.67	57.98	48.91	63.29
	$\xi = -0.7$		$\xi = -0.7$, $t < 1982$	
$\mathbb{V}(\ln P_t/D_t)$	28.36	29.62	29.43	91.38
$\mathbb{E}(PD_t)$	26.55	54.61	24.01	60.48
$\tilde{\omega}$	20.35	40.53	23.05	2.39
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P = 0, I_t)$	42.77	53.47	40.16	60.87

This table reports the Relative Average Treatment Effect of capital gains and dividends tax cuts as defined in expression (59). The column "RATE-m" stands for the first component of the decomposition in equation (60). The labels signal departures from different assumptions of the baseline model.

blue line depicts the mean of simulations in a scenario without tax cuts. This approach offers a distinct advantage by allowing us to explore the timing of the effects of tax cuts on market valuations.

Interestingly, until the second half of the 1990s, the series with and without tax cuts appear remarkably similar, despite some tax cuts being implemented since the 1980s. This observation can be attributed to non-linearities in the model. The relationship between taxes and volatility is highly convex, such that tax cuts have particularly large effects when tax rates are already low.

The model without tax cuts still produces a boom and bust around the 2000s, suggesting that other forces are necessary to fully explain that bubble. For example, Bordalo et al. (2024) proposed that the overreaction of long-term expectations was a key element in explaining the bubble. However, the magnitude of the boom is notably smaller without tax cuts. Specifically, the PD ratio peaks around 70 instead of 90, representing a reduction of approximately 22%. Similarly, during the Great Recession, the model predicts a sharp drop from 45 to 22, which is significantly less severe than the observed decline from 57 to 27 in reality – a difference of about 23%. In terms of levels, the model without tax cuts still delivers a sustained run-up in valuations from 30 to 40, albeit not reaching the observed level of 50. This suggests that other explanations, such as the reduction in macroeconomic risk boosting risky investments proposed by Lettau, Ludvigson, and Wachter (2008), may also play a role in explaining the full extent of the sustained increase in market valuations.

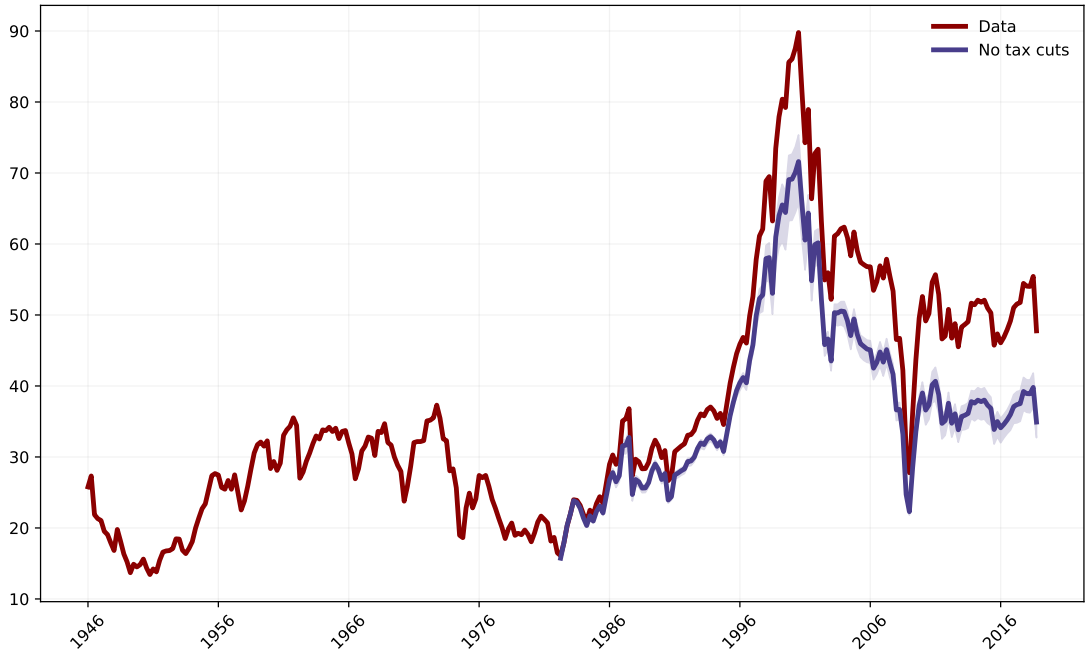
5.4. Informational Efficiency

Although the data and the model exhibit a decline in return predictability, the interpretation of this decline as an increase in efficiency is generally problematic by the Joint Hypothesis Problem (Fama, 1970). Instead, this section studies informational efficiency through the lens of the model.

In the model with Learning, prices are not informationally efficient. As pointed out by Adam, Marcet, and Beutel (2017), payoffs are discounted using a subjective factor that differs from the equilibrium discount factor —that is, the marginal rate of substitution in the real economy. Intuitively, prices reflect subjective beliefs, which are characterized by extrapolation and predictable forecast errors. Thus, price dynamics are not fully justified by fundamentals. This approach contrasts with RE, whereby price fluctuations are fully efficient but miss salient features of the data.

In this context, efficiency can be tested by tracking the deviations in valuations between a version of the model with Learning and a version with RE. Put differently, the exercise consists of a comparison between a model with good quantitative properties but inefficient prices and the

Figure IX: ***The Price-Dividend ratio with and without tax cuts.***



The red line is the data for the S&P 500 Price-Dividend ratio. The blue line plots a counterfactual with $\tau_t = 0.15$ *and* $\tau_t^d = 0.24$ *for all* $t > 1982.III$*. The confidence bands at 95% level are computed from 1000 simulations of the model. Simulations use a calibrated series of stock supply shocks.*

efficient counterpart of the same model. Mimicking a statistic used in portfolio analysis, let the tracking error be

$$\sigma_E = \sqrt{\mathbb{V}\Big(\frac{P_t^L}{D_t} - \frac{P_t^{RE}}{D_t}\Big)} \tag{62}$$

A higher σ_E signals higher deviations of Learning from the efficient benchmark and then, a lower informational efficiency of valuations. For the baseline parameterization, I obtain $\sigma_E = 5.02$ before the 1980s and $\sigma_E = 6.88$ afterward, which represent a decline in the informational efficiency of valuations of 37%. This reduction in efficiency can be formally tested using the following one-sided t-test:

$$\begin{aligned} H_0 &: \phi_1 = 0 \\ H_1 &: \phi_1 > 0 \end{aligned} \tag{63}$$

where ϕ_1 comes from the following models

$$\frac{P_t^L}{D_t} - \frac{P_t^{RE}}{D_t} = \pi_0 + \pi_1 \mathcal{D}_t + \epsilon_t \tag{64}$$

$$\epsilon_t^2 = \phi_0 + \phi_1 \mathcal{D}_t + \tilde{\epsilon}_t \tag{65}$$

with $\mathcal{D}_t$ being a dummy that takes 1 if $t > 1982.II$ and 0 otherwise. Under the null hypothesis, the variance of the Learning valuations deviations from RE valuations is the same in the two subsamples. It turns out $\hat{\phi}_1 = 22.02$ with a p-val $= 0.0001$. Thus, the null hypothesis is formally rejected, flagging a significant increase in inefficiency after the 1980s.

What was the role of tax cuts in the decline in efficiency? Without tax cuts, $\hat{\phi}_1^{\text{no cuts}} = 12.83$, which signals that tax cuts were responsible for $(\hat{\phi}_1 - \hat{\phi}_1^{\text{no cuts}})/\hat{\phi}_1 = 42\%$ of the fall in efficiency. Again, CGT plays a special role, accounting for 32% of the fall, while DT cuts alone account for 22%.

6. Conclusion

This paper argues that lower taxes on personal capital income contribute to increased asset price fluctuations. In a consumption-based asset pricing framework, an inverse relationship is established between capital tax levels and asset price volatility. The key mechanism driving this relationship is the impact of taxes on the pass-through from belief fluctuations to asset prices. While this finding proves robust across various models of expectation formation, it has dynamic implications when

investors extrapolate from prices. In such cases, taxes mitigate the price-expectations spiral that emerges.

The theory has been applied to the United States to link the resurgence of asset price fluctuations to the sharp decline in capital taxes since the 1980s. Using a structurally estimated model that is quantitatively consistent with a comprehensive set of statistics, I have found that tax cuts account for approximately 30% of the higher valuations and 40% of their increased fluctuations. Thus, tax cuts posed a mean-variance trade-off. Moreover, the model implies a significant reduction in stock market informational efficiency, attributed to a higher pass-through from beliefs to prices, partly induced by lower taxes.

Although the research has focused on the U.S. in the postwar period, capital income taxes have exhibited notable variation across countries and time. Anecdotal evidence suggests various tax cuts that preceded bubbles. In the U.S., the rise and fall of Wall Street in the Roaring Twenties was preceded by sharp cuts at the beginning of the 1920s; tax cuts also preceded episodes such as the Scandinavian Crisis or the Housing Bubble in countries like Spain. In the opposite direction, Japan introduced a tax on stock capital gains in 1989 to curb its bubble. Exploiting all this historical variation appears to be an interesting research avenue to test the model's hypotheses and explore the effects of capital taxes on financial instability globally.

Due to the simplicity of the setup, the potential feedback from asset prices to the macroeconomy has been ignored. Embedding the asset pricing model with capital taxes in macroeconomic models can be used to revisit questions about the consequences of tax changes on revenues, investment, inequality, growth, and welfare. Following this approach, the issue of optimal capital gains taxation might be further explored. As Aguiar, Moll, and Scheuer (2024), integrating richer asset pricing frameworks into optimal policy models might offer new insights into the theory of capital taxation, particularly from a macroprudential viewpoint.

Economists have long been in search of a tax instrument to mitigate financial instability. Although a Financial Transactions Tax has been a prominent proposal, its impact on price volatility remains unclear. Against this backdrop, this article shows that capital taxes, especially on capital gains, have first-order consequences for price dynamics, suggesting that they can be effective tools in preventing asset price bubbles.

References

Abel, Andrew B (1990). "Asset prices under habit formation and catching up with the Joneses". In: *The American Economic Review, Papers and Proceedings* 80.2, pp. 38–42.

Adam, Klaus (2020). *Monetary policy challenges from falling natural interest rates*. Tech. rep. ECB conference.

Adam, Klaus and Albert Marcet (2011). "Internal rationality, imperfect market knowledge and asset prices". In: *Journal of Economic Theory* 146.3, pp. 1224–1252.

Adam, Klaus, Albert Marcet, and Johannes Beutel (2017). "Stock price booms and expected capital gains". In: *American Economic Review* 107.8, pp. 2352–2408.

Adam, Klaus, Albert Marcet, and Juan Pablo Nicolini (2016). "Stock market volatility and learning". In: *The Journal of Finance* 71.1, pp. 33–82.

Agersnap, Ole and Owen Zidar (2021). "The tax elasticity of capital gains and revenue-maximizing rates". In: *American Economic Review: Insights* 3.4, pp. 399–416.

Aguiar, Mark, Benjamin Moll, and Florian Scheuer (2024). "Putting the 'Finance'into 'Public Finance': A Theory of Capital Gains Taxation". In.

Anagnostopoulos, Alexis, Eva Cárceles-Poveda, and Danmo Lin (2012). "Dividend and capital gains taxation under incomplete markets". In: *Journal of Monetary Economics* 59.7, pp. 599–611.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro (2017). "Measuring the sensitivity of parameter estimates to estimation moments". In: *The Quarterly Journal of Economics* 132.4, pp. 1553–1592.

Bansal, Ravi and Amir Yaron (2004). "Risks for the long run: A potential resolution of asset pricing puzzles". In: *The journal of Finance* 59.4, pp. 1481–1509.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer (2018). "Extrapolation and bubbles". In: *Journal of Financial Economics* 129.2, pp. 203–227.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, Matthew O'Brien, and Andrei Shleifer (2023). "Long Term Expectations and Aggregate Fluctuations". In: *NBER Chapters*.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2024). "Belief overreaction and stock market puzzles". In: *Journal of Political Economy* 132.5, pp. 1450–1484.

Brun, Lıdia and Ignacio González (2017). "Tobin's Q and Inequality". In: *Available at SSRN 3069980*.

Burnside, Craig (1998). "Solving asset pricing models with Gaussian shocks". In: *Journal of Economic Dynamics and control* 22.3, pp. 329–340.

Buss, Adrian and Bernard Dumas (2019). "The dynamic properties of financial-market equilibrium with trading fees". In: *The Journal of Finance* 74.2, pp. 795–844.

Bybee, J Leland (2023). "The ghost in the machine: Generating beliefs with large language models". In: *arXiv preprint arXiv:2305.02823*.

Campbell, John Y and John H Cochrane (1999). "By force of habit: A consumption-based explanation of aggregate stock market behavior". In: *Journal of political Economy* 107.2, pp. 205–251.

Campbell, John Y and Robert J Shiller (1988). "The dividend-price ratio and expectations of future dividends and discount factors". In: *The Review of Financial Studies* 1.3, pp. 195–228.

Cappelletti, Giuseppe, Giovanni Guazzarotti, and Pietro Tommasino (2017). "The stock market effects of a securities transaction tax: quasi-experimental evidence from Italy". In: *Journal of Financial Stability* 31, pp. 81–92.

Cassella, Stefano and Huseyin Gulen (2018). "Extrapolation bias and the predictability of stock returns by price-scaled variables". In: *The Review of Financial Studies* 31.11, pp. 4345–4397.

Chen, Andrew Y., Fabian Winkler, and Rebecca Wasyk (2019). "In full-information estimates, long-run risks explain at most a quarter of p/d variance, and habit explains even less". In: *Critical Finance Review* 10, No. 3, pp. 329–381.

Cochrane, John H (2009). *Asset pricing: Revised edition*. Princeton university press.

— (2017). "Macro-finance". In: *Review of Finance* 21.3, pp. 945–985.

Constantinides, George M (1990). "Habit formation: A resolution of the equity premium puzzle". In: *Journal of political Economy* 98.3, pp. 519–543.

Dai, Zhonglan, Edward Maydew, Douglas A Shackelford, and Harold H Zhang (2008). "Capital gains taxes and asset prices: capitalization or lock-in?" In: *The Journal of Finance* 63.2, pp. 709–742.

Dai, Zhonglan, Douglas A Shackelford, and Harold H Zhang (2013). "Capital gains taxes and stock return volatility". In: *The Journal of the American Taxation Association* 35.2, pp. 1–31.

Dammon, Robert M, Chester S Spatt, and Harold H Zhang (2001). "Optimal consumption and investment with capital gains taxes". In: *The Review of Financial Studies* 14.3, pp. 583–616.

Dávila, Eduardo (2023). "Optimal financial transaction taxes". In: *The Journal of Finance* 78.1, pp. 5–61.

Dowd, Tim, Robert McClelland, and Athiphat Muthitacharoen (2015). "New evidence on the tax elasticity of capital gains". In: *National Tax Journal* 68.3, pp. 511–544.

Ehling, Paul, Stathis Tompaidis, and Chunyu Yang (2019). "Tax Collection from Realized Capital Gains on Equity". In: *Available at SSRN 3349058*.

Evans, George W, Seppo Honkapohja, and Noah Williams (2010). "Generalized stochastic gradient learning". In: *International Economic Review* 51.1, pp. 237–262.

Fama, Eugene F (1970). "Efficient capital markets: A review of theory and empirical work". In: *The journal of Finance* 25.2, pp. 383–417.

Feenberg, Daniel and Elisabeth Coutts (1993). "An introduction to the TAXSIM model". In: *Journal of Policy Analysis and management* 12.1, pp. 189–194.

Gavin, William T, Benjamin D Keen, and Finn E Kydland (2015). "Monetary policy, the tax code, and the real effects of energy shocks". In: *Review of Economic Dynamics* 18.3, pp. 694–707.

Gavin, William T, Finn E Kydland, and Michael R Pakko (2007). "Monetary policy, taxes, and the business cycle". In: *Journal of Monetary Economics* 54.6, pp. 1587–1611.

Gemmill, Robert F (1956). "The Effect of the Capital Gains Tax on Asset Prices". In: *National Tax Journal* 9.4, pp. 289–301.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus (2021). "Five facts about beliefs and portfolios". In: *American Economic Review* 111.5, pp. 1481–1522.

Gourio, Francois and Jianjun Miao (2010). "Firm heterogeneity and the long-run effects of dividend tax reform". In: *American Economic Journal: Macroeconomics* 2.1, pp. 131–168.

Gravelle, Jane G (2020). "Capital gains tax options: Behavioral responses and revenues". In: *CRS Report for Congress*. Vol. 41364.

Greenwood, Robin and Andrei Shleifer (2014). "Expectations of returns and expected returns". In: *The Review of Financial Studies* 27.3, pp. 714–746.

Hakansson, Nils H. (1970). "Optimal Investment and Consumption Strategies Under Risk for a Class of Utility Functions". In: *Econometrica* 38.5, pp. 587–607.

Haugen, Robert A and A James Heins (1969). "The effects of the personal income tax on the stability of equity value". In: *National Tax Journal* 22.4, pp. 466–471.

Haugen, Robert A and Dean W Wichern (1973). "The diametric effects of the capital gains tax on the stability of stock prices". In: *The Journal of Finance* 28.4, pp. 987–996.

Jin, Lawrence J and Pengfei Sui (2022). "Asset pricing with return extrapolation". In: *Journal of Financial Economics* 145.2, pp. 273–295.

Judd, Kenneth L (1992). "Projection methods for solving aggregate growth models". In: *Journal of Economic theory* 58.2, pp. 410–452.

Keynes, John Maynard (1936). *The general theory of employment, interest, and money*. London: MacMillan.

Kilian, Lutz and Robert J Vigfusson (2011). "Are the responses of the US economy asymmetric in energy price increases and decreases?" In: *Quantitative Economics* 2.3, pp. 419–453.

Kueng, Lorenz (2014). *Tax news: The response of household spending to changes in expected taxes*. Tech. rep. National Bureau of Economic Research.

Lettau, Martin, Sydney C Ludvigson, and Jessica A Wachter (2008). "The declining equity premium: What role does macroeconomic risk play?" In: *The Review of Financial Studies* 21.4, pp. 1653–1687.

Ljung, Lennart (1977). "Analysis of recursive stochastic algorithms". In: *IEEE transactions on automatic control* 22.4, pp. 551–575.

Lucas, Robert E. (1978). "Asset prices in an exchange economy". In: *Econometrica: Journal of the Econometric Society*, pp. 1429–1445.

Malmendier, Ulrike and Stefan Nagel (2016). "Learning from inflation experiences". In: *The Quarterly Journal of Economics* 131.1, pp. 53–87.

Marcet, Albert (1988). "Solving nonlinear stochastic growth models by parametrizing expectations." In.

Martin, Alberto and Jaume Ventura (2018). "The macroeconomics of rational bubbles: a user's guide". In: *Annual Review of Economics* 10, pp. 505–539.

McGrattan, Ellen R (2023). "Taxes, regulations, and the value of US corporations: A reassessment". In: *Review of Economic Dynamics* 50, pp. 131–145.

McGrattan, Ellen R and Edward C Prescott (2003). "Average debt and equity returns: Puzzling?" In: *American Economic Review* 93.2, pp. 392–397.

— (2005). "Taxes, Regulations, and the Value of US and UK Corporations". In: *The Review of Economic Studies* 72.3, pp. 767–796.

Nagel, Stefan and Zhengyang Xu (2022). "Asset pricing with fading memory". In: *The Review of Financial Studies* 35.5, pp. 2190–2245.

Piketty, Thomas (2014). *Capital in the twenty-first century*. Harvard University Press.

Piketty, Thomas and Gabriel Zucman (2014). "Capital is back: Wealth-income ratios in rich countries 1700–2010". In: *The Quarterly journal of economics* 129.3, pp. 1255–1310.

Rosenthal, Steve and Lydia Austin (2016). "The dwindling taxable share of US corporate stock". In: *Tax Notes* 151.6.

Sarin, Natasha, Lawrence Summers, Owen Zidar, and Eric Zwick (2022). "Rethinking how we score capital gains tax reform". In: *Tax Policy and the Economy* 36.1, pp. 1–33.

Shiller, Robert J (2000). *Irrational exuberance.* Princeton university press.

Sialm, Clemens (2006). "Stochastic taxation and asset pricing in dynamic general equilibrium". In: *Journal of Economic Dynamics and Control* 30.3, pp. 511–540.

— (2009). "Tax changes and asset pricing". In: *American Economic Review* 99.4, pp. 1356–83.

Somers, Harold M (1948). "An economic analysis of the capital gains tax". In: *National Tax Journal* 1.3, pp. 226–232.

— (1960). "Reconsideration of the capital gains tax". In: *National Tax Journal* 13.4, pp. 289–309.

Stantcheva, Stefanie (2020). "Dynamic taxation". In: *Annual Review of Economics* 12.1, pp. 801–831.

Stiglitz, Joseph E (1983). "Some aspects of the taxation of capital gains". In: *Journal of Public Economics* 21.2, pp. 257–294.

— (1989). "Using tax policy to curb speculative short-term trading". In: *Regulatory Reform of Stock and Futures Markets: A Special Issue of the Journal of Financial Services Research*, pp. 3–17.

Tobin, James (1978). "A proposal for international monetary reform". In: *Eastern economic journal* 4.3/4, pp. 153–159.

Umlauf, Steven R (1993). "Transaction taxes and the behavior of the Swedish stock market". In: *Journal of Financial Economics* 33.2, pp. 227–240.

Williams, Noah (2024). "A Learning Model of Financial Instability". In: *Online manuscript.*

Appendix A: Data Sources

Stock market data. Stock prices, dividends and CPI inflation come from Robert Shiller's website. The nominal risk-free rate is the 90-day T-Bill, from the FRED database. The monthly data has been transformed into quarterly frequency by taking the last month of the quarter. In addition, nominal variables have been transformed to real terms using Shiller's CPI index. Annual dividends have been converted to quarterly figures by dividing them by 4 to compute quarterly returns. Data on stock repurchases is from Refinitiv.

Macroeconomic data. Consumption data is the BEA real quarterly personal consumption expenditures series. Wages are the BEA compensation of employees. When computing the Wage-Dividends ratio, I use the Net Dividends from the BEA (Corporate Profits after tax with IVA and CCAdj: Net Dividends).

Capital tax rates. The base effective average marginal rates on dividends, short and long capital gains and interests are supplied by the TAXSIM program of the National Bureau of Economic Research (NBER). See Feenberg and Coutts (1993) for a description of the program. The data can be found at here. These rates are offered on an annual basis from 1960 to 2018 at federal level and from 1979 to 2008 at state level. I took the rates computed using 1984 data in/deflated to law year, which reflects changes in tax rates without conflating changes in the income distribution. The tax data was downloaded in February 2022.

Following Sialm (2009), I adjusted for state and local taxes before 1979 and after 2008 as well as for the distinction between qualified and nonqualified dividends from 2003 onward to get a complete series for the 1960-2018 period. Before 1960, the tax rates are interpolated using data from Sialm (2009). Details on the computation of the taxable share are available on Appendix C.

Survey expectations. The main survey is the UBS Index of Investor Optimism. The quantitative question on stock market expectations has been surveyed over the period 1998:Q2-2007:Q4 with 702 responses per month on average. To make the data consistent with the model, I have run some adjustments. First, the series have been deflated by using inflation expectations from the Michigan Surveys of Consumers, available at `https://data.sca.isr.umich.edu/data-archive/mine.php`. Second, I transformed real returns expectations into capital gains expectations by subtracting the mean dividend growth along the period over each period of the price-dividend ratio. In addition, I use historical data from the Philadelphia Fed's Livingston Survey on the Stock Price Index since 1952; The IBES survey, detailing analysts' forecasts on long-term real earnings growth, which starts in 1984 and was obtained from Stephan Nagel's website; the American Association of Individual

Investors' Sentiment Survey, offering monthly data since 1987; the LLM-generated sentiment indicator of Bybee (2023), to who I thank for sharing. See Appendix C for details on the treatment of this data.

Appendix B: Proofs.

B1.- Proof of Lemma 1.

Consider the pricing formula (6) in the main text. If follows $P_t/D_t = p(\boldsymbol{\beta}_t^g)$ with $\boldsymbol{\beta}_t^g = (\beta_t^D, \beta_t^P, \beta_t^M)$. Take a first-order Taylor approximation around $\boldsymbol{b} \in \mathbb{R}^3$. Then,

$$\frac{P_t}{D_t} \approx p(\boldsymbol{b}) + \boldsymbol{\omega}(\boldsymbol{\beta}_t^g - \boldsymbol{b}) \tag{66}$$

with $\boldsymbol{\omega} = \nabla p(\boldsymbol{b})$. Take the variance of both sides such that

$$\mathbb{V}\Big[\frac{P_t}{D_t}\Big] \approx \boldsymbol{\omega}(\tau)\boldsymbol{\Sigma}\boldsymbol{\omega}(\tau)^T \tag{67}$$

where $\boldsymbol{\Sigma}$ is the variance-covariance matrix of beliefs.

B2.- Proof of Lemma 2.

The lemma holds that

$$\frac{d\omega}{d\tau}\Sigma\omega^T < 0 \tag{68}$$

I start by characeterizing $\boldsymbol{\omega}$, the partial derivatives of the PD ratio with respect to beliefs (keeping them as a function of $\boldsymbol{\beta}_t^g$):

$$\omega_1 = \frac{1}{1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M} \tag{69}$$

$$\omega_2 = \frac{\beta_t^D(1-\pi\tau)}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2} \tag{70}$$

$$\omega_3 = \frac{\beta_t^D\pi\tau}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2} \tag{71}$$

Since $\pi\tau \in (0,1)$, $\boldsymbol{\beta}_t^g > 0$, ω_2 and ω_3 are positive. In turn, ω_1 is also positive since the denominator was assume positive to guarantee positive prices.

Let us look now at $\frac{\partial}{\partial\tau}\boldsymbol{\omega}$:

$$\omega_1' = -\frac{\pi(\beta_t^P - \beta_t^M)}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2} \tag{72}$$

$$\omega_2' = -\frac{\pi\beta_t^D}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2}\Big(1 + \frac{2(1-\pi\tau)(\beta_t^P - \beta_t^M)}{1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M}\Big) \tag{73}$$

$$\omega_3' = \frac{\pi\beta_t^D}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2}\Big(1 - \frac{2\pi\tau(\beta_t^P - \beta_t^M)}{1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M}\Big) \tag{74}$$

Note that $\omega_1', \omega_2' \leq 0$ and $\omega_3' \geq 0$ under the Lemma assumptions. Hence, given the assumptions about the signs of the elements in Σ, the only elements that might have a positive sign in $\frac{d\omega}{d\tau}\Sigma\omega^T$ are $(\omega_1'\omega_3 + \omega_1\omega_3')\Sigma_{13}$, $(\omega_2'\omega_3 + \omega_2\omega_3')\Sigma_{23}$ and $\omega_3'\omega_3\Sigma_{33}$. I derive three sufficient conditions for them not to dominate. First, $\beta_t^P \leq 1$ ensures $(\omega_1'\omega_3 + \omega_1\omega_3') \geq 0$. Then, since $\Sigma_{13} \leq 0$, the contribution to the total derivative is zero or negative. Second, $\beta_t^M \leq \beta_t^P$ and $\tau \leq 1/(2\pi)$ is sufficient for $(\omega_2'\omega_3 + \omega_2\omega_3') \geq 0$. Then, since since $\Sigma_{23} \leq 0$, the contribution to the total derivative is zero or negative. Finally, condition iii) guarantees that $\omega_3\omega_3'\mathbb{V}(\beta_t^M) \leq \omega_2\omega_2'\mathbb{V}(\beta_t^P)$, such that the positive contribution in the LHS is offset by the negative contribution in the RHS.

B3.- Proof of Lemma 3 and Proposition 2.

Case 1: Rational Expectations.

Consider the RE pricing formula (13). It follows $P_t^{RE}/D_t = p_1(d_t)$. Take a first-order Taylor approximation around the constant μ. Then,

$$\frac{P_t^{RE}}{D_t} \approx p_1(\mu) + \omega^{RE}(d_t - d) \tag{75}$$

with $\omega^{RE} = \frac{\partial P_t^{RE}/D_t}{\partial d_t}$ evaluated at μ. Taking the variance of both sides

$$\mathbb{V}\Big[\frac{P_t}{D_t}\Big] \approx (\omega^{RE})^2\mathbb{V}(d_t) \tag{76}$$

From (11), $\mathbb{V}(d_t) = \frac{\sigma_d^2}{1-\rho^2}$. Then, taxes affect the variance of the PD ratio only through ω. To study this channel, consider

$$\frac{\partial(\omega^{RE})^2}{\partial\tau} = 2\omega^{RE}\frac{\partial\omega^{RE}}{\partial\tau} \tag{77}$$

ω^{RE} reads as

$$\omega^{RE} = \sum_{j=1}^{\infty} f_0(\tau, j) f_1(\mu, j) \frac{\rho}{1-\rho}(1-\rho^j) \tag{78}$$

which has a positive sign since $0 < \rho < 1$. Then, the sign of the effect of taxes on the volatility of the PD ratio is determined by

$$\frac{\partial \omega^{RE}}{\partial \tau} = \sum_{j} f_1(\mu, j) \frac{\rho}{1-\rho}(1-\rho^j) \frac{\partial f_0(\tau, j)}{\partial \tau} \tag{79}$$

Given,

$$f_0(\tau, j) = \Big(\frac{\delta}{1-\delta\pi\tau}\Big)^j (1-\pi\tau)^{j-1} \tag{80}$$

$\partial f_0(\tau, 1)/\partial \tau = \delta^2 \pi/(1-\delta\pi\tau)^2 > 0$ while for any $j > 1$,

$$\frac{\partial f_0(\tau, j)}{\partial \tau} = \frac{\pi \delta^j (1-\pi\tau)^{j-1}}{(1-\delta\pi\tau)^j} \Big(\frac{j\delta(1-\pi\tau)}{1-\delta\pi\tau} - (j-1) \Big) \tag{81}$$

whose sign is negative only if

$$j > \frac{1-\delta\pi\tau}{1-\delta} \equiv \tilde{j}$$

In other words, the negative effect starts to dominate only in the distant future. Hence, for $\frac{\partial \omega^{RE}}{\partial \tau} < 0$, it must be that the negative terms outweigh the positive ones, that is,

$$\frac{\partial \omega^{RE}}{\partial \tau} = \frac{\rho}{1-\rho} \Bigg(\underbrace{\sum_{j=1}^{\tilde{j}} f_1(\mu, j)(1-\rho^j) \frac{\partial f_0(\tau, j)}{\partial \tau}}_{\text{Factor 1>0}} + \underbrace{\sum_{\tilde{j}=j+1}^{\infty} f_1(\mu, j)(1-\rho^j) \frac{\partial f_0(\tau, j)}{\partial \tau}}_{\text{Factor 2<0}} \Bigg) \tag{82}$$

I want to show |Factor 1| < |Factor 2|. For that, it is enough to show that $\frac{\partial f_0(\tau, j)}{\partial \tau}$ becomes increasingly negative with j and $f_1(d_t, j)$ increasingly positive. Towards that end, note

$$\lim_{j \to \infty} \frac{\partial f_0(\tau, j)}{\partial \tau} = -\infty \tag{83}$$

Additionally, consider $f_1(\mu, j)(1-\rho^j)$, which affects the final balance. Since $\rho \in (0, 1)$, $(1-\rho^j)$ gets larger with j. Besides, it turns out that

$$\frac{\partial f_1(\mu, j)}{\partial j} = f_1(d_t, j) \Bigg(\mu + \frac{\sigma_d^2}{2(1-\rho)^2} \Big(1 + \frac{2\rho^j j}{1-\rho} - \frac{2\rho^{2j+1} j}{1-\rho^2} \Big) \Bigg) \tag{84}$$

The sign of the derivative is determined by $1 - \rho^2 - \rho^j + \rho^{j+1}$, which converges to $1 - \rho^2 > 0$ when

j tends to ∞. Hence, as j gets larger, the negative effects of τ increase.

Case 2: Learning.

Consider the equilibrium pricing function under Learning, equation (23) in the main text. It follows $P_t^L/D_t = p_2(\boldsymbol{x}_t)$ with $\boldsymbol{x}_t = (\ln\beta_t, d_t)$. Take a first-order Taylor approximation around $\boldsymbol{x} = (\mu, \mu)$. Then,

$$\frac{P_t^L}{D_t} \approx p_2(\boldsymbol{x}) + \boldsymbol{\omega^L}(\boldsymbol{x_t} - \boldsymbol{x}) \tag{85}$$

with $\boldsymbol{\omega}^L = \nabla p_2(\boldsymbol{x})$. Take the variance of both sides such that

$$\mathbb{V}\Big[\frac{P_t}{D_t}\Big] \approx \boldsymbol{\omega}^L \boldsymbol{\Sigma}^L (\boldsymbol{\omega}^L)^T \tag{86}$$

The proposition claims $\frac{\partial \boldsymbol{\omega}^L}{\partial \tau} < 0$. Note that

$$\omega_1^L = \rho \frac{\delta exp\Big\{\mu + \sigma_d^2/2\Big\}}{1 - (1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau} \tag{87}$$

which is always positive under the assumptions made. Its derivative with respect to τ reads as

$$\frac{\partial \omega_1^L}{\partial \tau} = -\frac{\rho\delta exp\Big\{\mu + \sigma_d^2/2\Big\}\pi\delta\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)}{\Big(1 - (1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau\Big)^2} \tag{88}$$

A sufficient condition for it to be negative is $\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big) > 0$, that is, agents expected positive price growth. In addition,

$$\omega_2^L = \frac{\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}(1-\pi\tau)\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)}{\Big(1 - (1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau\Big)^2} \tag{89}$$

which is positive since $\pi\tau < 1$. Take its derivative with respect to τ

$$\begin{aligned}\frac{\partial \omega_2^L}{\partial \tau} = &-\frac{\rho\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)}{\Big(1 - (1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau\Big)^2} \\ &- \frac{2\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}(1-\pi\tau)\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)}{\Big(1 - (1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau\Big)^3}\delta\pi\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)\end{aligned} \tag{90}$$

which is positive provided agents expect positive price growth. Hence, $\frac{\partial \boldsymbol{\omega}^L}{\partial \tau} < 0$ in an expected growth environment.

Now I move to characterize Σ^L. First, consider the following 2nd order difference equation characterizing the dynamics of $\ln\beta_t$

$$\ln\beta_t = \ln\beta_{t-1} + g\left(\ln\frac{exp\Big\{(1-\rho)\mu + \rho d_{t-1} + \sigma_d^2/2\Big\}}{exp\Big\{(1-\rho)\mu + \rho d_{t-2} + \sigma_d^2/2\Big\}}\frac{1-(1-\pi\tau)\delta exp\Big\{\ln\beta_{t-2} + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau}{1-(1-\pi\tau)\delta exp\Big\{\ln\beta_{t-1} + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau} + d_{t-1} - \ln\beta_{t-1}\right) \tag{91}$$

Note $\ln\beta_t = f(\ln\beta_{t-1}, \ln\beta_{t-2}, d_{t-1}, d_{t-2})$. Use a first-order Taylor approximation around $\boldsymbol{x} = (\mu, \mu, \mu, \mu)$. Then,

$$\ln\beta_t \approx \kappa + \mathcal{A}\ln\beta_{t-1} + \mathcal{B}\ln\beta_{t-2} + \mathcal{C}d_{t-1} + \mathcal{D}d_{t-2} \tag{92}$$

with κ collecting all the constants in the approximation and caligraphic letters being the derivative of $\ln\beta_t$ with respect to every input evaluated at $\boldsymbol{x}$. Then, the dynamics of $(\ln\beta_t, d_t)$ can be represented as the following VAR(2):

$$\begin{bmatrix}\ln\beta_t \\ d_t\end{bmatrix} = \begin{bmatrix}\kappa \\ (1-\rho)\mu\end{bmatrix} + \begin{bmatrix}\mathcal{A} & \mathcal{C} \\ 0 & \rho\end{bmatrix}\begin{bmatrix}\ln\beta_{t-1} \\ d_{t-1}\end{bmatrix} + \begin{bmatrix}\mathcal{B} & \mathcal{D} \\ 0 & 0\end{bmatrix}\begin{bmatrix}\ln\beta_{t-2} \\ d_{t-2}\end{bmatrix} + \begin{bmatrix}0 \\ \varepsilon_t^d\end{bmatrix} \tag{93}$$

Transform it into a VAR(1):

$$\begin{bmatrix}\ln\beta_t \\ d_t \\ \ln\beta_{t-1} \\ d_{t-1}\end{bmatrix} = \begin{bmatrix}\kappa \\ (1-\rho)\mu \\ 0 \\ 0\end{bmatrix} + \begin{bmatrix}\mathcal{A} & \mathcal{C} & \mathcal{B} & \mathcal{D} \\ 0 & \rho & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}\ln\beta_{t-1} \\ d_{t-1} \\ \ln\beta_{t-2} \\ d_{t-2}\end{bmatrix} + \begin{bmatrix}0 \\ \varepsilon_t^d \\ 0 \\ 0\end{bmatrix} \tag{94}$$

or, in matrix form,

$$\boldsymbol{Z}_t = \tilde{\boldsymbol{\kappa}} + \Phi\boldsymbol{Z}_{t-1} + \boldsymbol{\eta}_t \tag{95}$$

with

$$\boldsymbol{\eta}_t \sim \mathcal{N}\left[\begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0 & 0 & 0 & 0\\ 0 & \sigma_d^2 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\right] \tag{96}$$

Assume stationarity. Apply the variance operator on both sides of the last expression. Then,

$$\Gamma = \Phi\Gamma\Phi^T + \tilde{\Omega} \tag{97}$$

where Γ is the covariance matrix of $\boldsymbol{Z}_t$ and $\tilde{\Omega}$ is the covariance matrix of $\boldsymbol{\eta}_t$. Vectorize the matrixes to represent the matrix product using the Kronecker product,

$$\text{vec}(\Gamma) = (\Phi \otimes \Phi)\text{vec}(\Gamma) + \text{vec}(\tilde{\Omega}) \tag{98}$$

Then,

$$\text{vec}(\Gamma) = (I - \Phi \otimes \Phi)^{-1}\text{vec}(\tilde{\Omega}) \tag{99}$$

Solving this equation, we get

$$\mathbb{V}(\ln\beta_t) = \nu\frac{\sigma_d^2}{(1-\rho^2)} \tag{100}$$

with

$$\nu \equiv \frac{2\mathcal{C}\mathcal{D}(\mathcal{A} + \rho(1 - \mathcal{B}^2 + \mathcal{A}\mathcal{B}\rho)) + (\mathcal{C}^2 + \mathcal{D}^2)(1 + \mathcal{A}\rho + \mathcal{B}(\rho(\mathcal{A} + \rho(1 - \mathcal{B}))))}{((1 + \mathcal{B})(1 + \mathcal{A} - \mathcal{B})(1 - \mathcal{A} - \mathcal{B})(1 - \rho(\mathcal{A} + \mathcal{B}\rho))} \tag{101}$$

In turn,

$$\mathbb{C}ov(\ln\beta_t, d_t) = \upsilon\frac{\sigma_d^2}{(1-\rho^2)} \tag{102}$$

with

$$\upsilon \equiv \frac{\rho(\mathcal{C} + \rho\mathcal{D})}{(1 - \rho(\mathcal{A} + \rho\mathcal{B}))} \tag{103}$$

The conditions required for stationarity ensure $\mathbb{V}(\ln\beta_t) > 0$. In turn, it can shown that $\mathbb{C}ov(\ln\beta_t, d_t) > 0$ if[16]

$$g < \frac{1}{\rho(1 - \rho\kappa)} \tag{104}$$

with

$$\kappa \equiv \frac{(1 - \pi\tau)\delta exp\left\{\mu + \frac{\sigma^2 + \sigma_b^2 + \sigma_P^2}{2}\right\}}{1 - (1 - \pi\tau)\delta exp\left\{\mu + \frac{\sigma^2 + \sigma_b^2 + \sigma_P^2}{2}\right\} - \delta\pi\tau} \tag{105}$$

Now, I analyze the effect of taxes on the covariance matrix. It can be shown that $\partial\mathbb{V}(\ln\beta_t)/\partial\tau < 0$ when

$$g < \frac{1}{1 + \kappa} \tag{106}$$

Given $\rho < 1$, it is clear that the inequality (106) is more restrictive than (104). Let $\bar{g} \equiv \frac{1}{1+\kappa}$.

[16]The details of the following derivations are available upon request.

Focus now on the effect of taxes on the covariance:

$$\frac{\partial \mathbb{C}ov(\ln\beta_t, d_t)}{\partial \tau} = -\frac{(\delta-1)exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}g^2\pi\rho^2(\rho^2-\rho-1)}{(1+\rho)\left(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}(-1+\rho-2g\rho+g\rho^2)(1-\pi\tau)+(1+(-1+g)\rho)(1-\delta\pi\tau)^2\right)^2}\sigma_d^2 \tag{107}$$

It is clear that the denominator is positive. The numerator is too since $\delta < 1$ so that $(\delta - 1) < 0$ and $0 < \rho < 1$ implies $\rho^2 < \rho$ and then $(\rho^2 - \rho - 1) < 0$. Hence, the derivative has a negative sign.

B4.- Proof of Proposition 3.

The proposition holds that τ^D negatively influences ω but has no effect on Σ. I start by characterizing $\boldsymbol{\omega}$ at the approximation point $\boldsymbol{x} = (\mu, \mu)$ for the equilibrium PD ratio with learning and a tax on dividend (equation (40) in the main text).

$$\omega_1^L = \rho\frac{\delta(1-\tau^D)exp\Big\{\mu + \sigma_d^2/2\Big\}}{1-\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}} \tag{108}$$

which is always positive given $\tau^D < 1$. Its derivative with respect to τ reads as

$$\frac{\partial \omega_1^L}{\partial \tau} = -\rho\frac{\delta exp\Big\{\mu + \sigma_d^2/2\Big\}}{1-\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}} \tag{109}$$

which is always negative. In addition,

$$\omega_2^L = \frac{\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}(1-\tau^D)}{\left(1-\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}\right)^2} \tag{110}$$

which is positive since $\tau^D < 1$. Take its derivative with respect to τ

$$\frac{\partial \omega_2^L}{\partial \tau} = -\frac{\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}}{\left(1-\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}\right)^2} \tag{111}$$

which is always negative. Hence, τ^D reduces the pass-through from beliefs to prices.

To characterize the effect on Σ is it enough to show that τ^D plays no role in the 2nd order difference equation characterizing the dynamics of $\ln\beta_t$. This equation comes from the combination

of two equilibrium equations, the Kalman updating equation

$$\ln\beta_t = \ln\beta_{t-1} + g\Big(\ln\frac{P_{t-1}}{P_{t-2}} - \ln\beta_{t-1}\Big) \tag{112}$$

and the equilibrium PD ratio, which implies that

$$\frac{P_{t-1}}{P_{t-2}} = \frac{\delta(1-\tau^D)\mathbb{E}^{\mathcal{P}}_{t-1}\left[\frac{D_t}{D_{t-1}}\right]}{\delta(1-\tau^D)\mathbb{E}^{\mathcal{P}}_{t-2}\left[\frac{D_{t-1}}{D_{t-2}}\right]}\frac{1-\delta\mathbb{E}^{\mathcal{P}}_{t-2}\left[\frac{P_{t-1}}{P_{t-2}}\right]}{1-\delta\mathbb{E}^{\mathcal{P}}_{t-1}\left[\frac{P_t}{P_{t-1}}\right]} \tag{113}$$

It is clear that τ^D cancels out and then, τ^D plays no role in the expectation formation process.

Appendix C: Expectations measurement in the long-run

In this appendix, I explore the relationship between the survey-like expectations measure used in the main text and other available measures. One challenge in measuring expectations is that surveys often cover a short time period, typically not extending beyond the late 1980s. To overcome this limitation, I have followed the procedure outlined by Adam, Marcet, and Beutel (2017) in the text to derive a series of survey-like capital gains expectations dating back to 1946. This appendix provides a detailed description of the procedure and conducts a comparison between the computed beliefs and other available measures.

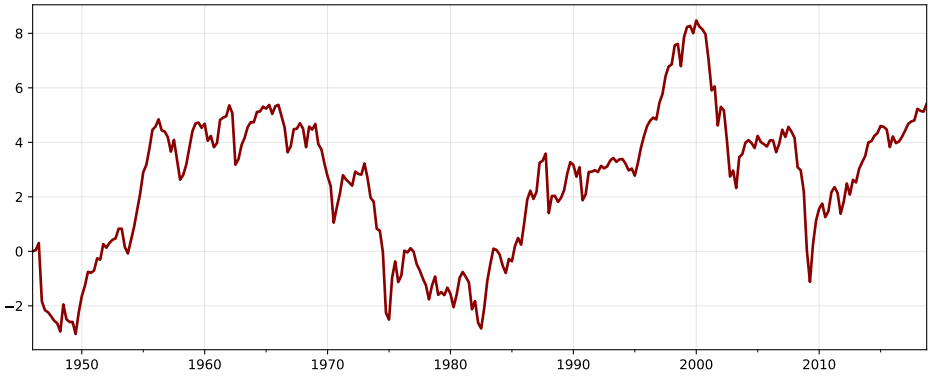
The basic data are derived from the UBS Index of Investor Optimism, which reports the expected one-year-ahead return for the period from Q2 1998 to Q4 2007. As explained in Appendix A, a series of capital gains beliefs can be obtained from the raw data by using dividend growth and price-dividend ratio data. Following Adam, Marcet, and Beutel (2017), I employ an updating algorithm that consists of a Kalman filter, as detailed by equation (47) in the main text. An initial condition, a value for the Kalman gain, and realized price growth data suffice to establish a series of beliefs. The initial condition is set equal to zero growth for 1946:Q1. The Kalman gain is estimated to optimize the fit of the model to historical data as follows:

$$g = \arg\min_{\mathcal{G}} \sum_{t=1998.Q2}^{2007:Q4} \left(\mathcal{R}_t(\mathcal{G}) - \mathbb{E}^{UBS}_t\Big(\frac{P_{t+1}}{P_t}\Big)\right)^2 \tag{114}$$

with $\mathcal{R}_t \equiv \mathbb{E}^s_t\left[\frac{P_{t+1}}{P_t}\right]$, the survey-like beliefs given by equation (47) in the main text. Figure X plots the time series.

This analysis juxtaposes the derived series with other available long-term survey data. First, I

Figure X: ***Survey-like capital gains beliefs from 1946 to 2018.***



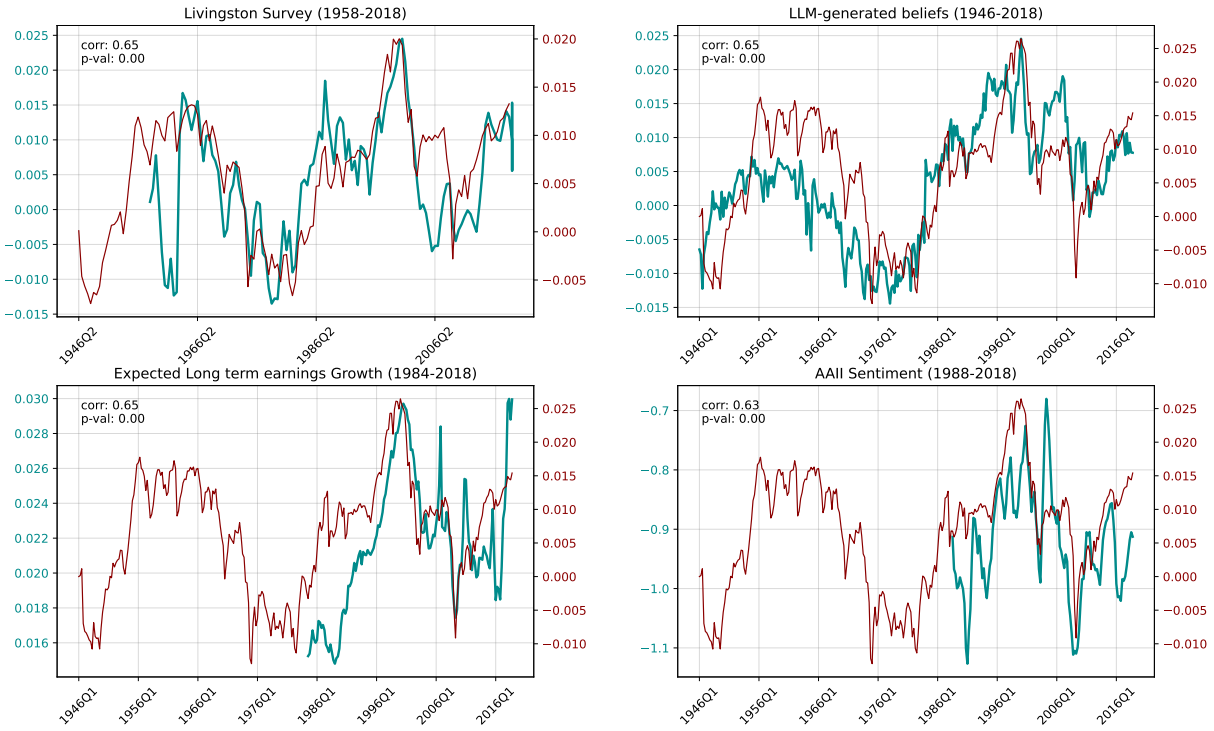
The graph plots the series of capital gains expectations implied by equation (47) in the main text with $g = 0.0276$ *and initial condition of zero growth.*

utilize historical data from the Philadelphia Fed's Livingston Survey on the Stock Price Index. This series, commencing in 1952, presents unique longitudinal insights despite certain inconsistencies (for example, forecasts specifically on the SP 500 Index only begin in 1990, whereas earlier predictions pertained to various iterations of the SP's Industrials). Despite these variations, it is the only available resource for examining trends over the long-run. By calculating the index's changes and adjusting for expected inflation from the same dataset, I extract a metric for anticipated real price growth. Since the data is quite noisy, I employ a five-year moving average to smooth the series, noting that the choice of mean or median calculations does not materially affect the outcomes.

A novel source of data is introduced through the recent work of Bybee (2023), who leverages a Large Language Model to analyze newspaper articles, thereby generating a sentiment indicator akin to the bullish-bearish spread observed in traditional surveys. This innovative approach facilitates the construction of extensive time series, significantly enriching our data set. Given the scarcity of long-run data, I exploit this source to approximate capital gains beliefs. To that end, assume the LLM sentiment measure to be perfectly correlated with a quantitative measure of expected gross returns and rescale it appropriately to match the mean realized returns. Then, a procedure like the one used to obtain capital gains beliefs from the raw UBS data can also be applied to this data.

Additionally, survey data from the 1980s onward provide further insight. The IBES survey, detailing analysts' forecasts on long-term real earnings growth, starts in 1984. The data was obtained from Stephan Nagel's website. In standard payoff discount models of asset prices, beliefs about longer-term growth are equivalent to the beliefs about capital gains. Through those lenses, this measure can be seen as a proxy for capital gains beliefs. Moreover, the American Association

Figure XI: ***Measures of capital gains expectations in the long-run.***



The figure plots the expectations implied by a Kalman algorithm (in red) along with the expectations coming from the Livingston survey, LLM-generated beliefs by Bybee (2023), IBES Long Term Earning Growth forecast and AAII bullish-bearish spread (in cyan). Correlation reported with p-values that uses Newey-West standard errors with a maximum of 4 lags. All the series shows expected quarterly price growth, except for the Livingston panel that shows 6-month forecasts.

of Individual Investors' Sentiment Survey, offering monthly data since 1987, contributes to our understanding of market sentiment. To mitigate the noise within this series, I apply a one-year moving average to the bullish-bearish spread, selecting end-of-quarter values for analysis.

The data, as illustrated in Figure XI, underscores a strong correlation between the beliefs derived from the Learning equation and those obtained through other measures, with correlation coefficients consistently exceeding 0.60 with highly statistical significance. Notably, this correlation remains robust across various sample periods, including both pre- and post-1980s data. Furthermore, the analysis reveals significant correlations not only with direct measures of capital gains expectations, but also with long-term growth forecasts and broader sentiment indicators.

While Bordalo et al. (2023) holds that the Dotcom boom was associated to over-extrapolation of news about fundamentals rather than to price extrapolation, the series generated via the learning from prices exhibits a very similar pattern. This fact suggests either a complementary between the two extrapolations or, potentially, that a way of learning about longer-term fundamentals is precisely by using price data. The same boom is observed in sentiment-based measures.

Appendix D: Computing the non-taxable share

The evolution of the effective capital tax rates depends essentially on two factors: statutory rates and regulations. Legal regulations are accounted for by the NBER TaxSim rates. The important exception is the amount of capital income accruing to non-taxable units, as pension funds, IRAs or non-profit institutions. The Financial Accounts of the United States, run by the Fed, report the household share of corporate equity. Some takes that as a proxy for the taxable share of ownership, but that overestimate it given the inclusion of IRAs (see Rosenthal and Austin (2016) for a critical review of the different measures). Therefore, the goal is to get an estimate of the fraction of equities hold by households in taxable accounts. I follow Rosenthal and Austin (2016).

Table VI reports the steps followed to compute the taxable share. Essentially, it amounts to an adjustment of the Fed's households equity share, considering IRAs, indirect holdings and so on. Here I detail the abbreviations dictionary: CE = corporate equities; HHNPI = households and nonprofit institutions; RoW = rest of the world; ETF = exchange traded fund; CEF = closed-end fund; REIT = real estate investment trust; C-CE = C corporations CE; MF = mutual funds; IRA = investment retirement accounts. The variables comes from the Federal's Reserve Financial Accounts of the United States, except for those variables whose construction is explained in the table. Besides, as in Rosenthal and Austin (2016), the stock held in self- directed IRAs is based on data from the Investment Company Institute. Calculations files are available upon request. Pre-1951, the share is interpolated from Sialm (2009).

Appendix E: Parameterized Expectations Algorithm

The optimality conditions of the model involve the following three functions of subjective conditional expectations. For stocks,

$$\frac{C_t}{D_t} = \left\{ \mathbb{E}_t^{\mathcal{P}} \left[\delta \Big(\frac{C_{t+1}}{D_{t+1}} \Big)^{-\gamma} \Big(\frac{D_{t+1}}{D_t} \Big)^{1-\gamma} \Big(1 - \tau_{t+1}^D + \frac{P_{t+1}}{D_{t+1}} \Big) \frac{D_t}{P_t} - \mu_{t+1} \Big(\frac{P_{t+1}}{D_{t+1}} - \frac{P_t}{D_t} \frac{D_t}{D_{t+1}} \Big) \frac{D_t}{P_t} \right] \right\}^{-1/\gamma} \equiv \mathcal{E}_1(\boldsymbol{X}_t) \tag{115}$$

For bonds,

$$\frac{C_t}{D_t} = \left\{ \mathbb{E}_t^{\mathcal{P}} \left[\delta \Big(\frac{C_{t+1}}{D_{t+1}} \Big)^{-\gamma} \Big(\frac{D_{t+1}}{D_t} \Big)^{1-\gamma} \right] (1 + (1 - \tau_t^b) r_t^b) \right\}^{-1/\gamma} \equiv \mathcal{E}_2(\boldsymbol{X}_t) \tag{116}$$

Table VI: ***Taxable share estimation.***

Step	Operation	Definition
	Total CE HHNPI	
1. Subtract foreign equities	- RoW x (Total CE HHNPI / Total CE All Sectors) = HHNPI domestic CE	
2. Subtract the stocks issued by the passthrough entities	- S Corporations – (ETF + CEF + REITs) x HH share of Mutual Funds = HHNPI domestic C-CE	
3. Subtract NPI holdings	- NPI domestic C-CE = HH domestic C-CE	NPI domestic C-CE = (NPI CE + MF stocks) x NNHPI [CE / (CE + MF)] (NPI CE + MF) given by the Fed after 1987 (CE + MF together). Before: NPI CE + MF = (HHNPI CE + MF) x (NPI CE + MF) 1987 / (HHNPI CE + MF) 1987
4. Subtract IRAs and 529 savings plans holdings	- IRA C-CE - 529 C-CE = Direct HH domestic C-CE	IRA C-CE = CE IRA x C-CE Fraction CE IRA = IRA Other Assets x 0.75 [a] 529 C-CE = College Savings Plans Assets x 0.5 [b] C-CE fraction = All sectors C-CE / (All sectors C-CE + ETF + CEF + REITs) All sectors C-CE = All sectors domestic CE – S corp – ETF - CEF – REITs
5. Add Indirect Holdings of C Corporation Equity	+ Indirect HH domestic C-CE = HH Taxable CE	Indirect HH C-CE = (CE MF + (CE ETF + CE CEF) x HH share of MF) x Direct HH domestic CE / Total HHNPI CE
6. Divide by the total C corporation equity	/ All sector C-CE	
	= Taxable share	

[a] Assumed 75% of IRA other assets are stocks.

[b] Assumed 50% of the assets in college savings plans were C corporation equity.

For the accumulation of unrealized capital gains,

$$\mu_t = \mathbb{E}_t^{\mathcal{P}}\left[\delta\mu_{t+1}\tau^{1+\xi}\left(1 - ln\Big(\frac{\hat{G}_{t+1}}{G_t}\frac{D_t}{D_{t+1}}\Big)\right)\right] \equiv \mathcal{E}_3(\boldsymbol{X}_t) \tag{117}$$

These subjective conditional expectations are generally unknown functions of the state variables. I approximate them using parametric functions of the state variables, as first suggested by Marcet (1988). To pick the approximation functions, I use two different approaches. For the consumption policy, I use a functional form rooted in economic theory, following the insights of Hakansson (1970) which were also used in Williams (2024). In particular, I conjecture

$$\frac{C_t}{D_t} = \mathcal{E}_1(\boldsymbol{X}_t) \approx \psi_1(\boldsymbol{X}_t;\chi) = c_t^y\Big(S_{t-1} + \frac{W_t}{D_t}\Big) + c_t^w\frac{P_t}{D_t}S_{t-1} \tag{118}$$

where $c_t^y \equiv 1 - \chi_1\delta(1-\hat{\tau}_t^D)\beta_t^d$ is the marginal propensity to consume out of income and $c_t^w \equiv 1-\chi_1\delta(1-\chi_2\hat{\tau}_t^{1+\xi})\beta_t$ is the marginal propensity to consume out of wealth. (χ_1,χ_2) are parameters of the approximation function to be estimated. $\psi_1(\boldsymbol{X}_t;\chi)$ has an economic meaning: if agents expect higher after-tax returns, they save more and consume less. Inspired by the analytical results in Section 2, after-tax capital gains beliefs are corrected by the tax elasticity of realization ξ and the parameter χ_2.

Using this function, the risk-free rate can be pinned down by

$$r_t^b = \left(\mathbb{E}_t^{\mathcal{P}}\left[\delta\Big(\frac{\psi_1(\boldsymbol{X}_{t+1};\chi)}{\psi_1(\boldsymbol{X}_t;\chi)}\Big)^{-\gamma}\Big(\frac{D_{t+1}}{D_t}\Big)^{1-\gamma}\right]^{-1} - 1\right)(1-\tau_t^b)^{-1} \tag{119}$$

The conditional expectation is computed by numerical integration considering agents subjective models of wages, prices, dividends and taxes. Numerical integration is performed using the Gauss-Hermite quadrature rule extended to the multidimensional case.

Finally, I use a flexible form to approximate $\mathcal{E}_3(\boldsymbol{X}_t)$. It turns out that the exponential of a second-order polynomial on some of the states performs well

$$\begin{aligned}\mathcal{E}_3(\boldsymbol{X}_t) \approx \psi_3(\boldsymbol{X}_t;\chi) = \chi_3 exp\Big\{&\chi_4\tau_t + \chi_5\frac{P_t}{D_t} + \chi_6\frac{W_t}{D_t} + \chi_7\tau_t^D + \chi_8\beta_t \\ &+\chi_9\tau_t^2 + \chi_{10}\Big(\frac{P_t}{D_t}\Big)^2 + \chi_{11}\Big(\frac{W_t}{D_t}\Big)^2 + \chi_{12}(\tau_t^D)^2 + \chi_{13}\beta_t^2\Big\}\end{aligned} \tag{120}$$

The vector of parameters χ is unknown. It is customary in the PEA literature to estimate it via Montecarlo integration. However, in the current setup this approach is very costly as the structural parameters of the model are themselves to be estimated via SMM. Thus, for every try

of θ, χ should be estimated, being an estimation within an estimation.

The problem is avoided by a guess-and-verify approach. I use an educated guess for χ, say $\tilde{\chi}$, coming from the experience with model simulations. For the given $\tilde{\chi}$, I obtain an estimated $\hat{\theta}$. Then, I verify that $(\tilde{\chi}, \hat{\theta})$ implies an acceptable accuracy level. If accuracy is unacceptable, I use an alternative guess for χ and reestimate θ. This approach differs from conventional PEA, whereby the structural parameters are fixed and the approximation parameters are estimated.

PEA uses the fact that the conditional expectation is the best predictor of the variables within the expectation for the given probability measure. Thus, a good approximation can be evaluated by the size of the prediction errors. Let $\phi\Big(z_{t+1}, \varepsilon_{t+1}, z_t\Big)$ be the vector of expressions inside the conditional expectation, with ϕ being a function of current and future endogenous variables z and shocks ε. Thus, the prediction errors are given by

$$u_t = \phi\Big(z_{t+1}, \varepsilon_{t+1}, z_t\Big) - \psi(\boldsymbol{X}_t; \chi) \tag{121}$$

This vector includes the errors in all the optimality conditions, that is, the two Euler Equations and the pricing equation for unrealized gains. In the spirit of Judd (1992), I normalize the error by the variable to be approximated to have a dimension-free quantity to evaluate the accuracy of the prediction

$$\tilde{M}_1 = \mathbb{E}\left[\left(\frac{u_{1,t}}{\frac{C_t}{D_t}}\right)^2\right] \tag{122}$$

$$\tilde{M}_2 = \mathbb{E}\left[\left(\frac{u_{2,t}}{\frac{C_t}{D_t}}\right)^2\right] \tag{123}$$

$$\tilde{M}_3 = \mathbb{E}\left[\left(\frac{u_{3,t}}{\mu_t}\right)^2\right] \tag{124}$$

The statistic $\tilde{M}$ measure the mean square relative residuals. Mimicking the SMM approach, the expectation is computed using $K = 1000$ simulations of length T, with T being the number of quarters from 1946.I to 2018.IV. The values of the statistics are $(\tilde{M}_1, \tilde{M}_2, \tilde{M}_3) = (0.0006, 0.0012, 0.0024)$. Thus, the prediction errors that agents make by using the approximating functions are of a magnitude of around 1 out of 1000.

Appendix F: Details of the Quantitative Model

F.1. Learning about dividends and taxes

Keeping the structure of the subjective model of prices given by equations (17) and (18), the subjective model of dividends growth reads as

$$\ln\frac{D_t}{D_{t-1}} = \ln\tilde{d}_t + \varepsilon_t^d \tag{125}$$

$$\ln\tilde{d}_t = \ln\tilde{d}_{t-1} + \nu_t^{\tilde{d}} \tag{126}$$

where $\tilde{d}_t$ is unobserved and innovations are Normal i.i.d. with constant variances σ_d^2 and $\sigma_{\tilde{d}}^2$ respectively. As with prices, they use a Kalman algorithm to filter out the noise and the posterior expectations with period t information are given by $\ln\tilde{d}_t|\omega^t \sim \mathcal{N}(\ln\beta_t^d, \sigma^2)$ with

$$\ln\beta_t^d = \ln\beta_{t-1}^d + g\Big(\ln\frac{D_t}{D_{t-1}} - \ln\beta_{t-1}^d\Big) \tag{127}$$

It follows that

$$\mathbb{E}_t^{\mathcal{P}}\Big[\frac{D_{t+1}}{D_t}\Big] = \beta_t^d exp\{\frac{\sigma^2 + \sigma_d^2 + \sigma_{\tilde{d}}^2}{2}\} \tag{128}$$

Likewise, investors think of taxes using an unobserved component model which implies

$$\boldsymbol{\tau}_t = \tilde{\boldsymbol{\tau}}_t + \varepsilon_t \tag{129}$$

$$\tilde{\boldsymbol{\tau}}_t = \tilde{\boldsymbol{\tau}}_{t-1} + \varepsilon_t^q \tag{130}$$

where $\boldsymbol{\tau} = [\tau, \tau^D, \tau^B]'$, $\tilde{\boldsymbol{\tau}}$ is unobserved and innovations are i.i.d. Normally distributed with zero mean and constant variance. As before, the posterior distribution of the unobserved component is given by $\tilde{\boldsymbol{\tau}}_t|\omega^t \sim \mathcal{N}(\hat{\boldsymbol{\tau}}_t, \Omega^\tau)$, with Ω^τ being the steady state Kalman filter uncertainty. Then

$$\mathbb{E}_t^{\mathcal{P}}\Big[\boldsymbol{\tau}_{t+1}\Big] = \hat{\boldsymbol{\tau}}_t \tag{131}$$

and

$$\hat{\boldsymbol{\tau}}_t = \hat{\boldsymbol{\tau}}_{t-1} + g(\boldsymbol{\tau}_t - \hat{\boldsymbol{\tau}}_{t-1}) \tag{132}$$

Table VII: ***Alternative estimations of the parameters.***

Parameters included in SMM	$\bar{\theta}$				
		Baseline	No change in σ_W	$\xi = -0.70$	$\xi = -0.70$, $t < 1982$
Average wage-dividend ratio	w	13.86	14.79	16.52	15.19
Standard deviation WD shock (pre-1982)	σ_W	0.0543	0.0582	0.0559	0.0574
Standard deviation WD shock (post-1982)	σ_W	0.0649	0.0582	0.0693	0.0574
Kalman gain	g	0.0287	0.0285	0.0292	0.0273
Tax elasticity of realization	ξ	-0.51	-0.61	-0.70	-0.70

The table reports the value of the parameters included in the SMM estimations for the baseline version of the model and alternative versions, always with $M = 18$.

F.2. Projection facility

The equilibrium PD ratio given by 55 faces a discontinuity. For this reason, simulation requires to set up a following modified belief updating equation to ensure non-negative prices. Following Adam, Marcet, and Nicolini (2016),

$$\beta_{t+1} = w\left(exp\left\{ln\beta_t(1-g) + gln\frac{P_t}{P_{t-1}}\right\}\right) \tag{133}$$

where

$$w(x) = \begin{cases} x & \text{if} \quad x \leq \beta_t^L \\ \beta_t^L + \frac{x-\beta_t^L}{x+\beta_t^U-2\beta_t^L}(\beta_t^U - \beta_t^L) & \text{if} \quad x > \beta_t^L \end{cases} \tag{134}$$

and

$$\beta_t^q = \frac{PD^q\bar{S}_t - \chi_2\delta(1-\hat{\tau}_{t+1}^d)\beta_t^d\left(\frac{W_t}{D_t} + \bar{S}_{t-1}\right)}{\chi_2\delta(1-\chi_1\hat{\tau}_t^{1+\xi})PD^q\bar{S}_{t-1}} \tag{135}$$

for $q = L, U$. Thus, this projection facility starts to dampen belief coefficients that imply a price-dividend ratio equal to PD^L and sets an effective upper bound at PD^U. Projection facilities are usual devices in this sort of algorithms (see Ljung (1977)); particularly, (134) is similar to the one used by Adam, Marcet, and Nicolini (2016). It can be understood in a Bayesian sense, so that agents attach zero probability to beliefs coefficients implying a PD ratio higher than PD^U. In simulations, I use $PD^L = 500$ and $PD^U = 700$ such that the facility is only used very rarely; on average only the 0.1% of the periods of each simulation.

F.3. Alternative estimations of the parameters

Appendix G: The Equity Premium

The model aligns with both the level and the change in the equity premium, as illustrated in table III in the main text. It achieves this by utilizing realistic consumption and dividend growth processes, a non-negative discount factor, and a low risk aversion coefficient—key components of the risk premium puzzle identified by Cochrane (2017). This appendix explores why.

I study the accounting drivers of stock returns both in the data and in the model through the following decomposition of their geometric mean, first suggested in Adam, Marcet, and Nicolini (2016):

$$\left(\prod_{t=1}^{N}\frac{P_t+D_t}{P_{t-1}}\right)^{\frac{1}{N}}=\underbrace{\left(\prod_{t=1}^{N}\frac{D_t}{D_{t-1}}\right)^{\frac{1}{N}}}_{R_1}\underbrace{\left(\frac{PD_N+1}{PD_0}\right)^{\frac{1}{N}}}_{R_2}\underbrace{\left(\prod_{t=1}^{N-1}\frac{PD_t+1}{PD_t}\right)^{\frac{1}{N}}}_{R_3} \tag{136}$$

The mean gross return can be understood as the product of three elements. The first term (R_1) is the mean dividend growth. The second term (R_2) is the ratio of the terminal to the initial value of the PD ratio, which may be related to the existence of a time trend. The last term (R_3) is a convex function of the PD ratio in period t that increases with the volatility of the PD time series and decreases with its mean.

Table VIII presents the decomposition using both empirical and simulated data. Since the dividend growth process has been parameterized directly from the data, the models reproduce R_1 accurately. Regarding R_2, the model captures its increase in the second subperiod but overstates its value in the first subperiod, as tax cuts began before 1982, causing the model to produce a slight increase before that period. The model also replicates the decline in R_3, although not as strongly as observed in the data, because the increase in the mean PD ratio produced by the model is insufficiently strong. Overall, the model captures the direction of the change in the different components, despite not targeting any of these statistics specifically.

While consumption-based asset pricing models typically require either a high level of risk aversion or high fundamental volatility to match mean returns, this paper explores alternative forces. In line with the Learning literature, non-fundamental volatility arising from subjective beliefs reconciles realistic income processes and low risk aversion with high and volatile stock returns (Adam, Marcet, and Nicolini (2016), Adam, Marcet, and Beutel (2017)). However, belief volatility alone cannot account for the entire effect. Adam, Marcet, and Nicolini (2016), for instance, achieve only about half of the equity premium, partly due to a too low R_2. The introduction of tax cuts proved

Table VIII: ***Decomposition of the stock return geometric mean.***

	US data		Baseline model	
	1946-1982	1982-2018	1946-1982	1982-2018
$\mathbb{E}(R_1)$	1.0183	1.0310	1.0183	1.0310
$\mathbb{E}(R_2)$	0.9876	1.0275	1.0075	1.0209
$\mathbb{E}(R_3)$	1.0422	1.0239	1.0480	1.0362
$\mathbb{E}(1+r_t^s)$	1.0481	1.0846	1.0752	1.0906

The table shows the stock returns mean decomposition according to expression (136). The first column uses U.S. data; the second, simulated data using the model with baseline the parameters in table II. Rates are annualized.

to be fruitful in two dimensions. First, it imparts a trend in the PD ratio, helping to obtain a larger R_2. Second, tax cuts contribute to a higher mean and variance in the second subperiod, inducing a fall in R_3.

In summary, the interaction of subjective beliefs and tax cuts can account for the level and change in stock returns since 1946. Crucially, these elements do not influence the determination of the risk-free rate, thereby avoiding a new version of the risk-free rate puzzle.