Capital Gains Taxation, Learning and Bubbles

PAU BELDA*
Federal Reserve Board

Latest version here.

Abstract

The paper argues that a low capital gains tax fuels asset price instability. The reason is that a lower tax increases the price sensitivity to investors' beliefs which, when agents learn from prices, makes self-fulfilling booms and busts more likely. Applying this theory to the United States, I establish a connection between capital gains tax cuts and the proliferation of stock market fluctuations since the 1980s. The structurally estimated model is consistent with many asset pricing facts and suggests that tax cuts account for 40% of the observed rise in the S&P 500's price-dividend ratio volatility. Moreover, while the model replicates the observed decrease in excess return predictability, it also implies a decline in the stock market's informational efficiency since the 1980s.

Keywords: Capital Taxation, Asset Pricing, Learning, Predictability, Informational Efficiency.
JEL codes: D83, D84, E32, E44, E62, G12, G14.

*Economist at the Federal Reserve Board. Email: pau.belda@frb.gov. First version: July 2020. This version: February 2024. I am especially grateful to Albert Marcet for his guidance during the PhD, of which this paper was part. Besides, I have benefited from comments from many people at bilateral meetings, seminars, and conferences. A special mention goes to Luis E. Rojas and Jordi Caballé, from the UAB. 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 (ERC) under the European Union's Horizon2020, research and innovation program GA project number 788547 (APMPAL-HET). The views expressed on this paper are only personal and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System. The usual disclaimers apply.

1.- Introduction

Economists have long been in search of a tax instrument to mitigate financial instability. A Financial Transactions Tax has been a prominent proposal, endorsed by influential figures such as Keynes (1936), Tobin (1978) or Stiglitz (1989). Despite its widespread popularity, research has consistently indicated that, while it may curb speculative trading, its impact on price volatility remains unclear.[1] Against this backdrop, this paper argues for the Capital Gains Tax (CGT) as a more effective tool in ensuring asset price stability, offering new theoretical insights and empirical evidence to support its efficacy.

Capital gains are at the heart of a long-standing view in finance that characterizes booms and busts as self-fulfilling prophecies. The sequence goes as follows: Initial positive news ignites investor optimism, leading to an increase in demand for the asset and a rise in its price; these gains further stoke the initial optimism, thereby reinforcing the process. However, at a certain point, a "Minsky Moment" occurs, flipping this cycle into a downward phase of pessimism and price deflation.[2] This paper argues that a CGT mitigates this expectations-price spiral. Specifically, a higher tax would cool down the optimism induced by good news, thereby reducing the pass-through from beliefs to asset demands and prices; on the contrary, a lower tax exacerbates the effect of bullish beliefs on asset prices.

I formalize these points using an asset pricing model à la Lucas (1978) with a probabilistic tax on realized capital gains. I start with a version that is general (nesting models with habits, long-run risk, or subjective beliefs), but takes the sequence of expectations as given. In a world with taxes, agents need to pencil in the probability of tax payments and, as a result, the tax rate times the probability of tax payments appear as a wedge between (pre-tax) expectations and prices. Thus, the tax level shields price volatility from belief fluctuations. The finding holds under minimal constraints on the given trajectory of beliefs and is compatible with many available expectations models.

Nonetheless, as suggested above, taxes could also influence beliefs dynamics. To explore that possibility, the model is specialized, and beliefs are made model-consistent. First, I assume full information rational expectations. However, since this hypothesis has been consistently rejected when survey expectations are used (e.g., Greenwood and Shleifer (2014), Adam et al. (2017), Bordalo et al. (2019)), I then relax the information endowments. In particular, agents have only

[1] See, for instance, Buss et al. (2016) , Buss and Dumas (2019), or Dávila (2023) for theoretical analysis and Umlauf (1993) or Cappelletti et al. (2017) for empirical results that cast doubts on the ability of a tax on financial transactions to stabilize asset prices.

[2] Keynes (1936), Minsky (1976) or Kindleberger (1978) are classic references of this view.

imperfect market knowledge such that they cannot deduce the true model of price formation despite behaving rationally; instead, they learn about it from price data as in Adam and Marcet (2011).[3] In this case, an expectations-price spiral emerges, with an Euler Equation connecting expectations to prices and Bayesian updating linking prices back to expectations.[4] This expectations formation process replicates key features of survey data such as extrapolation and sluggish adjustments.

With full information rational expectations, beliefs fluctuations only reflect fundamental risk. Similarly to the case with given beliefs, taxes reduce the transmission of fundamental risk to price volatility. Under learning, though, expectations vary not only due to fundamental risk, but also due to self-referential logic. It turns out that this self-fulfilling dynamic is influenced by the tax level. For instance, a lower tax makes expectations self-fulfillment easier, as prices react more to a change in beliefs (as under rational expectations), which in turn induces further changes in beliefs. Thus, a low tax eases the de-anchoring of beliefs from their fundamental value, exacerbating the non-fundamental component of price volatility.

Although this theory finds precedent in Haugen and Heins (1969) and Haugen and Wichern (1973), the most common view among this scarce literature is that a CGT destabilizes asset markets (e.g., Somers (1948), Somers (1960), Gemmill (1956), Stiglitz (1983)).[5] Since higher taxes increase the penalty of selling assets, investors might hold their assets longer than otherwise, which has been known as the lock-in effect. This effect would become particularly acute when price booms increase the amount of unrealized gains, as the lock-in decreases the available supply of shares in the market, intensifying upward price pressures. To address this concern, the probabilistic tax is replaced with an explicit decision about when to realize the stock of capital gains. In particular, each investor manages a stock of unrealized capital gains facing accumulation costs as in Gavin et al. (2007) and Gavin et al. (2015), such that the deferral of gains realization is penalized.

If the lock-in effect is sufficiently pronounced, a high tax rate could significantly reduce the realization of capital gains. This reduction may lead to a diminished tax burden, completely counteracting the stabilization effect previously mentioned. The magnitude of the lock-in effect hinges on the tax elasticity of realization, a key parameter within the cost function. Particularly,

[3] Other papers have emphasized the idea of learning about fundamentals, often linked to new technologies, as Pástor and Veronesi (2006) or Pástor and Veronesi (2009). Along this line, Bordalo et al. (2023) pointed out that over-reaction of long term earnings expectations generates booms and busts in stock prices. These mechanisms share with Learning the idea of subjective beliefs fluctuations driving excess price volatility. A tax on capital gains (which in these models are related to longer term payoffs) would affect price volatility similarly.

[4] In the model, stock demand is a mediator: higher taxes makes demand less responsive to beliefs, which translate into less responsive prices. This is entirely consistent with the empirical evidence in Giglio et al. (2021) that, using survey and administrative data from Vanguard investors, showed that they adjust their portfolio following their beliefs less when they face higher capital gains taxes.

[5] See Dai et al. (2008) for a review on the effects of a CGT on the price and returns *levels*.

if this elasticity is markedly negative, surpassing a threshold value of -1, the lock-in effect becomes predominant. Thus, the model can yield an increase in stability or instability when taxes are cut depending on the values of the tax elasticity of realization.[6]

In the second part of the paper, this theory is employed to examine the recent history of the US stock market. A remarkable fact has been that, despite the Great Moderation in macroeconomic aggregates, stock market volatility did not fall. This disconnect between the stock market and the macroeconomy is an observation difficult to explain for many macro-finance models. For instance, in models with external habits as Campbell and Cochrane (1999), a less volatile consumption growth would lead to a more stable risk premium and prices; in models with long-run risk as Bansal and Yaron (2004), lower risks might account for a run-up in valuations but not for larger swings; even in models of subjective beliefs as Adam et al. (2016), smaller macroeconomic shocks would lead to smaller forecast errors, yielding more stable beliefs and prices. While a popular belief is that the prolonged decline in the risk-free rate would have prompted the emergence of rational bubbles (e.g., Martin and Ventura (2018)), this paper suggests another complementary possibility: the successive cuts in the CGT observed since the late 1970s would have destabilized the stock market, counteracting the stabilizing effects of lower macroeconomic risk.[7,8,9]

To quantitatively evaluate this hypothesis, I first outline five facts characterizing changes in the stock market since the 1980s with respect to the previous four decades. Apart from the lack of a Great Moderation in the stock market (e.g., the variance of the log PD ratio doubled after the 1980s), excess stock return predictability went down substantially, mostly related to a change from a positive to a negative correlation between returns and dividend growth. Furthermore, the valuation levels became exuberant as a result of a strong surge in capital gains the outpaced the increase in mean dividend growth, but the equity premium was very similar to the pre-1980s level. Finally, the response of stock prices to identical shocks was about 60% higher since the 1980s and prices also exhibited a larger sensitivity to investors beliefs as measured by surveys.

[6]Sialm (2006) also included taxes in a Lucas (1978) framework. However, he focuses on a flat "consumption tax" that affects both the acquisition of new shares and investors' wealth. On the contrary, I focus on dividends and capital gains taxes with clear empirical counterparts. Besides, in their setup tax changes drive all the action while I emphasize the role of the tax level in regulating the overall volatility.

[7]While the risk-free was low around the 2000s, the condition of a risk-free rate below the economy growth rate would also be met in the 1950s and 1960s, but bubbles did not follow in those decades. The quantitative investigation of how much rational bubbles can explain and complement the tax theory of this paper is left for future research.

[8]The average effective marginal tax rate on capital gains fell from a maximum of 18% in the early 1970s to a minimum of 5% in 2008. This fall was the consequence of statutory tax reforms, in particular in 1978 and 1997, but also of different regulatory changes that encouraged a movement of assets to tax-free accounts. See Section 3.

[9]The link between lower taxes and volatility echoes a previous episode: the rise and fall of Wall Street in the Roaring Twenties was preceded by a cut in the top Capital Gain Tax (CGT) rate from 73% to 12.5% in 1921. In the opposite direction, Japan introduced a tax on stock capital gains in 1989 to curve its bubble.

Then, a quantitative version of the model, that includes stochastic taxes and stock repurchases, is structurally estimated via an extension of the Simulated Method of Moments, following Adam et al. (2016). A remarkable observation is the model's ability to pass almost all individual t-tests, with t-stats generally below 2, indicating a robust fit to the data. Importantly, the estimated parameters are within the range of other studies. In particular, the baseline estimated tax elasticity of realization is -0.30, within the range obtained by Agersnap and Zidar (2021); the risk-aversion is slightly above 1, as in many macroeconomic applications; and the learning speed is 0.03, virtually identical to the estimates using real-world data.

I use the model to estimate the average effect of capital gains tax cuts. For this purpose, I simulate a "no tax cuts" scenario, holding the tax rate steady at its average pre-1980s level throughout the entire period following the 1980s. Then, I calculate the difference in a specific statistic (e.g., the variance of the PD ratio) between scenarios with tax cuts and without them and compare it to the observed change in the empirical data. Thus, a relative average treatment effect is obtained. For the full-fledged model, tax cuts account for 40% of the observed increase in the volatility and 10% of the higher mean PD ratio.[10]

If the lock-in effect is ignored, the impact of tax cuts is magnified, generating an increase in volatility equal to 80% of the observed one. This provides a quantitative measure of the lock-in effect relevance. On the contrary, a model with Rational Expectations attributes 25% of the increase in volatility to tax cuts, which illustrates that although learning is crucial to have a better data fit, the basic mechanism linking lower taxes to greater volatility is not learning-dependent.[11,12] Alternative assumptions about taxes such as near-term tax foresight or tax learning do not significantly change the estimated effect.

While the paper has emphasized the role of capital gains taxes, dividend tax cuts were also important. In line with McGrattan and Prescott (2005), the model shows that the dividend tax cuts were relevant to explain the increase in the PD ratio, accounting for 20% of the valuation run-up; however, their contribution to the increase in the variance is more modest, below 10%. This result suggests a complementarity between dividends and capital gains taxes in terms of their effect in the mean and variance of stock prices, respectively. Furthermore, the model attributes

[10] This complements the time series evidence presented in Sialm (2009) and Brun and González (2017) relating lower taxes to higher price levels.

[11] There is a large literature on how learning might contribute to solving asset pricing puzzles. Some references are Timmermann (1993), Bullard and Duffy (2001), Cogley and Sargent (2008), Adam et al. (2016), Adam et al. (2017), Jin and Sui (2022).

[12] However, the version with Rational Expectations perform poorly in other dimensions. In particular, the PD ratio under Rational Expectations falls short of serial correlation, as it inherits the correlation of the dividend growth process that is notably lower than that of the PD ratio.

an important role to shares buybacks, which explain about 35% of the increase in the PD ratio volatility. While this figure resembles the one for CGT cuts, the channel through which it operates is different. Using a Campbell and Shiller (1988) decomposition, it turns out that CGT cuts (buybacks) affect the covariance between future returns (future dividends) and current prices.

In line with the model's ability to mirror the variance decomposition of the PD ratio, it also effectively captures the observed reduction in the predictability of excess stock returns. However, the interpretation of this decline as an increase in efficiency is generally problematic by the Joint Hypothesis Problem (Fama (1970)). Instead, I use the model to explore the implications of tax cuts for informational efficiency. While in the model with learning prices are not informationally efficient as they reflect the imperfect information of investors (Adam et al. (2017)), the version with rational expectations represents an efficient benchmark. Then, a measure of efficiency boils down to tracking the deviations of the model with learning from the one with rational expectations. It turns out that in the full-fledged model, the tracking error exhibits a 60% escalation since the 1980s. CGT cuts account for 58% of this reduction in informational efficiency as lower taxes make prices more sensitive to non-fundamental fluctuations coming from subjective beliefs.

While the main topic of the paper is stock market volatility, the model has also some implications for the equity premium. It generates an equity premium of about 4% with a relatively low risk-free rate, realistic consumption and dividend growth processes, a non-negative discount factor and a low risk aversion coefficient, which are the elements of the equity premium puzzle identified by Cochrane (2017). Differently from standard macro-finance models that typically need either a high risk-aversion level or high fundamental volatility to match the mean returns, this paper resorts to two alternative forces. In line with the Learning literature, non-fundamental volatility coming from subjective beliefs makes compatible realistic income processes with high and volatile stock returns (Adam et al. (2016), Adam et al. (2017)). Besides, the inclusion of tax cuts (as suggested in McGrattan and Prescott (2003)) imparts a trend on the PD ratio that helps generating high returns without exaggerating non-fundamental volatility.

Finally, the model offers a potential rationale for the existent statistical cross-sectional evidence. Using a difference-in-difference approach around the 1978 and the 1997 CGT cuts, Dai et al. (2013) documented that the stock returns volatility of portfolios with large unrealized capital gains or non-dividend paying stocks went significantly up with respect to portfolios with less unrealized gains or dividend-paying securities. They suggest that this increase in volatility following tax cuts might be related to a reduction in the risk-sharing with the government, an idea going back at least to Lerner (1943) and further explored in Sikes and Verrecchia (2012). However, simulations of the estimated model suggest that if taxes are not rebated, investors' income ends up being more -not less- volatile, as fluctuations in the realization of capital gains then have an actual impact on investors' budgets.[13] In contrast, the idea of tax cuts boosting price-expectations spirals offers an alternative explanation.

The rest of the paper proceeds as follows. Section 2 studies the relationship between a CGT and asset price volatility in an asset pricing model. Section 3 applies that theory to examine the recent history of the US stock market, presenting estimates of the effects of tax cuts on stock market volatility. Section 4 briefly concludes by summarizing the results and suggesting some venues for future research.

2.- Theory

This section examines the impact of the CGT on the volatility of the PD ratio using an asset pricing model. In Section 2.1. a general consumption-based model is set up. It includes a tax on realized capital gains, where realization is driven by liquidity shocks. Keeping the generality of this model, Section 2.2. shows how taxes influence the pass-through from beliefs fluctuations to prices, taking beliefs as given. In Section 2.3., the model is specialized to derive additional results when beliefs are internal to the model. It is shown that when agents learn from prices, beliefs dynamics are also affected by taxes. Finally, in Section 2.4., investors also decide when to realized the accumulated capital gains so that taxes give rise to the lock-in effect, which counteracts the previous results to a degree determined by the tax elasticity of realization.

2.1.- The model

In this section, a general consumption-based asset pricing model is set up. Its basic layer is the Lucas (1978)'s tree model, generalize to allow for the presence of habits, different fundamental growth processes, a general probability measure and a tax on realized capital gains.

The economy is populated by a unit mass of infinitely living investors. This is a stochastic exchange economy with 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

[13] Furthermore, the tax law treats gains and losses very asymmetrically. Investors can only claim a compensation for losses up to $\$3,000$ which is a tiny amount, while the rest is transferred as a tax credit to compensate for future tax liabilities. In other words, the tax credit is transferred to future idiosyncratic good states (for instance, states with gains), being then a poorly risk-sharing device. I thank Andrea Prestipino for pointing this out.

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. $z^i_t \sim Bernoulli(\pi)$ is a random variable indicating that possibility, with π being the probability of that event occurring. If the event materializes ($z^i_t = 1$), investor i has to sell all her stock holdings. There is a tax $\tau \in [0, 1)$ on realized capital gains.[14,15] Capital gains and losses are treated symmetrically. Tax revenues are transferred back to each individual according to her contribution, represented by T^i_t.

Investors know the stochastic processes followed by D_t and z^i_t. Besides, they take T^i_t 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 $\omega = \{D_t, z^i_t, T^i_t, P_t\}^\infty_{t=0}$ as a typical element, $\mathcal{B}$ denotes the σ-algebra of Borel subsets of Ω and $\mathcal{P}$ 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 $\{C^i_t, S^i_t, X^i_t\}^\infty_{t=0}$ by solving the following optimization program:

$$\max_{\{C^i_t, S^i_t, X^i_t\}^\infty_{t=0}} \mathbb{E}^{\mathcal{P}^i}_0 \sum_{t=0}^{\infty} \delta^t U(C^i_t, H^i_t) \tag{1}$$

subject to the budget constraint

$$C^i_t + P_t X^i_t \leq D_t S^i_{t-1} + z^i_t \Big(S^i_{t-1} P_t - \tau (G^i_{t-1} + (P_t - P_{t-1}) S^i_{t-1}) \Big) + T^i_t \tag{2}$$

the stock holdings law of motion

$$S^i_t = (1 - z^i_t) S^i_{t-1} + X^i_t \tag{3}$$

[14]Note realization is exogenous, driven by the liquidity shock. In section 2.3, I endogeneize π by including portfolio adjustment costs. The introduction of this liquidity shock is equivalent to assume that investors expect that a fixed proportion of the capital gains will be realized as in Sialm (2009).

[15]In equilibrium, a fraction π of agents sells their assets and pay taxes (let $Z_t = \frac{1}{n}\sum_{i=1}^n z^i_t$; by the LLN, $Z_t \xrightarrow{p} \pi$); hence, the effective rate on total capital gains is $\pi \tau^K$.

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 internal and external habit formation (e.g., 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 with exogenous beliefs

In the previous model, the following Euler Equation 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 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:[16]

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

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. Assuming for a moment that $\boldsymbol{\beta}_t$ is given, the following proposition holds.

Proposition 1. *Let Σ be the beliefs covariance matrix and $\mathbb{V}(P_t/D_t)$ the variance of the PD ratio. Assume*

[16] 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).

i. $\beta_t^P, \beta_t^M < 1$;

ii. $\beta_t^P - \beta_t^M > \frac{1-\beta_t^P}{3\pi\tau}$;

iii. $\mathbb{V}ar(\beta_t^M) < \mathcal{Q}\mathbb{V}ar(\beta_t^P)$, *with* $\mathcal{Q} \equiv \frac{(1-\beta_t^P+\pi\tau(\beta_t^P-\beta_t^M)+2(1-\pi\tau)(\beta_t^P-\beta_t^M))(1-\pi\tau)}{(1-\beta_t^P+\pi\tau(\beta_t^P-\beta_t^M)-2\pi\tau(\beta_t^P-\beta_t^M))\pi\tau}$.

Then, the matrix of cross-derivatives $\frac{\partial}{\partial\tau}\nabla_\Sigma\mathbb{V}(P_t/D_t)$ *that measures the sensitivity of* $\mathbb{V}(P_t/D_t)$ *to* Σ *is dominated by its negative elements.*

Proof. Appendix B.1

Conditions i.-iii. imposes some limits on beliefs. Condition i. puts an upper bound, which is analog to the one required in Adam et al. (2016) to ensure positive prices. Condition ii. requires the gap between beliefs about capital gains and risk changes to be above certain threshold; if this gap is too small or even negative, investors expect a subsidy rather than a tax, so τ boosts returns rather than diminishing them. In a boom, the threshold becomes smaller as β_t^P approaches 1; besides, the gap $\beta_t^P - \beta_t^M$ increases as price growth expectations increases while expected risk decreases. Finally, iii. requires that beliefs about future risk are not too volatile with respect to risk-adjusted capital gains. The scaling constant $\mathcal{Q}$ is typically larger than 1 (all it takes is $\pi\tau < 0.5$; historical values are well below this threshold), so the condition is rather permissive for expected risk variations. Note that these three conditions are also sufficient for the pricing formula (6) to have a positive denominator.

The proof in Appendix B.1 shows that the variance of the PD ratio can be approximated as

$$\mathbb{V}\left[\frac{P_t}{D_t}\right] \approx \boldsymbol{\omega}(\tau)\boldsymbol{\Sigma}\boldsymbol{\omega}(\tau)^T \tag{7}$$

where $\boldsymbol{\omega} \equiv \nabla p(b)$ with $\boldsymbol{b} \in \mathbb{R}^3$ being the approximation point; $\omega(\tau)$ is used to highlight the dependence of ω on taxes. To understand how τ affects the mapping from Σ to $\mathbb{V}(P_t/D_t)$, I look at the signs and magnitudes of the elements of $\frac{\partial}{\partial\tau}\nabla_\Sigma\mathbb{V}(P_t/D_t)$. Conditions i-iii ensure almost all its elements are negative. The only exception is the element in the right bottom corner, $\frac{\partial^2 P_t/D_t}{\partial\beta_t^M\partial\tau}\frac{\partial P_t/D_t}{\partial\beta_t^M}$, which might be positive under plausible parameter values. Condition iii. ensures that the potential counteracting effect of this element is always dominated. Thus, it can be concluded that a lower tax makes prices more sensitive to beliefs dynamics.

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 an insight into the direct effect of a CGT on price stability across asset pricing theory. Nonetheless, beliefs might be affected in different ways by taxes, potentially amending *Proposition 1*. In the next section, I study this indirect effect in particular setups.

2.3.- Particular results 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:

$$d_t = (1-\rho)\mu + \rho d_{t-1} + \varepsilon_t^d \tag{8}$$

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.

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:

$$P_t = \mathbb{E}_t \left[\sum_{j=1}^{\infty} \Big(\frac{\delta}{1 - \delta \mathbb{E}_t(z_{t+1}\tau)} \Big)^j D_{t+j} \prod_{l=1}^{j-1} (1 - z_{t+l}\tau) \right] \tag{9}$$

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

$$\frac{P_t^{RE}}{D_t} = \sum_{j=1}^{\infty} f_0(\tau, j) f_1(d_t, j) \tag{10}$$

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 $\left(\frac{\delta}{1-\delta\pi\tau}\right)(1-\pi\tau)exp\left\{\mu + \frac{\sigma_d^2}{2(1-\rho)^2}\right\} < 1$.[17] 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

$$\frac{P_t^{RE2}}{D_t} = \frac{\delta exp\{\mu + \sigma_d^2/2\}}{1-(1-\pi\tau)\delta exp\{\mu+\sigma_d^2/2\} - \delta\pi\tau} \tag{11}$$

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.

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

$$P_t = \frac{\delta}{1-\delta\mathbb{E}_t^{\mathcal{P}^i}(z_{t+1}\tau)}\mathbb{E}_t^{\mathcal{P}^i}\left[D_{t+1} + P_{t+1}(1 - z_{t+1}^i\tau)\right] \tag{12}$$

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

$$\frac{P_t}{D_t} = \frac{\mathbb{E}_t^{\mathcal{P}}\left[\frac{D_{t+1}}{D_t}\right]}{1 - \mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}(1-z_{t+1}\tau)\right] - \delta\tau\mathbb{E}_t^{\mathcal{P}}(z_{t+1})} \tag{13}$$

that is, the ratio depends on beliefs about dividends and price growth and the probability of

[17] 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, $\mathbb{C}ov_t\left(\prod_{l=1}^{j}\frac{D_{t+l}}{D_{t+l-1}}, \prod_{l=1}^{j-1}(1-z_{t+l}\tau)\right) = 0$.

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 (13) 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[18]

$$\ln\frac{P_t}{P_{t-1}} = \ln b_t + \varepsilon_t^P \tag{14}$$

$$\ln b_t = \ln b_{t-1} + \nu_t^b \tag{15}$$

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 $\ln b_t|\omega^{t-1} \sim \mathcal{N}(\ln\beta_{t-1},\sigma^2)$ 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 given by $\ln b_t|\omega^t \sim \mathcal{N}(\ln\beta_t,\sigma^2)$ with[19]

$$\ln\beta_t = \ln\beta_{t-1} + g\Big(\ln\frac{P_{t-1}}{P_{t-2}} - \ln\beta_{t-1}\Big) \tag{16}$$

and

$$g \equiv \frac{\sigma^2}{\sigma^2+\sigma_P^2} \tag{17}$$

Hence, subjective price growth expectations are given by

$$\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}\right] = \bar{\rho}\beta_t \tag{18}$$

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

$$\mathbb{E}_t^{\mathcal{P}}\left[z_{t+1}\frac{P_{t+1}}{P_t}\right] = \mathbb{E}_t^{\mathcal{P}}\left[z_{t+1}(\bar{\rho}\beta_t + u_t + \vartheta_{t+1} + \varepsilon_{t+1}^P)\right] = \mathbb{E}_t^{\mathcal{P}}\left[z_{t+1}\right]\bar{\rho}\beta_t = \pi\bar{\rho}\beta_t \tag{19}$$

[18] 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.

[19] As it is standard in the Learning literature, equation (56) 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 et al. (2017) for a discussion.

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 (12), the equilibrium PD ratio with imperfect information and learning reads as

$$\frac{P_t^L}{D_t} = \frac{\delta \bar{\beta}_t^{\ D}}{1 - (1 - \pi\tau)\delta\bar{\rho}\beta_t - \delta\pi\tau} \tag{20}$$

assuming that $1 > (1 - \pi\tau)\delta\bar{\rho}\beta_t + \delta\pi\tau$ hodls. 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.

Taxes, beliefs and price instability. Having characterized model-consistent beliefs, now the relationship between the CGT and the variance of the PD ratio can be revisited. While before beliefs were taken as given and the main object of the interest was the mapping from beliefs to prices, now both the effect of taxes on beliefs dynamics and the total of effect of taxes on price instability can be analyzed. The following propositions summarize the results.

Proposition 2.1 *Assume rationality and full information such that the equilibrium PD ratio is given by equation (10). Then, up to a first-order approximation, $\mathbb{V}(P_t^{RE}/D_t)$ is a decreasing function of τ because the sensitivity of prices to beliefs ω^{RE} is decreasing on τ. Σ^{RE} is unaffected by taxes.*

Proposition 2.2 *Assume rationality, imperfect information and Learning about prices such that the equilibrium PD ratio is given by equation (20). Then, up to a first-order approximation, $\mathbb{V}(P_t^L/D_t)$ is a decreasing function of τ because both ω^L and Σ^L are decreasing functions of τ.*

Proof. Appendix B.2.

The proof shows that under Rational Expectations, a first-order approximation delivers the following PD variance:

$$\mathbb{V}\Big(\frac{P_t^{RE}}{D_t}\Big) \approx (\omega^{RE})^2 \Sigma^{RE} \tag{21}$$

with

$$\omega^{RE} = \frac{\rho}{1-\rho} \sum_{j=1}^{\infty} f_0(\tau, j) f_1(d_t, j)(1 - \rho^j) \tag{22}$$

and

$$\Sigma^{RE} = \frac{\sigma_d^2}{1 - \rho^2} \tag{23}$$

The proof shows $\partial\omega^{RE}/\partial\tau < 0$. Thus, *Proposition 1* goes through: taxes reduce the transmission of belief fluctuations to price volatility. Additionally, now the fluctuations of expectations can be characterized. It turns out they are fully determined by the logarithmic growth rate d_t so that the covariance matrix of beliefs boils down to the variance of the dividend growth. Hence, Σ^{RE} is independent of τ. Altogether, taxes reduce the variance of the PD ratio

$$\frac{d\mathbb{V}(P_t^{RE}/D_t)}{d\tau} < 0 \tag{24}$$

by reducing ω^{RE} without affecting Σ^{RE}. In the case of i.i.d. growth, $\mathbb{V}(P_t^{RE2}/D_t) = 0$, and taxes cannot play any role.

When there is imperfect information and agents cope with it by learning from price data, taxes also influence the belief covariance matrix. Appendix B.2 shows that

$$\mathbb{V}\Big[\frac{P_t^L}{D_t}\Big] \approx \omega^L \Sigma^L (\omega^L)^T \tag{25}$$

In this expression, ω^L is given by

$$\begin{aligned}\omega^L = &\Bigg(\rho \frac{\delta exp\Big\{\mu + \sigma_d^2/2\Big\}}{1-(1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau}, \\ &\frac{\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}(1-\pi\tau)\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)}{\Big(1-(1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau\Big)^2}\Bigg)\end{aligned} \tag{26}$$

It is easy to check $\omega^L > 0$ and it decreases on τ whenever investors expect positive capital gains. Furthermore, the proof shows that

$$\Sigma^L = \begin{bmatrix} \mathbb{V}(\ln\beta_t) & \mathbb{C}ov(\ln\beta_t, d_t) \\ \mathbb{C}ov(\ln\beta_t, d_t) & \mathbb{V}(d_t) \end{bmatrix} = \underbrace{\frac{\sigma_d^2}{1-\rho^2}}_{=\Sigma^{RE}} \underbrace{\begin{bmatrix} \nu(\tau) & \upsilon(\tau) \\ \upsilon(\tau) & 1 \end{bmatrix}}_{\equiv M} \tag{27}$$

where $\nu(\tau)$ and $\upsilon(\tau)$ are constants defined in Appendix B.2., and the notation emphasizes their dependence on τ. Σ^L contains two elements: i) the fundamental risk captured by Σ^{RE}; ii) a matrix M showing how this risk is propagated through expectations fluctuations. Under RE, beliefs about future payoffs are determined by d_t; thus, fundamental risk is the only driver of beliefs volatility. Under Learning, though, d_t fully determines beliefs about short-term dividend growth but not beliefs about capital gains, which are also affected by their past values. In this case, fundamental risk is amplified through this price beliefs autoregression, giving rise to extra volatility.

It turns out that taxes influence the degree of fundamental risk amplification through subjective.

In particular,

$$\frac{\partial \mathbb{V}(\ln\beta_t)}{\partial \tau} = \frac{\partial \nu(\tau)}{\partial \tau}\Sigma^{RE} < 0 \tag{28}$$

that is, lower taxes amplify the effect of fundamental volatility in belief fluctuations and then, on price volatility. Furthermore, $\mathbb{C}ov(\ln\beta_t, d_t) > 0$ and

$$\frac{\partial \mathbb{C}ov(\ln\beta_t, d_t)}{\partial \tau} = \frac{\partial \upsilon(\tau)}{\partial \tau}\Sigma^{RE} < 0 \tag{29}$$

Thus, lower taxes increase the comovement between short-run fundamental growth and capital gains beliefs and, as result, asset prices react more to fundamental shocks.

Altogether, when agents learn from prices, taxes reduce the variance of the PD ratio

$$\frac{d\mathbb{V}(P_t^L/D_t)}{d\tau} < 0 \tag{30}$$

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. These beliefs amplify the fundamental risk, wtih 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 powerful regulator of price volatility.

2.4.- The lock-in effect

In this section, investors also 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 since the value of selling gets reduced. This is known as the lock-in effect. The literature has very much emphasized this effect, arguing that a CGT would boost -rather than dampen- price volatility (e.g., Somers (1948), Somers (1960), Stiglitz (1983)). Now, I include this effect and study the conditions under which it dominates the stabilizing effect analyzed in the previous section.

In this version, the liquidity shock z_t^i is replaced by a choice of the timing of realization. Following Gavin et al. (2007) and Gavin et al. (2015), each investor manages a stock of unrealized capital gains G_t facing portfolio management costs. G follows this law of motion

$$G_t = G_{t-1} + (P_t - P_{t-1})S_{t-1} - g_t + AC_t \tag{31}$$

with g_t are the realized capital gains, and AC_t stands for adjustment costs. It is assumed $AC_t = g_t - G_{t-1} - \phi(\bar{\pi}_t)G_{t-1}$ for $\bar{\pi}_t \equiv g_t/G_{t-1}$, 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 g_t capital gains and pay taxes $\tau^K g_t$ today or defer the realization and face an extra tax liability in the future.

Altogether, the investor's problem consists now in choosing sequences of consumption, stock holdings, stock purchases, realized and unrealized gains $\{C_t, S_t, X_t, g_t, G_t\}_{t=0}^{\infty}$ to maximize lifetime welfare (1), subject to the stock bounds, the stocks law of motion (3), the unrealized capital gains law of motion (31) and the budget constraint

$$C_t + P_t X_t \leq D_t S_{t-1} - \tau g_t + T_t \tag{32}$$

Let $\phi(\bar{\pi}_t) = (\tau)^{1+\xi} ln(\bar{\pi}_t)$. It 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, $\lim_{g_t \to 0} -\phi(\bar{\pi}_t) = +\infty$.

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

$$P_t = \delta \mathbb{E}_t^{\mathcal{P}} \Big[D_{t+1} + P_{t+1} - \mu_{t+1}(P_{t+1} - P_t) \Big] \tag{33}$$

$$\bar{\pi}_t = \mu_t \tau^{\xi} \tag{34}$$

$$\mu_t = \tau^{1+\xi} \mathbb{E}_t^{\mathcal{P}} \Big[\delta \mu_{t+1} \Big(1 - ln(\bar{\pi}_{t+1}) \Big) \Big] \tag{35}$$

Equation (33) 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 (31). 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 (34), the optimal realization of capital fraction $\bar{\pi}_t$ 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^K. If $\xi < 0$, a higher τ^K lowers the optimal realization, expressing the lock-in effect. From this equation is clear that ξ is the tax elasticity of realization

$$\frac{\partial \bar{\pi}_t}{\partial \tau}\frac{\tau}{\bar{\pi}_t} = \mu_t \xi (\tau)^{\xi-1}\frac{\tau}{\mu_t \tau^{\xi}} = \xi \tag{36}$$

which is a structural parameter of the model that can be directly estimated from the data. Finally, equation (35) is a first-order stochastic difference equation determining the shadow price of unrealized capital gains. Shifting it one-period ahead:

$$\mu_{t+1} = \tau^{1+\xi}\mathbb{E}^{\mathcal{P}}_{t+1}\left[\delta\mu_{t+2}\Big(1 - ln(\mu_{t+2}\tau^{\xi})\Big)\right] = \tau^{1+\xi}M_{t+1} \tag{37}$$

Then, operating on the Euler Equation (33)

$$P_t = \delta(1-\tau^D)\mathbb{E}^{\mathcal{P}}_t\left[\frac{D_{t+1}}{D_t}\right]D_t + \delta\mathbb{E}^{\mathcal{P}}_t\left[\frac{P_{t+1}}{P_t}(1-\tau^{1+\xi}M_{t+1})\right]P_t + \delta\tau^{1+\xi}P_t\mathbb{E}^{\mathcal{P}}_t\Big[M_{t+1}\Big] \tag{38}$$

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

$$\mathbb{E}^{\mathcal{P}}_t\left[\frac{P_{t+1}}{P_t}M_{t+1}\right] = \mathbb{E}^{\mathcal{P}}_t\Big[(\bar{\rho}\beta_t + u_t + \vartheta_{t+1} + \varepsilon^P_{t+1})M_{t+1}\Big] \approx \bar{\rho}\beta_t m_t \tag{39}$$

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

$$\frac{P_t}{D_t} = \frac{\delta\bar{\beta}_t^{\,D}}{1-(1-\tau^{1+\xi}m_t)\delta\bar{\rho}\beta_t - \delta\tau^{1+\xi}m_t} \tag{40}$$

This formula embeds the one under exogenous realization (20) 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 $\xi < -1$, that is

$$\xi \lesseqgtr -1 \;\;\Rightarrow\;\; \frac{\partial P_t/D_t}{\partial\tau} \gtreqless 0 \tag{41}$$

However, for equation (40) to be an equilibrium equation, m_t must be characterized. Equation (35) does not have an analytical solution but can be computed numerically.[20] 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 1 plots the propagation of a one-off dividend shock via capital gains beliefs under Learn-

[20]The algorithm to do so combines the Parameterized Expectations Algorithm with numerical integration. See Appendix F.

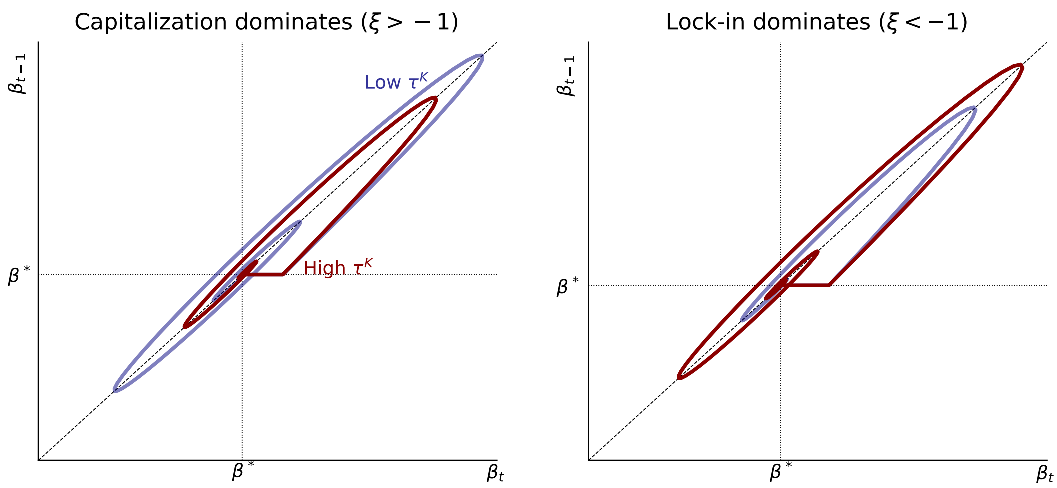

Figure 1: ***Shock propagation at different tax levels with both the capitalization and the lock-in effect.*** *The graph illustrates the dynamics of subjective capital gains expectations (β) when shocked by a one-off dividend shock. The blue (red) line uses $\tau = 0.1(= 0.4)$. The LHS (RHS) graph uses $\xi = -0.5$ ($\xi = -1.2$).*

ing. Note that the response of a capital gains beliefs to a fundamental shock under RE would be none. Hence, the graphs shows that 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 dampens the propagation, stabilizing beliefs and prices. Altogether, the model is able to deliver the stabilizing logic outlined before as well as the destabilizing effect from the lock-in effect as emphasized by Stiglitz (1983) depending on the value of ξ.[21]

3.- An application to the US Stock Market

This section uses the theory to link tax cuts to the proliferation of fluctuations in the US stock market since the 1980s. The emergence of large booms and busts, such as the Dotcom bubble, in a macroeconomic context characterized by the Great Moderation is a troubling observation for many macrofinance models. Following lower consumption growth volatility, models based on macroeconomic fundamentals would predict more stable prices due to a less volatile stochastic discount factor (e.g., Campbell and Cochrane (1999)). Besides, theories that explicitly detach prices from fundamentals, as some models of learning, would predict lower belief fluctuations and

[21] 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 (41) is approximately correct.

then more stable prices driven by smaller forecast errors (e.g., Adam et al. (2016)). Even models that link lower macroeconomic risk with higher demand for risky assets can explain part of the stock prices' run-up but not much of their larger swings (e.g., Lettau et al. (2008)).

In this section, I argue that capital gains tax cuts fed the larger fluctuations, partly counteracting the stabilizing force coming from the macroeconomy. First, in Section 3.1., I documented a list of asset pricing facts that highlight significant changes after the 1980s. Section 3.2. sets up a quantitative model that extends the one presented in Section 2, introduces a novel application of the Parameterized Expectations Algorithm to solve it and discusses the parameterization, which includes including a structural estimation using the Simulated Method of Moments. Section 3.3. presents the estimation results, a decomposition of different channels in the model, and an estimation of the average effect of tax cuts. Section 3.4. explores the implications of the model for the informational efficiency of the stock market. Finally, Section 3.5. discusses the ability of the model to produce a high enough equity premium.

3.1.- Facts

This section documents asset pricing facts using US data from 1946 to 2018. Observations are split into two halves of 36 years each for two reasons: i) to highlight relevant changes that have occurred in recent decades; ii) to measure changes in volatility more reliably, since fluctuations in valuation ratios are large and persistent, making difficult to measure volatility changes at short horizons. I report a list of standard asset pricing statistics involving volatility, predictability, valuations and the equity premium. Additionally, I include statistics reflecting investors' expectations. The evidence is summarized in the following five facts, which are collected in Table 1.

Fact 1: Lack of a Great Moderation in the stock market.[22] The fluctuations of the PD ratio, captured by its second moment, increased substantially after the 1980s. The Campbell and Shiller (1988)'s "dynamic accounting equation" has become a classical framework to understand the most proximate drivers of this variance. The equation reads as

$$\mathbb{V}(\ln P_t/D_t) \approx \quad \mathbb{C}ov\Big(\ln P_t/D_t, \underbrace{\sum_{j=1}^{\infty} \hat{\rho}^{j-1} \Delta \ln D_{t+j}}_{\equiv \bar{d}_t}\Big) - \mathbb{C}ov\Big(\ln P_t/D_t, \underbrace{\sum_{j=1}^{\infty} \hat{\rho}^{j-1} \ln R_{t+j}}_{\equiv \bar{r}_t}\Big) \qquad (42)$$

[22] In the paper, I focus on the unconditional variance. However, an increase in volatility is also observed for conditional variance. See Appendix H.

where $\hat{\rho} = PD/(1 + PD)$, PD is the mean PD ratio in the sample and $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$ is the gross stock return. The approximate accounting relationship points out that variations in the PD ratio must be associated with fluctuations in either future dividend growth or future returns.

The results of this decomposition, computed using a VAR projection as proposed by Campbell and Shiller (1988), are shown in the first panel of Table 1. The figures for the 1946-1982 subsample point out two results widely shared by the literature (see, for instance, Cochrane (2009)). First, variations in the PD ratio are only weakly related to future dividend dynamics (-10%); instead, they are mostly associated with variations in returns. Second, the association between prices and future dividend growth is negative. The two results are puzzling for any theory that emphasizes future dividends as the main force driving current stock prices.

However, for the 1982-2018 subperiod, the story is substantially different. Dividend growth accounts for a larger share of the PD variance (17%) and the correlation with prices becomes positive, more aligned with standard theories. Put differently, returns and dividend growth moved together before the 1980s but decoupled afterward. This is quantitatively important: the reversal accounts for 47% of the increase in the variance of the PD ratio. The other half is related to the larger covariance between prices and future returns.

Additionally, the variance of stock returns remained largely stable during this period. This absence of a decrease in stock market volatility stands in sharp opposition to the broader economic trend. Table 1 also includes the standard deviations of both dividend and consumption growth rates, which experienced a 20% and 50% reduction, respectively.

Fact 2: A fall in excess return predictability. Price-Dividend volatility is closely related to stock return predictability. As it is well known, the covariance between the PD ratio and the discounted sum of future returns, which appears in the right-hand side of expression 42, would also be the denominator of the coefficient of an OLS projection of the log PD ratio into the accumulated discounted returns (see, for instance, Cochrane (2009)). And, since accumulated excess returns are mostly driven by accumulated stock returns, that covariance should be similar to the denominator of a predictability regression like

$$xr_{t,t+h} = c_0 + c_1^h \frac{P_t}{D_t} + u_t^{xr} \tag{43}$$

with

$$xr_{t,t+h} = \prod_{j=1}^{h} R_{t+j} - \prod_{j=1}^{h} R_{t+j}^b \tag{44}$$

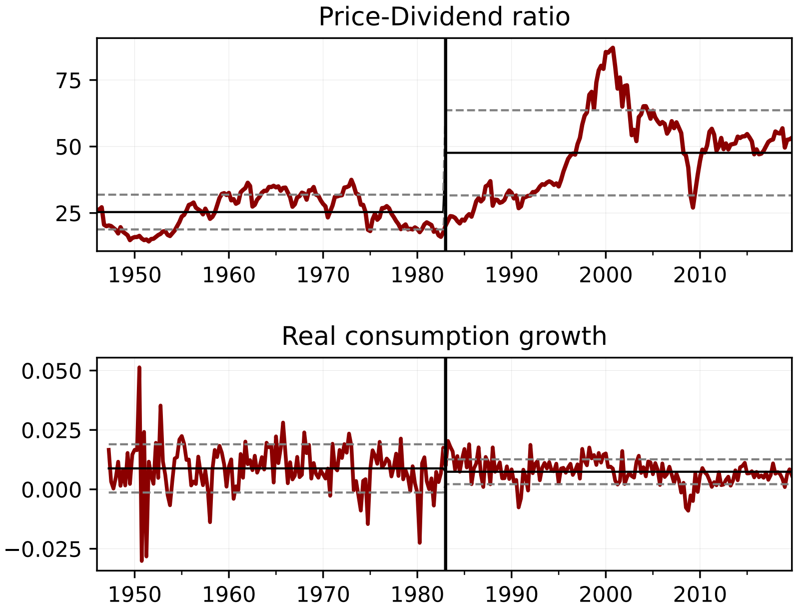

Figure 2: ***No signs of a Great Moderation in the stock market.*** *The graph plots the evolution of the PD ratio and real aggregate consumption growth in the 1946:I-2018:IV period. The continuous lines plot the mean of each subperiod 1946:I-1982:II and 1982:III - 2018:IV. The dotted bands show ± one standard deviation.*

where R_t^b is the gross real return on 90 days T-bills. I run the regression for $h = 4, 12, 20$ quarters for the two sub-samples. The result is always the same: a significant reduction in the magnitude of the coefficient c_1^h. In Table 1, I include the results for $h = 20$. At this horizon, the coefficient drops from -0.17 pre-1980s to just -0.04 post-1980s, that is, a decline of more than 75% in predictability (the drop in predictability is close to 70% for $h = 12$ and 65% for $h = 4$).

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 1, this is not the case; persistence is statistically the same across the two subsamples. Instead, the Campbell-Shiller's decomposition suggests that the predictability drop is related to the fact that the increase in $\mathbb{V}(\ln P_t/D_t)$ has been greater than the increase in $\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$ due to the substantial rise in $\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$.

Fact 3: A surge in capital gains. The mean PD ratio almost doubled after the 1980s. This fact has been extensively documented in the literature (e.g., Shiller (2000), McGrattan and Prescott (2005), Brun and González (2017)) and is illustrated in figure 2. In accounting terms,

$$\frac{P_{t+n}/D_{t+n}}{P_t/D_t} = exp\Big\{\Delta \ln P_{t+n} - \Delta \ln D_{t+n}\Big\} \tag{45}$$

Hence, the accounting reason for a higher PD ratio is that the increase in price growth exceeds the dividend growth. Indeed, table 1 shows that the average price growth was 1.48% post-1982 from 0.48% pre-1982 while dividends grew only 0.75% post- from 0.49% pre-1982.

Fact 4: A stable equity premium. The historical average of the difference between the real stock return and the real 90 days T-Bill was statistically the same in both samples. 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. Then, sample averages are very much the same. The same is true for stock returns.

Fact 5: A higher sensitivity of prices to beliefs and shocks. A key insight of the model was that low taxes would increase volatility by increasing the sensitivity of prices to beliefs. This sensitivity can be measured using survey data. Since finding investors' survey data going back to 1946 is problematic, I follow Adam et al. (2017) to extend the UBS Gallup survey for the 1946-2018 period using the following adapting algorithm:

$$\ln\mathbb{E}_t^s\left(\frac{P_{t+1}}{P_t}\right) = (1-g)\ln\mathbb{E}_{t-1}^s\left(\frac{P_t}{P_{t-1}}\right) + g\,\ln\frac{P_{t-1}}{P_{t-2}} \tag{46}$$

where $\mathbb{E}^s$ stands for survey expectations, P_t is the real SP500 price, 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 $g = 0.0266$ and the initial condition is set equal to zero growth. This equation, which is consistent with a Kalman filter, turns out to replicate the survey data points very well, exhibiting a correlation of 0.85 in the period that both series overlap. The approach of obtaining survey-like data by projecting the existing surveys on other variables has also been used in Nagel and Xu (2022) . Besides, a constant gain scheme has proven to be a parsimonious and reasonable way of modeling other expectations series (see, for instance, Malmendier and Nagel (2016)).

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

$$PD_t = \alpha + \zeta\ln\mathbb{E}_t^s\left(\frac{P_{t+1}}{P_t}\right) + \varepsilon_t^s \tag{47}$$

As pointed out by Adam et al. (2017) and illustrated in figure 3, ζ went up substantially after the 1980s. Thus, although the correlation between the PD ratio and beliefs was virtually the same across both periods, their relationship became much steeper.

Additionally, the model emphasized that the sensitivity of the PD ratio to price growth shocks

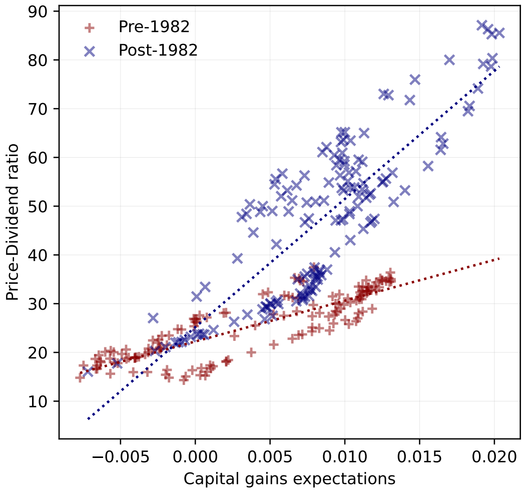

Figure 3: ***The PD ratio became more reactive to capital gains survey-like expectations.*** *The graph plots the quarterly PD ratio of the SP500 against the log real mean price growth expectations implied by an extended version of the UBS Gallup survey. The periods are 1946:I–1982:II and 1982:III–2018:IV.*

would be more pronounced in environments with lower taxes. A standard way of computing this sensitivity is by measuring the response of the PD ratio at the horizon h defined as

$$\mathbb{E}(PD_{t+h}|\varepsilon_t^P = 1, I_t) - \mathbb{E}(PD_{t+h}|\varepsilon_t^P = 0, I_t) \tag{48}$$

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 PD ratio to have a description of the data. As shown in figure 4, the response is significantly larger in the second subsample. For instance, at the 1-year horizon, the point estimate for the response is 70% higher in the post-1982 sample.

The previous changes in the stock market have occurred in a period where many paramount institutional changes were put in place. I highlight here two of the ones more directed to investors' payoffs.

A decline in personal capital taxes. Taxes on personal investment income (dividends, capital gains, and interests) decreased substantially over the last decades (McGrattan and Prescott (2005), Sialm (2009)). A standard measure is the effective average marginal rate, that is, a value-weighted mean of the marginal tax rates of investors in the various income brackets once adjusted for the features of the tax code (such as maximum and minimum taxes, partial inclusion of social security, or phaseouts of the standard deduction).[23] The NBER's TAXSIM program computes these

[23] 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.

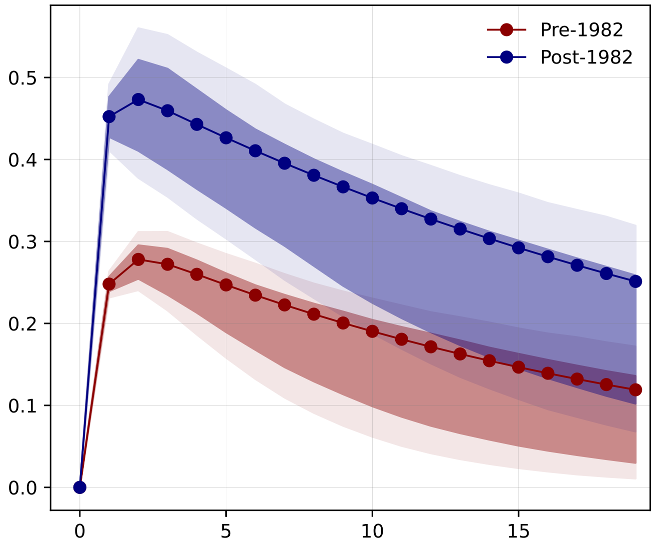

Figure 4: ***Response of the PD ratio to an equivalent shock.*** *The graph plots the response function of the PD ratio to a price growth shock at a quarterly frequency. Confidence bands at 68% and 95% levels are shown, coming from a bootstrap procedure with 1000 repetitions.*

rates. It does not account, though, for regulatory changes involving pension savings vehicles that led to a massive change in asset holdings from taxable to nontaxable accounts, effectively reducing taxes on the aggregate investment income (see McGrattan and Prescott (2005)). To correct for that, I follow the literature and adjust the NBER rates as follows:

$$\tau_t^D = \tau_t^d(1-\eta_t) \tag{49}$$

$$\tau_t = (\phi\tau_t^{skg} + (1-\phi)\tau_t^{lkg})(1-\eta_t) \tag{50}$$

$$\tau_t^B = \tau_t^b(1-\eta_t) \tag{51}$$

In the previous expressions, τ_t^d, τ_t^{skg}, τ_t^{lkg} and τ_t^b are the NBER effective average marginal rates on dividends, short, long capital gains and interest income, respectively; ϕ is the average weight of short capital gains on total capital gains; η_t is the nontaxable share. Data sources are in Appendix A; computation details on the nontaxable share are in Appendix C. As illustrated in figure 5, taxes exhibited a substantial decline which represented a movement towards a generally lower tax environment. This overall tax decline was the result of the joint action of statutory tax reforms along with changes in regulations that drove a movement to nontaxable accounts.[24] According

[24]Important reforms were the reduction of capital gains by Carter in 1978 and Clinton in 1997, partially counter-

Table 1: ***Five changes in the US Stock Market: 1946-1982 versus 1982-2018.*** *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 using 9 lags. The residuals for the response of the PD ratio to a price growth shock are obtained via bootstrapping with 1000 repetitions.*

		1946-1982	1982-2018
	Fact 1: Lack of a Great Moderation		
Volatility of the PD ratio	$\mathbb{V}ar(\ln P_t/D_t)$	7.15	13.97
		(1.35)	(3.56)
Comovement PD - dividends	$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	-0.81	2.38
		(0.20)	(0.55)
Comovement PD - returns	$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	-7.96	-11.60
		(1.54)	(3.03)
Stock returns volatility	$\sigma(r_t^s)$	7.87	7.41
		(0.72)	(0.81)
Volatility of dividend growth	$\sigma(D_t/D_{t-1}-1)$	2.52	1.97
		(0.41)	(0.35)
Volatility of consumption growth	$\sigma(C_t/C_{t-1}-1)$	1.00	0.53
		(0.15)	(0.06)
	Fact 2: A fall in excess return predictability		
5y Regression coefficient	c^{20}	-0.17	-0.04
		(0.02)	(0.01)
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	$\mathbb{E}(PD_t)$	25.48	47.09
		(1.55)	(4.04)
Dividend growth	$\mathbb{E}(D_t/D_{t-1}-1)$	0.49	0.75
		(0.35)	(0.34)
Stock price growth	$\mathbb{E}(P_t/P_{t-1}-1)$	0.48	1.84
		(0.64)	(0.63)
	Fact 4: A stable equity premium		
Real bond returns	$\mathbb{E}(r_t^b)$	0.42	0.38
		(0.02)	(0.03)
Real stock returns	$\mathbb{E}(r_t^s)$	4.73	4.34
		(0.76)	(0.76)
	Fact 5: A higher sensitivity of prices to beliefs and shocks		
Correlation of prices and beliefs	$corr(PD_t, \ln\beta_t)$	0.84	0.84
		(0.06)	(0.07)
Sensitivity of prices to beliefs	ζ	0.84	2.63
		(0.08)	(0.26)
Response of the PD ratio to a shock	$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P=1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P=0, I_t)$	0.26	0.44
		(0.06)	(0.10)

to my estimates and in line with the literature (McGrattan and Prescott (2005), Sialm (2009), Rosenthal and Austin (2016)), the share of equity income paying taxes dropped from 87% in 1946 to just 30% in 2018.

The rise in stock repurchases. In 1982, the Securities and Exchange Commission's Rule

acted by Reagan in 1986. When it comes to dividends, Reagan 1982 and Bush 2001 and 2003 represented substantial tax cuts.

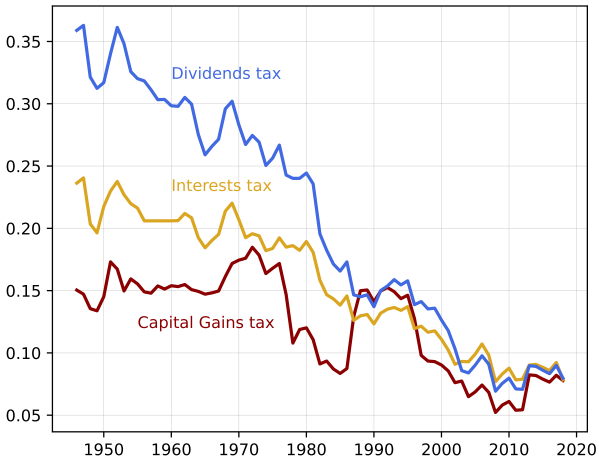

Figure 5: ***Capital tax rates during the postwar period****. The graph plots the effective average marginal tax rates on dividends (blue), interest income (yellow), and capital gains (red) as defined by equation (49),(50), and (51). Annual series 1946-2018. See Appendix A for data sources and Appendix C for details on the computations.*

10B-18 provided a safe harbor for stock repurchases under the Rule´s conditions.[25] With this change, stock buybacks began to grow substantially, as figure 6 illustrates. In fact, they became the dominant form of corporate payout for the SP500 firms since 1997. As opposed to dividends, earnings growth volatility did not fall. Hence, there is a possibility that part of the extra price-dividend volatility simply reflects a fundamental volatility that is channeled through repurchases rather than dividends.

3.2.- A Quantitative Model

This section extends the model outlined in Section 2.4. to remove certain simplifying assumptions. It also introduces a new application of the Parameterized Expectations Algorithm to solve it. Finally, I discuss the parameterization, which involves observable parameters picked from the US data and a structural estimation of the rest via the Simulated Method of Moments.

The model is extended in five dimensions. First, the assumption of risk neutrality is abandoned. Instead, investors are allowed to dislike risk in a Constant Relative Risk Aversion (CRRA) sense, with a parameter γ regulating its level of risk aversion.

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 et al.

[25] See `this file` for a summary.

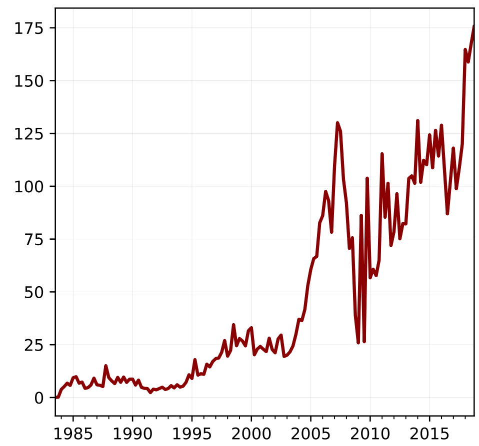

*Figure 6: **Stock repurchases by SP500 firms**. The graph plots the quarterly repurchases performed by all SP500 1983.III-2018.IV in billions of dollars.*

(2017), it is assumed agents get a wage endowment W_t each period, following this process:

$$\ln\Big(1+\frac{W_t}{D_t}\Big) = (1-p)\ln(1+w) + p\ln\Big(1+\frac{W_{t-1}}{D_{t-1}}\Big) + \varepsilon_t^w \tag{52}$$

D_t are aggregate dividends, $1+w$ is the average wage-dividend ratio and $p \in [0,1)$ is its quarterly persistence. The innovations to the log dividends growth (given by equation (8)) and wage-dividends are jointly distributed as follows

$$\begin{pmatrix} \varepsilon_t^d \\ \varepsilon_t^w \end{pmatrix} \sim ii\mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_D^2 & \sigma_{DW} \\ \sigma_{DW} & \sigma_W^2 \end{pmatrix} \right)$$

Third, taxes are stochastic. In particular, each type of tax follows a unit root process, that is

$$\tau_t^j = \tau_{t-1}^j + \epsilon_t^{\tau^j} \tag{53}$$

where $\epsilon_t^{\tau_j} \sim ii\mathcal{N}(0, s_{\tau_j}^2)$.[26] Tax shocks are assumed to be orthogonal to dividend and consumption shocks.

Fourth, risk-free bonds with a gross rate return R_t^b are introduced such that investors can choose between stock and bonds to store value. Fifth, the informational assumptions about dividends are relaxed such that investors are now uncertain about the dividend process. In their mind, the process

[26]When the observed tax time series are fit into an AR(1) model, the estimated coefficients are not statistically different from 0 (intercept) and 1 (slope). Thus, the unit root process constitutes a realistic representation of the tax process. Moreover, their residuals behave as Gaussian white noise. Normality has been tested via the Shapiro-Wilk Normality test.

looks like

$$\ln\frac{D_t}{D_{t-1}} = \ln\tilde{d}_t + \varepsilon^d_t \tag{54}$$

$$\ln\tilde{d}_t = \ln\tilde{d}_{t-1} + \nu^{\tilde{d}}_t \tag{55}$$

where $\tilde{d}_t$ is unobserved and innovations are Normal i.i.d. with constant variances $\sigma^2_D$and $\sigma^2_{\tilde{d}}$ 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 $\ln\tilde{d}_t|\omega^t \sim \mathcal{N}(\ln\beta^d_t, \sigma^2)$ with

$$\ln\beta^d_t = \ln\beta^d_{t-1} + g^d\Big(\ln\frac{D_t}{D_{t-1}} - \ln\beta^d_{t-1}\Big) \tag{56}$$

It follows that

$$\mathbb{E}^{\mathcal{P}}_t\Big[\frac{D_{t+1}}{D_t}\Big] = \beta^d_t exp\{\sigma^2_D/2\} \tag{57}$$

Finally, buybacks are included in the form of negative supply shocks

$$\bar{S}_t = 1 - \varepsilon^s_t \tag{58}$$

where $\bar{S}_t$ is the aggregate stock supply available in the market.

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, I conjecture that the consumption policy is linear in wealth, with the propensity to consume depending on after-tax return expectations.[27] A stock demand function S^d_t can be then derived from the budget constraint and equilibrium prices can be obtained using the equity market clearing condition:

$$S^d_t = S^d\Big(\frac{P_t}{D_t}, \cdot\Big) = \bar{S}_t \tag{59}$$

where $\bar{S}$ is the aggregate supply in the market. Due to the handy policy function, this equation

[27]See, for instance, Hakansson (1970).

can be analytically solved for prices such that

$$\frac{P_t}{D_t} = \frac{\chi_1 \delta (1-\tau_t^D)\beta_t^d exp\{\sigma_d^2/2\}}{\bar{S}_t - \chi_1 \delta (1-\psi_2(\boldsymbol{X}_t;\chi_2)\tau_t^{1+\xi})\beta_t} \left(1 + \frac{W_t}{D_t}\right) \tag{60}$$

where χ_1 is a parameter of the consumption approximating function and $\psi_2(\boldsymbol{X}_t;\chi_2)$ is approximating the expected value of the Lagrange multiplier in the unrealized capital gains accumulation equation.[28] Despite containing very few parameters, both approximating functions perform well, with errors in terms of consumption equivalent to \$1 out of \$1,000 and \$1 out of \$1,000,000, respectively. See Appendix D for all the details about the PEA implementation and accuracy measures.

Parameterization. The parameterization strategy is twofold. A subset of parameters related to income processes, the vector $\widetilde{\theta} = \{\sigma_D, \sigma_W, \sigma_{WD}, p\}$, is picked directly from US data, distinguishing between the two subsamples. Parameter values are specified in panel a) of Table 2 and data sources are reported in Appendix A.

The remaining parameters are collected in the vector $\theta = \{\delta, g, \gamma, \xi, w, PD^L, PD^U, B_0\}$, where PD^L and PD^U are parameters of the projection facility.[29] This vector is estimated via an extension of the Simulated Method of Moments, following Adam et al. (2016). To test the power of taxes to explain the various observed changes, estimated parameters are kept fixed throughout the sample. Hence, a total of n=8 parameters are estimated to match to a subset of M moments from those reported in table 1. The vector θ is chosen to minimize the distance between the model $\widetilde{\mathcal{S}}(\theta)$ and the data $\hat{\mathcal{S}}$ statistics, that is,

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[\hat{\mathcal{S}} - \widetilde{\mathcal{S}}(\theta)\right]' \hat{\Sigma}_{\mathcal{S}}^{-1} \left[\hat{\mathcal{S}} - \widetilde{\mathcal{S}}(\theta)\right] \tag{61}$$

where $\hat{\mathcal{S}}$ an $\widetilde{\mathcal{S}}(\theta)$ are $M \times 1$ vectors and $\hat{\Sigma}_{\mathcal{S}}$ is a $M \times M$ weighting matrix, which determines the relative importance of each statistic deviation from its target. A diagonal weighting matrix is used, whose diagonal is composed of the inverse of the estimated variances of the data statistics. The model is fed with the observed dividend growth, tax, and buybacks shocks; wage shocks are simulated through a Monte Carlo experiment with 1000 realizations. I formally test the hypothesis that any of the individual model statistics differs from its empirical counterpart using t-statistics.

[28] As it is common in similar models, there is a discontinuity in the formula. 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 et al. (2016)). Appendix D contains the details.

[29] The projection facility starts to dampen belief coefficients that imply a PD ratio equal to PD^L and sets an effective upper bound at PD^U. See Appendix D.

Table 2: ***Parameters.*** *Panel a) table reports the values of the parameters coming from US data. Panel b) reports the estimated parameters from the Simulated Method of Moments procedure described in the text.*

a) Parameters from the data		1946-1982	1982-2018
Dividend growth standard deviation	σ_D	0.0252	0.0197
Wage-dividend ratio standard deviation	σ_W	0.0261	0.0196
Covariance (wage-dividend, dividend growth)	σ_{WD}	-0.0006	-0.0004
Persistence wage-dividend ratio	p	0.91	0.99
b) Estimated parameters from SMM			
Discount factor	δ	1.00	
Kalman gain	g	0.0260	
Risk aversion	γ	1.24	
Tax elasticity of realization	ξ	-0.30	
Average wage-dividend ratio	w	5.83	
Starting value of the projection facility	PD^L	731.20	
Upper bound of the projection facility	PD^U	715.82	
Initial buybacks shocks	B_0	0.0023	

3.3.- Estimation results

In this section, the estimation results are reported. I split the list of statistics in the Table 1 in two groups. The first group, M_0, is included in the SMM procedure; the excluded statistics are used as out-of-estimation tests of the model. The following statistics are included in M_0 :

$$M_0 = \left\{ \mathbb{E}(PD_t), \mathbb{E}(P_t/P_{t-1}), \mathbb{E}(r_t^s), \mathbb{V}ar(\ln P_t/D_t), \mathbb{C}ov(\ln P_t/D_t, \bar{d}_t), \mathbb{C}ov(\ln P_t/D_t, \bar{r}_t), \sigma(r_t^s), \zeta \right\} \tag{62}$$

It includes key statistics related to the average price level, volatility and the sensitivity of prices to beliefs.[30] Since all of them are included in each sample, that makes a total of 16 statistics to be matched with 7 parameters.

The estimated parameter vector $\hat{\theta}$ is reported in panel b) of table 2. In the analysis, δ is constrained to fall within the interval of 0 and 1, with the estimation indicating that it approaches the upper limit of this range. The Kalman gain g is calculated to be 0.0260. This figure not only aligns very closely with the 0.2660 estimate derived from survey data but also concurs with various other estimations reported in the literature, such as those presented by Adam et al. (2017). Furthermore, the estimated level of risk aversion stands at 1.24, a figure that is considered to be

[30]Expanding the set of statistics does not change the results much. The only difference is the inclusion of the risk-free rate, which pushes the model towards a very low risk-aversion parameter and a too high volatility of beliefs to compensate for the former.

within the bounds of what is generally deemed reasonable. Perhaps the most critical parameter determined in our study is the tax elasticity of realization, $\xi = -0.3$. This parameter suggests that although the lock-in effect is observable, it is overshadowed by the tendency of lower taxes to increase market volatility. The identification is driven by the constraints imposed by the variance decomposition. Thus, the higher ξ, the higher is $\mathbb{V}ar(\ln P_t/D_t)$; however, it also generates an excessive increase in $\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$ (and a strong increase in $\sigma(r_t^s)$ that is completely at odds with the data). Notably, $\xi = -0.3$ is within the interval provided by Zidar (2019), who leveraged state-level variations for his calculations.

Table 3 presents the baseline 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-tests, with t-stats generally below 2, indicating a robust fit to the data. This analysis reveals a significant increase in the mean Price-to-Dividend (PD) ratio, primarily driven by an increase in capital gains. Both the mean level of stock returns and its virtual constancy are well replicated. Furthermore, the model captures an important increase in the variance of the PD ratio as well as the variance decomposition. This includes a transition from negative to positive correlation between the PD ratio and $\bar{d}_t$, as well as a more pronounced negative covariance between the PD ratio and $\bar{r}_t$. Additionally, the model successfully approximates the level of stock return volatility, where, despite a mild increase, it remains within the confidence bands. Another significant finding is the marked rise in ζ, the sensitivity of the PD ratio to beliefs, corroborating the theoretical mechanisms discussed in Section 2.

Beyond the statistics incorporated into the SMM procedure, the model exhibits several noteworthy features. It predicts a considerable reduction in the predictability of 5-year excess stock returns, maintaining a high and relatively stable serial correlation of the PD ratio. Additionally, capital gains expectations are strongly pro-cyclical, consistent with survey data. Moreover, the model demonstrates an enhanced response of the PD ratio to price growth shocks, indicating a movement in the intended direction, albeit with an excessive magnitude. These findings underscore the model's quantitative strengths, which not only adhere to the observed dynamics within financial markets but also perfectly replicate the observed decline in macroeconomic volatility, as evidenced by the standard deviations of dividends and consumption growth.

Decomposition by channels. To gain deeper insights into the specific contributions of each channel within our model, Table 4 provides statistics derived from scenarios where individual channels are selectively deactivated. Due to spatial constraints, this table is focused on presenting only the mean Price-Dividend (PD) ratio and the breakdown of variance, which are the most critical

Table 3: ***Baseline estimation results.*** *This table reports the M statistics included in the SMM estimation. The first four columns report the observed statistics for the US data. The next four columns report model-implied statistics and their t-statistics using the Table 2's parameters. The rates of growth, returns, variances and covariances have been multiplied by 100.*

	US data		Baseline 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 SMM estimation					
$\mathbb{E}(PD_t)$	25.48	47.09	27.31	-1.18	39.73	1.82
$\mathbb{E}(P_t/P_{t-1}-1)$	0.48	1.84	0.82	-0.53	1.74	0.16
$\mathbb{E}(r_t^s)$	4.73	4.34	4.78	-0.07	4.66	-0.42
$\mathbb{V}ar(\ln P_t/D_t)$	7.15	13.98	6.48	0.49	11.68	0.64
$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	-0.81	2.38	-0.30	-2.62	2.87	-0.88
$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	-7.96	-11.60	-6.76	-0.78	-8.83	-0.92
$\sigma(r_t^s)$	7.87	7.41	7.16	0.98	8.59	-1.45
ζ	0.84	2.63	0.87	-0.41	2.20	1.60
	Non-included in SMM estimation					
c^{20}	-0.17	-0.04	-0.13	-1.91	-0.03	-2.10
$corr(PD_t, PD_{t-1})$	0.96	0.98	0.96	-0.59	0.94	-0.16
$corr(PD_t, \beta_t)$	0.84	0.84	0.88	-0.81	0.87	-0.44
$\mathbb{E}(PD_{t+4}\|\varepsilon_t^P=1, I_t) - \mathbb{E}(PD_{t+4}\|\varepsilon_t^P=0, I_t)$	0.26	0.44	0.64	-6.40	1.36	-9.15
$\mathbb{E}(D_t/D_{t-1}-1)$	0.49	0.75	0.49	0.00	0.75	0.00
$\sigma(D_t/D_{t-1}-1)$	2.52	1.97	2.52	0.00	1.97	0.00
$\sigma(C_t/C_{t-1}-1)$	1.00	0.53	1.00	0.00	0.53	0.00

metrics for our analysis.

In the absence of a lock-in effect (i.e., $\xi = 0$ and $m_t = \pi$), lower taxes boost optimism without increasing the realization of gains so tax cuts unequivocally encourage higher and more fluctuating prices.[31] Echoing Stiglitz (1983), the lock-in effect generates lower volatility from tax cuts. Similarly, without transfers ($T_t = 0$), investors' income fluctuates more due to the taxes paid on realized capital gains.This extra volatility gets transmitted into greater price fluctuations. On the contrary, the absence of buybacks ($\varepsilon_t^s = 0, \forall t$) leads to fewer shocks and lower volatility. Importantly, buybacks appear to be critical to changing $\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$ from negative to positive. Furthermore, ignoring the Great Moderation by maintaining the volatility of aggregate consumption growth at its 1946–1982 level throughout the second subsample leads to higher volatility, highlighting the stabilizing role of the Great Moderation.

[31] A calibration of π is implemented to align $\mathbb{E}(PD_t)$ in the initial subsample to the baseline scenario, ensuring comparability.

A usual benchmark is Rational Expectations. In this case, investors' expectations about capital gains are time-varying as follows from their knowledge about dividend growth being an AR(1) process, but the expectations-price spiral is absent. The statistics reported in table 4 show that even without Learning, the model produces a substantial increase in both the mean and the variance of the PD ratio. It is worth mentioning, though, that the RE version performs poorly in different dimensions. The biggest failure comes from a too low PD ratio persistence, 0.17 and 0.32 respectively, as their fluctuations are driven by the dividend growth rate whose persistence is low. Table xx in appendix X reports all the statistics for the model with RE. Thus, learning from prices appears as pivotal for generating a change in volatility along with a highly persistent PD ratio.

Table 4: ***Model statistics in different versions of the model.*** *This table reports selected statistics for different versions of the model. The "full-fledged model" reproduces the results in Table 3. All versions use table 2's parameters. Variances and covariances have been multiplied by 100.*

	Full-fledged model		No Lock-in Effect		No transfers	
	1946-1982	1982-2018	1946-1982	1982-2018	1946-1982	1982-2018
$\mathbb{E}(PD_t)$	27.31	39.73	27.57	43.26	26.13	39.66
$\mathbb{V}ar(\ln P_t/D_t)$	6.48	11.68	5.86	14.85	5.95	12.40
$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	-0.30	2.87	-0.04	3.17	-0.29	1.05
$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	-6.76	-8.83	-5.87	-11.51	-6.23	-11.14
	No buybacks		**No learning (RE)**		**No Great Moderation**	
	1946-1982	1982-2018	1946-1982	1982-2018	1946-1982	1982-2018
$\mathbb{E}(PD_t)$	27.37	37.40	27.09	38.96	27.40	40.20
$\mathbb{V}ar(\ln P_t/D_t)$	7.23	9.19	8.15	12.43	6.43	12.01
$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	-0.61	-1.07	0.42	0.88	-0.29	3.10
$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	-7.84	-10.24	-7.73	-11.56	-6.71	-8.94

Average Effect of Capital Gains Tax Cuts. A natural question is: How much of the increase in volatility is imputable to the reduction in τ? This effect can be estimated by running the following counterfactual: fix τ_t at a particular level $\bar{\tau}$ for all t in 1982-2018; compute the difference between the baseline model statistic $\widetilde{S}_j(\hat{\theta})^{\text{post}}$ and the statistic for the constant tax $\widetilde{S}_j(\hat{\theta})^{\text{post}}|\tau_t = \bar{\tau}$. This difference is an average treatment effect that shows what would have happened without tax cuts after 1982. It is then divided by the increase in the statistic observed using US data, $\hat{S}_j^{\text{post}} - \hat{S}_j^{\text{pre}}$. This delivers a relative average treatment effect: the increase in the statistic due to tax cuts relative to the observed increase. Thus,

$$RATE_j = \frac{\widetilde{S}_j(\hat{\theta})^{\text{post}} - (\widetilde{S}_j(\hat{\theta})^{\text{post}}|\tau_t = \bar{\tau})}{\hat{S}_j^{\text{post}} - \hat{S}_j^{\text{pre}}} \times 100 \tag{63}$$

I selected $\bar{\tau} = 0.15$ as the baseline tax rate, representing the mean tax rate observed in the initial subsample and the highest level recorded between 1982 and 2018. This choice serves as a reasonable benchmark for a scenario without tax cuts in historical terms.

Table 5 reports the estimation for the main statistics and different versions of the model. Tax cuts explain approximately 40% of the observed increase in the variance of the PD ratio, a result robust to the exclusion of various model channels. Thus, even after excluding stock buybacks, transfers, or accounting for the absence of the Great Moderation, the proportion attributed to tax cuts remains relatively stable, ranging from 36.49% to 40.94%. However, the presence of the lock-in effect significantly alters the impact, accounting for 85% of the total increase in volatility. Therefore, considering the lock-in effect is essential to accurately evaluate the overall effect of tax cuts. Notably, when agents hold Rational Expectations, the impact of tax cuts on volatility is significant, 26.54%, although a bit lower than with Learning. That implies that Learning accounts for about 1/3 of the tax cut-induced increase of volatility.

Tax cuts primarily affect volatility by influencing the covariance between the PD ratio and future returns. In the full-fledged model, tax cuts account for approximately 70% of this increase, whereas their influence on the covariance between the PD ratio and future dividend growth remains modest, below 10%. The effect of tax cuts on the mean level of the PD ratio is also relatively modest, hovering slightly above 10% across various channels.

Table 5: ***Relative Average Effect of Capital Gains Tax Cuts (in %).*** *This table reports the relative average effect of tax cuts as defined in formula 68 for* $\bar{\tau} = 0.15$ *for a selected subsample of statistics. The columns are different versions of the model. All the versions are parameterized according to Table 2.*

	Full-fledged model	No Lock-in Effect	No transfers
$\mathbb{E}(PD_t)$	13.02	27.63	15.37
$\mathbb{V}ar(\ln P_t/D_t)$	39.83	84.98	56.62
$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	8.20	14.08	-13.54
$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	70.18	144.14	116.71
	No buybacks	**No learning (RE)**	**No Great Moderation**
$\mathbb{E}(PD_t)$	12.26	9.65	13.36
$\mathbb{V}ar(\ln P_t/D_t)$	36.49	26.54	40.94
$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	-24.14	1.30	22.22
$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	90.77	48.69	62.09

Alternative assumptions about taxes. How robust are these effects to changes in assumptions about tax news? So far, the assumption is that tax changes are unexpected. There exists, however, some evidence that markets show substantial ability to forecast taxes (see, for instance, Kueng (2014)). To address this concern, I try two different schemes. First, I assume perfect foresight of near-term taxes (which are indeed the only relevant, given the pricing function is unknown), such that $\mathbb{E}_t^{\mathcal{P}}(\tau_{t+1}) = \tau_{t+1}$. Additionally, I conjecture that investors are uncertain about the stochastic process followed by taxes and then they learn about it. Mimicking the price and dividends models they use, investors think of taxes using an unobserved component model which implies

$$\boldsymbol{\tau}_t = \boldsymbol{q}_t + \varepsilon_t \tag{64}$$

$$\boldsymbol{q}_t = \boldsymbol{q}_{t-1} + \varepsilon_t^q \tag{65}$$

where $\boldsymbol{\tau} = [\tau, \tau^D, \tau^B]'$, $\boldsymbol{q}$ 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 $\boldsymbol{q}_t|\omega^t \sim \mathcal{N}(\tilde{\boldsymbol{\tau}}_t, \Omega^\tau)$, with Ω^τ being the steady state Kalman filter uncertainty. Then,

$$\mathbb{E}_t^{\mathcal{P}}\Big[\boldsymbol{\tau}_{t+1}\Big] = \tilde{\boldsymbol{\tau}}_t \tag{66}$$

and

$$\tilde{\boldsymbol{\tau}}_t = \tilde{\boldsymbol{\tau}}_{t-1} + g^\tau(\boldsymbol{\tau}_t - \tilde{\boldsymbol{\tau}}_{t-1}) \tag{67}$$

The first and second columns of table 6 report the relative average treatment effect of capital gains tax cuts for this alternative tax expectations specifications.[32] The bottom line is that results are very similar to the baseline case. For tax foresight, tax cuts generate an increase in the variance of 38%, very close to the 40% of the baseline model. With learning about taxes, the effect is even stronger, reaching almost 50%. This larger increase occurs despite the tax expectations time series under learning is smoother, displaying less tax changes. While the original series exhibits a strong tax cut in the late 1970s that is largely reversed in the early 1980s, the smoother series displays a clearer downward trend since the 1970s which unambiguously leads to higher volatility. Whether agents have perfect foresight about the experiment or learn about it makes very little difference. Finally, the fact that tax-induced volatility goes mostly through a larger covariance between prices and returns also holds for both cases.

A key parameter is the tax elasticity of realization ξ. The estimated value is in the lower bound of the estimation in Zidar (2019). How would results change with a more negative value? The effect of CGT tax cuts on the mean PD ratio is barely changed, staying around 10%. Unsurprisingly, the effect of tax cuts on volatility is a third lower, accounting for 27% rather than 40% of the observed

[32] g^τ is set equal to g. Alternative g^τ barely affects the results.

increase.

The Average Effect of Capital Tax Cuts. While the paper has emphasized the role of capital gains taxes, dividend tax cuts were even more important. McGrattan and Prescott (2005) argued that these cuts would be behind an important part of the rise in the stock market capitalization to GDP, a statistic with a high comovement with the PD ratio. Are dividend tax cuts also behind the increase in the PD ratio variance? To address this question, I compute the relative average treatment effect of tax cuts. Following the above, procedure, I compute $\widetilde{S}_j(\hat{\theta})^{\text{post}}|\tau_t^D = \bar{\tau}^D$ with $\bar{\tau}^D = 0.26$, which is $max(\{\tau_t^D\}_{1982.III}^{2018.IV})$.

The third column of the table 6 reports the results for the selected statistics. Consistent with McGrattan and Prescott (2005), dividend tax cuts are behind 20% of the increase in the mean PD ratio. Thus, they seem to be much more important than capital gains tax to explain the run-up of the stock market. However, their contribution to the increase in the variance is much more modest, below 10%. Furthermore, they influence the variance only through $\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$. In other words, dividend and capital gains taxes are quite complementary in their effects on asset prices.

Additionally, so far I have considered the effect of tax cuts one at a time, but how would the stock market have behaved without both dividends and capital gains tax cuts altogether?[33] I answer this question by computing $(\widetilde{S}_j(\hat{\theta})^{\text{post}}|\tau_t = \bar{\tau}, \tau_t^D = \bar{\tau}^D)$. The first column of the table 6's bottom panel reports the effect for selected statistics. It turns out that the joint tax cuts accounted for 31% of the increase in the mean level of the stock market and 47% of the greater variance. This is approximately equal to the sum of the effects of each tax cut, that is

$$\widetilde{S}_j(\hat{\theta})^{\text{post}} - (\widetilde{S}_j(\hat{\theta})^{\text{post}}|\tau_t = \bar{\tau}, \tau_t^D = \bar{\tau}^D) \approx \left[\widetilde{S}_j(\hat{\theta})^{\text{post}} - (\widetilde{S}_j(\hat{\theta})^{\text{post}}|\tau_t = \bar{\tau})\right] + \left[\widetilde{S}_j(\hat{\theta})^{\text{post}} - (\widetilde{S}_j(\hat{\theta})^{\text{post}}|\tau_t^D = \bar{\tau}^D)\right] \tag{68}$$

The Average Effect of Stock Buybacks. The effects of buybacks on both the mean and the variance of the PD ratio are very similar to those of CGT cuts, explaining 10.86% and 34.91%, respectively. The main difference is that while tax cuts mostly impact $\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$, buybacks impact $\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$. Put differently, both tax cuts and stock buybacks seem to increase the variance of the PD ratio, but through different channels.

[33] I have not included the effects of τ^B because they are very small, as they only affect the risk-free rate.

Table 6: ***Additional Relative Average Treatment Effects (in %).*** *This table reports the relative average treatment effect when agents have perfect foresight about CGT tax cuts, when agents learn about taxes and when the tax elasticity of realization is lower. Besides, it reports the effect of other variables: dividend tax cuts, joint capital tax cuts and buybacks. All the versions are parameterized according to Table 2.*

	τ foresight	Tax Learning	$\xi = -0.4$
$\mathbb{E}(PD_t)$	12.67	11.02	9.09
$\mathbb{V}ar(\ln P_t/D_t)$	37.83	49.92	26.62
$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	9.74	20.32	7.23
$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	66.79	77.92	45.77
	τ^D	τ and τ^D	Buybacks
$\mathbb{E}(PD_t)$	19.15	30.86	10.86
$\mathbb{V}ar(\ln P_t/D_t)$	8.99	46.54	34.91
$\mathbb{C}ov(\ln P_t/D_t, \bar{d}_t)$	20.72	25.50	122.51
$\mathbb{C}ov(\ln P_t/D_t, \bar{r}_t)$	-2.80	68.20	-40.13

3.4.- Informational Efficiency

In line with the model's ability to mirror the variance decomposition of the PD ratio, it also effectively captures the observed reduction in the predictability of excess stock returns. However, the interpretation of this decline as an increase in efficiency is generally problematic by the Joint Hypothesis Problem (Fama (1970)). Instead, given the good quantitative performance of the model, this section studies informational efficiency within the model.

In the model with Learning, prices are not efficient for two reasons. First, payoffs are discounted using a subjective factor that differs from the equilibrium discount factor (i.e., the marginal rate of substitution in the real economy) (Adam et al. (2017)). Second, the expected payoffs themselves are produced from subjective statistical models without general equilibrium considerations. In other words, investors do not internalize the effect of their expectations on market prices, giving rise to non-fundamental volatility. For the opposite reasons, given the model, market efficiency is equivalent to Rational Expectations.

Hence, efficiency can be tested by tracking the deviations between Learning and Rational Expectations. Put differently, I compare a model with good quantitative properties but inefficient prices with the efficient version of the same model. In particular, mimicking a statistic used in portfolio analysis, let the Tracking Error be

$$\sigma_E = \sqrt{\mathbb{V}\Big(\frac{P_t^L}{D_t} - \frac{P_t^{RE}}{D_t}\Big)} \tag{69}$$

It expresses how closely the PD ratio in the world with Learning follows that benchmark PD ratio, using the standard deviation of their distance as the measure.

Table 7 presents the tracking error metrics, highlighting a notable increase in the full-fledged model, which sees an approximate 60% escalation. This significant uptick indicates a marked deviation of market prices from the theoretical efficient benchmark, suggesting a decrease in market efficiency. Notably, this decrease in efficiency is not attributable to any specific mechanism within the model. For instance, the absence of the lock-in effect or the cessation of stock repurchases tends to exacerbate this error even further.

Furthermore, the table delineates the impact of capital gains tax reductions on the tracking error. It reveals that tax cuts are responsible for 60% of the observed decline in informational efficiency, a proportion that remains robust across various model specifications. This indicates that the reduction in efficiency tied to tax cuts is a consistent outcome, unaffected by the deactivation of different model channels. This underscores the critical role of capital gains tax policy in influencing market efficiency.

*Table 7: **Informational Efficiency Test.** This table reports two statistics: i) the tracking error σ_E as defined by equation (69); ii) the Relative Average Treatment Effect (RATE) of capital gains tax cuts on σ_E for $\bar{\tau} = 0.15$. All versions of the model use table 2's parameters.*

	Full-fledged model		No Lock-in Effect		No transfers	
	1946-1982	1982-2018	1946-1982	1982-2018	1946-1982	1982-2018
σ_E	16.93	27.37	16.47	35.67	15.59	33.71
$RATE$	58.06		74.01		66.38	
	No buybacks		No learning (RE)		No Great Moderation	
	1946-1982	1982-2018	1946-1982	1982-2018	1946-1982	1982-2018
σ_E	16.95	24.32	0.00	0.00	16.93	29.87
$RATE$	83.26		0.00		60.92	

3.5.- The Equity Premium

While the main topic of the paper is the volatility of the Price-Dividend ratio and related questions as predictability or informational efficiency, the model has also some implications for the equity premium. This section explores them.

Table 8 reports the mean returns for bonds and stocks implied by the model. The average mean return is always above 4% and it is barely changed after the 1980s as in the data; the mean risk-free rate is close to 1% both pre- and post-1980s and, at odds with the data, it increases in the second subperiod in the model. While it is clear that the model does not produce a low enough risk-free rate, it is still remarkable that it generates an equity premium of about 4% while using realistic consumption and dividend growth 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).

Table 8: ***The model performance regarding the equity premium.*** *This table reports the the mean stock and bond returns from. The first two columns report the statistics for the US data. The next four columns report model-implied statistics and their t-statistics using the Table 2's parameters. The returns are in real terms and annualized.*

	US data		Baseline 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
$\mathbb{E}(r_t^s)$	4.73	4.34	4.78	-0.07	4.66	-0.42
$\mathbb{E}(r_t^b)$	0.41	0.35	0.72	-18.77	1.14	-24.32

To explore the drivers of the relatively good performance, I first examine stock returns through the following decomposition of their geometric mean, first suggested in Adam et al. (2016).

$$\underbrace{\left(\prod_{t=1}^{N}\frac{P_t+D_t}{P_{t-1}}\right)^{\frac{1}{N}}}_{rs} = \underbrace{\left(\prod_{t=1}^{N}\frac{D_t}{D_{t-1}}\right)^{\frac{1}{N}}}_{R_1}\underbrace{\left(\frac{PD_N+1}{PD_0}\right)^{\frac{1}{N}}}_{R_2}\underbrace{\left(\prod_{t=1}^{N-1}\frac{PD_t+1}{PD_t}\right)^{\frac{1}{N}}}_{R_3} \tag{70}$$

Thus, 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 could be related to the existence of a time trend. Finally, the last term (R_3) is a convex function of period t PD ratio. It increases with the volatility of the PD time series, but decreases with its mean.

Table 9 reports the decomposition using empirical and simulated data. Since the dividend growth process has been parameterized directly from the data, the models exactly replicate R_1. Regarding R_2, the model moves in the right direction, although not as strongly as in the data. This relative mismatch of the model is partly due to the pre and post-1982 division; the CGT falls temporarily at the end of the first subsample and beginning of the second subsample, bringing the PD ratio up during the final quarter of the 1st subsample and the initial quarters of the 2nd subperiod. As a result, R_2 gets too high (low) in 1946-1982 (1982-2018). R_3 is reasonably close to the data, as expected from its good behavior in terms of the mean and the variance of the PD ratio.

While macro-finance models typically need either a high risk-aversion level or high fundamental

Table 9: ***Decomposition of the stock return geometric mean.*** *The table shows the stock returns mean decomposition according to expression (70). The first column uses U.S. data; the second, simulated data using the baseline model, using the parameters in table 2.*

	US data		Baseline model	
	1946-1982	1982-2018	1946-1982	1982-2018
$\mathbb{E}(R_1)$	0.46	0.74	0.46	0.74
$\mathbb{E}(R_2)$	-0.27	0.88	0.41	0.73
$\mathbb{E}(R_3)$	4.10	2.37	3.83	2.55
$\mathbb{E}(rs)$	4.30	4.03	4.74	4.06

volatility to match the mean returns, this paper resort to alternative forces. 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 is unable to do all the job (Adam et al. (2016), Adam et al. (2017)). Thus, the introduction of tax cuts helps in getting an increase in R_2. Crucially, these elements barely affect the risk-free rate, since there is no feedback loop affecting the bond price due to its one-period maturity.

4.- Conclusion

This paper has analyzed how a Capital Gains Tax influences asset price booms and busts. I have studied this issue using an asset pricing model à la Lucas (1978) with learning about prices as in Adam et al. (2017) and a tax on capital gains upon realization. The central theoretical result is that the variance of the Price-Dividend ratio is decreasing on the CGT level, since taxes dampen the pass-through from beliefs to prices. Although the existence of a lock-in effect counteracts this logic, the ability of taxes to stabilize taxes prevails if the tax elasticity of realization is not too negative.

The theory has been applied to the United States to connect the recurrence of asset price cycles in the middle of the Great Moderation, a troubling observation for many macro-finance models, to the observed decline in the effective CGT. The structural estimation of the model reveals that the CGT cuts account for 40% of the observed increase in the volatility of the stock market. The model also replicates a comprehensive list of asset pricing moments and their dynamics such as the rise in stock market valuations, a sizable equity premium, the reduction in return predictability, the procyclicality of beliefs, or the increase in the elasticity of prices to beliefs. Furthermore, the model implies a notable reduction in stock market informational efficiency, related to higher role of subjective beliefs partly induced by lower taxes.

The research has left some issues open. The decline in capital taxes since the 1980s was a global phenomenon, so an international analysis of its effects on financial instability and its possible interaction with financial deregulation and capital flows liberalization appears as an interesting research avenue. Besides, corporate taxes went down substantially in the same period, which might directly affect the payouts that I took as exogenous, as well as having interesting implications for investment or productivity dynamics. Furthermore, the theoretical framework has not addressed the issue of optimal capital gains taxation, considering both financial stability and public finance perspectives.

In conclusion, the paper's discussions imply that a tax on capital gains could play a relevant role in regulating asset price booms and busts. Hence, while the ability of a Financial Transaction Tax to prevent excessive price volatility has been widely questioned, this paper suggests that a Capital Gains Tax might be an effective alternative.

References

Abel, A. B. (1990). Asset prices under habit formation and catching up with the joneses. *The American Economic Review, Papers and Proceedings*, 80(2):38–42.

Adam, K. and Marcet, A. (2011). Internal rationality, imperfect market knowledge and asset prices. *Journal of Economic Theory*, 146(3):1224–1252.

Adam, K., Marcet, A., and Beutel, J. (2017). Stock price booms and expected capital gains. *American Economic Review*, 107(8):2352–2408.

Adam, K., Marcet, A., and Nicolini, J. P. (2016). Stock market volatility and learning. *The Journal of Finance*, 71(1):33–82.

Agersnap, O. and Zidar, O. (2021). The tax elasticity of capital gains and revenue-maximizing rates. *American Economic Review: Insights*, 3(4):399–416.

Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, 59(4):1481–1509.

Bordalo, P., Gennaioli, N., La Porta, R., O'Brien, M., and Shleifer, A. (2023). Long term expectations and aggregate fluctuations. *NBER Chapters*.

Bordalo, P., Gennaioli, N., Porta, R. L., and Shleifer, A. (2019). Diagnostic expectations and stock returns. *The Journal of Finance*, 74(6):2839–2874.

Brun, L. and González, I. (2017). Tobin's q and inequality. *Available at SSRN 3069980*.

Bullard, J. and Duffy, J. (2001). Learning and excess volatility. *Macroeconomic Dynamics*, 5(02):272–302.

Burnside, C. (1998). Solving asset pricing models with gaussian shocks. *Journal of Economic Dynamics and control*, 22(3):329–340.

Buss, A. and Dumas, B. (2019). The dynamic properties of financial-market equilibrium with trading fees. *The Journal of Finance*, 74(2):795–844.

Buss, A., Dumas, B., Uppal, R., and Vilkov, G. (2016). The intended and unintended consequences of financial-market regulations: A general-equilibrium analysis. *Journal of Monetary Economics*, 81:25–43.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, 107(2):205–251.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228.

Cappelletti, G., Guazzarotti, G., and Tommasino, P. (2017). The stock market effects of a securities transaction tax: quasi-experimental evidence from italy. *Journal of Financial Stability*, 31:81–92.

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

Cochrane, J. H. (2017). Macro-finance. *Review of Finance*, 21(3):945–985.

Cogley, T. and Sargent, T. J. (2008). The market price of risk and the equity premium: A legacy of the great depression? *Journal of Monetary Economics*, 55(3):454–476.

Constantinides, G. M. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of political Economy*, 98(3):519–543.

Dai, Z., Maydew, E., Shackelford, D. A., and Zhang, H. H. (2008). Capital gains taxes and asset prices: capitalization or lock-in? *The Journal of Finance*, 63(2):709–742.

Dai, Z., Shackelford, D. A., and Zhang, H. H. (2013). Capital gains taxes and stock return volatility. *The Journal of the American Taxation Association*, 35(2):1–31.

Dávila, E. (2023). Optimal financial transaction taxes. *The Journal of Finance*, 78(1):5–61.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

Feenberg, D. and Coutts, E. (1993). An introduction to the taxsim model. *Journal of Policy Analysis and management*, 12(1):189–194.

Gavin, W. T., Keen, B. D., and Kydland, F. E. (2015). Monetary policy, the tax code, and the real effects of energy shocks. *Review of Economic Dynamics*, 18(3):694–707.

Gavin, W. T., Kydland, F. E., and Pakko, M. R. (2007). Monetary policy, taxes, and the business cycle. *Journal of Monetary Economics*, 54(6):1587–1611.

Gemmill, R. F. (1956). The effect of the capital gains tax on asset prices. *National Tax Journal*, 9(4):289–301.

Giglio, S., Maggiori, M., Stroebel, J., and Utkus, S. (2021). Five facts about beliefs and portfolios. *American Economic Review*, 111(5):1481–1522.

Greenwood, R. and Shleifer, A. (2014). Expectations of returns and expected returns. *The Review of Financial Studies*, 27(3):714–746.

Hakansson, N. H. (1970). Optimal investment and consumption strategies under risk for a class of utility functions. *Econometrica*, 38(5):587–607.

Haugen, R. A. and Heins, A. J. (1969). The effects of the personal income tax on the stability of equity value. *National Tax Journal*, 22(4):466–471.

Haugen, R. A. and Wichern, D. W. (1973). The diametric effects of the capital gains tax on the stability of stock prices. *The Journal of Finance*, 28(4):987–996.

Jin, L. J. and Sui, P. (2022). Asset pricing with return extrapolation. *Journal of Financial Economics*, 145(2):273–295.

Judd, K. L. (1992). Projection methods for solving aggregate growth models. *Journal of Economic theory*, 58(2):410–452.

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

Kindleberger, C. P. (1978). Manias, panics and crashes: a history of financial crises.

Kueng, L. (2014). Tax news: The response of household spending to changes in expected taxes. Technical report, National Bureau of Economic Research.

Lerner, A. P. (1943). Functional finance and the federal debt. *Social research*, pages 38–51.

Lettau, M., Ludvigson, S. C., and Wachter, J. A. (2008). The declining equity premium: What role does macroeconomic risk play? *The Review of Financial Studies*, 21(4):1653–1687.

Ljung, L. (1977). Analysis of recursive stochastic algorithms. *IEEE transactions on automatic control*, 22(4):551–575.

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

Malmendier, U. and Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87.

Marcet, A. (1988). Solving nonlinear stochastic growth models by parametrizing expectations.

Martin, A. and Ventura, J. (2018). The macroeconomics of rational bubbles: a user's guide. *Annual Review of Economics*, 10:505–539.

McGrattan, E. R. and Prescott, E. C. (2003). Average debt and equity returns: Puzzling? *American Economic Review*, 93(2):392–397.

McGrattan, E. R. and Prescott, E. C. (2005). Taxes, regulations, and the value of us and uk corporations. *The Review of Economic Studies*, 72(3):767–796.

Minsky, H. P. (1976). *John Maynard Keynes*. Springer.

Nagel, S. and Xu, Z. (2022). Asset pricing with fading memory. *The Review of Financial Studies*, 35(5):2190–2245.

Pástor, L. and Veronesi, P. (2006). Was there a nasdaq bubble in the late 1990s? *Journal of Financial Economics*, 81(1):61–100.

Pástor, L. and Veronesi, P. (2009). Technological revolutions and stock prices. *American Economic Review*, 99(4):1451–1483.

Rosenthal, S. and Austin, L. (2016). The dwindling taxable share of us corporate stock. *Tax Notes*, 151(6).

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

Sialm, C. (2006). Stochastic taxation and asset pricing in dynamic general equilibrium. *Journal of Economic Dynamics and Control*, 30(3):511–540.

Sialm, C. (2009). Tax changes and asset pricing. *American Economic Review*, 99(4):1356–83.

Sikes, S. A. and Verrecchia, R. E. (2012). Capital gains taxes and expected rates of return. *The Accounting Review*, 87(3):1067–1086.

Somers, H. M. (1948). An economic analysis of the capital gains tax. *National Tax Journal*, 1(3):226–232.

Somers, H. M. (1960). Reconsideration of the capital gains tax. *National Tax Journal*, 13(4):289–309.

Stiglitz, J. E. (1983). Some aspects of the taxation of capital gains. *Journal of Public Economics*, 21(2):257–294.

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

Timmermann, A. G. (1993). How learning in financial markets generates excess volatility and predictability in stock prices. *The Quarterly Journal of Economics*, 108(4):1135–1145.

Tobin, J. (1978). A proposal for international monetary reform. *Eastern economic journal*, 4(3/4):153–159.

Umlauf, S. R. (1993). Transaction taxes and the behavior of the swedish stock market. *Journal of Financial Economics*, 33(2):227–240.

Zidar, O. (2019). Tax cuts for whom? heterogeneous effects of income tax changes on growth and employment. *Journal of Political Economy*, 127(3):1437–1472.

Appendix A: Data Sources

Stock market data. Stock prices, dividends and CPI inflation comes from Robert Shiller database. They can be downloaded here: `http://www.econ.yale.edu/~shiller/data.htm`. The risk-free rate is the 90 days T-Bill, from the FRED database `https://fred.stlouisfed.org/series/TB3MS`.

The data has been transformed into quarterly frequency by taking the last month of the considered quarter. Besides, the nominal variables have been transformed to real terms using Shiller's CPI inflation index. Finally, as is standard in the literature, I have deseasonalize dividends (by taking the average over the current and past 3 quarters) to compute the price-dividend ratio.

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. They can be found here `https://taxsim.nber.org/marginal-tax-rates/`. 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 national data for each state and year.

Following Sialm (2009), I adjusted for state and local taxes before 1979 and after 2008 as well as for the distinction between qualified and non-qualified dividends from 2003 on to get a complete series for the 1960-2018 period. Before 1960, τ_t^d, τ_t^{skg}, τ_t^{lkg} and rates are taken from Sialm (2009). τ_t^B are interpolated.

The weights for the convex combination are computed using the dividend, short and long capital gains yields offered by Sialm (2009). They are averaged over the 1954-2006 period. Letting them vary barely change the synthetic rate. For details on the taxable share, see Appendix C.

Capital gains. The total realized capital gains are a 5 year moving average on the capital gains reported in the adjusted gross income, coming from the IRS. As for total capital gains, I use a 5 year moving average of the nominal taxable gains, obtained from the Financial Accounts. I am grateful to Jacob Robbins for providing these data, coming from his paper **?**. The portion of capital gains coming from equities is obtained from the US Financial Accounts, covering the 1951-2018 on a quarterly basis. Finally, the portion of realized capital gains coming from equities is computed using data from the IRS for the year 1985 and 1997-2012.

Survey expectations. For the test of the tax indirect effect, I have used the UBS survey is the UBS Index of Investor Optimism. The quantitative question on stock market expectations has been surveyed over the period Q2:1998-Q4:2007 with 702 responses per month on average. To make the data consistent with the model, I have run some adjustment. 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 price-dividend ratio.

Appendix B: Proofs.

B1.- Proof of Proposition 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,

$$\frac{P_t}{D_t} \approx p(\boldsymbol{b}) + \boldsymbol{\omega}(\boldsymbol{\beta}_t^g - \boldsymbol{b}) \tag{71}$$

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

$$\mathbb{V}\Big[\frac{P_t}{D_t}\Big] \approx \boldsymbol{\omega}(\tau)\boldsymbol{\Sigma}\boldsymbol{\omega}(\tau)^T \tag{72}$$

where $\boldsymbol{\Sigma}$ is the variance-covariance matrix of beliefs. The effect of taxes on the mapping from $\boldsymbol{\Sigma}$ to the PD variance is determined by the signs and magnitudes of the elements of the following matrix

$$\frac{\partial}{\partial \tau}\nabla_{\Sigma}\mathbb{V}(P_t/D_t) = \begin{bmatrix} 2\omega_1\omega_1' & \omega_1'\omega_2 + \omega_1\omega_2' & \omega_1'\omega_3 + \omega_1\omega_3' \\ \omega_1'\omega_2 + \omega_1\omega_2' & 2\omega_2\omega_2' & \omega_2'\omega_3 + \omega_2\omega_3' \\ \omega_1'\omega_3 + \omega_1\omega_3' & \omega_2'\omega_3 + \omega_2\omega_3' & 2\omega_3\omega_3' \end{bmatrix} \tag{73}$$

where ω_j' is the derivative with respect to τ. I provide sufficient conditions which are reasonable to establish the signs of the matrix elements. Start with the partial derivatives of the PD ratio with respect to beliefs (keeping them as a function of $\boldsymbol{\beta}_t^g$):

$$\omega_1 = \frac{1}{1 - (1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M} \tag{74}$$

$$\omega_2 = \frac{\beta_t^D(1-\pi\tau)}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2} \tag{75}$$

$$\omega_3 = \frac{\beta_t^D\pi\tau}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2} \tag{76}$$

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 if $1 > \beta_t^P - \pi\tau(\beta_t^P - \beta_t^M)$. Condition i and ii are sufficient for that. Let us look now at $\frac{\partial}{\partial\tau}\boldsymbol{\omega}$:

$$\omega_1' = -\frac{\pi(\beta_t^P - \beta_t^M)}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2} \tag{77}$$

$$\omega_2' = -\frac{\pi\beta_t^D}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2}\Big(1 + \frac{2(1-\pi\tau)(\beta_t^P - \beta_t^M)}{1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M}\Big) \tag{78}$$

$$\omega_3' = \frac{\pi\beta_t^D}{(1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M)^2}\Big(1 - \frac{2\pi\tau(\beta_t^P - \beta_t^M)}{1-(1-\pi\tau)\beta_t^P - \pi\tau\beta_t^M}\Big) \tag{79}$$

$\beta_t^P > \beta_t^M$ (implied by condition ii) ensures ω_1' and ω_2' are negative. On the contrary, a sufficient condition for ω_3' to be positive is $\beta_t^P - \beta_t^M < \frac{1-\beta_t^P}{\pi\tau}$, which is likely to be met for standard parameters values. Given these signs, it is clear that all the elements in the matrix except those involving ω_3' are negative. Now I proceed to check these elements. It turns out that condition ii. is sufficient to ensure $\omega_1'\omega_3 + \omega_1\omega_3' < 0$ and $\omega_2'\omega_3 + \omega_2\omega_3' < 0$. Under conditions i and ii, the only positive element is $2\omega_3\omega_3'$. Condition iii guarantees that $\omega_3\omega_3'\mathbb{V}(\beta_t^M) < \omega_2\omega_2'\mathbb{V}(\beta_t^P)$. The only condition that has not been justified is $\beta_t^M < 1$. Take condition ii and isolate β_t^P such that $\beta_t^P > \frac{1+3\pi\tau\beta_t^M}{1+3\pi\tau}$. If $\beta_t^M \geq 1$, then $\beta_t^P \geq 1$, which contradicts condition i.

B2.- Proof of Proposition 2.

Proposition 2.1. Rational Expectations.

Consider the RE pricing formula (10). It follows $P_t^{RE}/D_t = p_1(d_t)$. Take a first-order Taylor approximation around the constant μ. Then,

$$\frac{P_t^{RE}}{D_t} \approx p_1(\mu) + \omega^{RE}(d_t - d) \tag{80}$$

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

$$\mathbb{V}\Big[\frac{P_t}{D_t}\Big] \approx (\omega^{RE})^2\mathbb{V}(d_t) \tag{81}$$

From (8), $\mathbb{V}(d_t) = \frac{\sigma_d^2}{1-\rho^2}$. Then, taxes affect the variance of the PD ratio only through ω. To study this channel, consider

$$\frac{\partial(\omega^{RE})^2}{\partial\tau} = 2\omega^{RE}\frac{\partial\omega^{RE}}{\partial\tau} \tag{82}$$

ω^{RE} reads as

$$\omega^{RE} = \sum_{j=1}^{\infty} f_0(\tau, j) f_1(\mu, j) \frac{\rho}{1-\rho}(1-\rho^j) \tag{83}$$

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

$$\frac{\partial\omega^{RE}}{\partial\tau} = \sum_j f_1(\mu, j)\frac{\rho}{1-\rho}(1-\rho^j)\frac{\partial f_0(\tau, j)}{\partial\tau} \tag{84}$$

Given,

$$f_0(\tau, j) = \Big(\frac{\delta}{1-\delta\pi\tau}\Big)^j (1-\pi\tau)^{j-1} \tag{85}$$

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

$$\frac{\partial f_0(\tau, j)}{\partial\tau} = \frac{\pi\delta^j(1-\pi\tau)^{j-1}}{(1-\delta\pi\tau)^j}\Big(\frac{j\delta(1-\pi\tau)}{1-\delta\pi\tau} - (j-1)\Big) \tag{86}$$

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,

$$\frac{\partial\omega^{RE}}{\partial\tau} = \frac{\rho}{1-\rho}\Bigg(\underbrace{\sum_{j=1}^{\tilde{j}} f_1(\mu, j)(1-\rho^j)\frac{\partial f_0(\tau, j)}{\partial\tau}}_{\text{Factor 1>0}} + \underbrace{\sum_{\tilde{j}=j+1}^{\infty} f_1(\mu, j)(1-\rho^j)\frac{\partial f_0(\tau, j)}{\partial\tau}}_{\text{Factor 2<0}}\Bigg) \tag{87}$$

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

$$\lim_{j\to\infty}\frac{\partial f_0(\tau, j)}{\partial\tau} = -\infty \tag{88}$$

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

$$\frac{\partial f_1(\mu,j)}{\partial j} = f_1(d_t,j)\left(\mu + \frac{\sigma_d^2}{2(1-\rho)^2}\left(1+\frac{2\rho^j j}{1-\rho} - \frac{2\rho^{2j+1}j}{1-\rho^2}\right)\right) \tag{89}$$

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.

Proposition 2.2. Learning.

Consider the equilibrium pricing function under Learning, equation (20) 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,

$$\frac{P_t^L}{D_t} \approx p_2(\boldsymbol{x}) + \boldsymbol{\omega^L}(\boldsymbol{x_t}-\boldsymbol{x}) \tag{90}$$

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

$$\mathbb{V}\left[\frac{P_t}{D_t}\right] \approx \boldsymbol{\omega}^L\boldsymbol{\Sigma}^L(\boldsymbol{\omega}^L)^T \tag{91}$$

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

$$\omega_1^L = \rho\frac{\delta exp\Big\{\mu+\sigma_d^2/2\Big\}}{1-(1-\pi\tau)\delta exp\Big\{\mu+\frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}-\delta\pi\tau} \tag{92}$$

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

$$\frac{\partial\omega_1^L}{\partial\tau} = -\frac{\rho\delta exp\Big\{\mu+\sigma_d^2/2\Big\}\pi\delta\Big(exp\Big\{\mu+\frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}-1\Big)}{\Big(1-(1-\pi\tau)\delta exp\Big\{\mu+\frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}-\delta\pi\tau\Big)^2} \tag{93}$$

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,

$$\omega_2^L = \frac{\delta^2 exp\Big\{\mu+\sigma_d^2/2\Big\}(1-\pi\tau)\Big(exp\Big\{\mu+\frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}-1\Big)}{\Big(1-(1-\pi\tau)\delta exp\Big\{\mu+\frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\}-\delta\pi\tau\Big)^2} \tag{94}$$

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

$$\frac{\partial \omega_2^L}{\partial \tau} = -\frac{\rho\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)}{\Big(1-(1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau\Big)^2}$$
$$-\frac{2\delta^2 exp\Big\{\mu + \sigma_d^2/2\Big\}(1-\pi\tau)\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big)}{\Big(1-(1-\pi\tau)\delta exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau\Big)^3}\delta\pi\Big(exp\Big\{\mu + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - 1\Big) \tag{95}$$

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 $\ln\beta_t$

$$\ln\beta_t = \ln\beta_{t-1} + g\left(\ln\frac{exp\Big\{(1-\rho)\mu + \rho d_{t-1} + \sigma_d^2/2\Big\}}{exp\Big\{(1-\rho)\mu + \rho d_{t-2} + \sigma_d^2/2\Big\}}\frac{1-(1-\pi\tau)\delta exp\Big\{\ln\beta_{t-2} + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau}{1-(1-\pi\tau)\delta exp\Big\{\ln\beta_{t-1} + \frac{\sigma^2+\sigma_b^2+\sigma_P^2}{2}\Big\} - \delta\pi\tau} + d_{t-1} - \ln\beta_{t-1}\right) \tag{96}$$

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,

$$\ln\beta_t \approx \kappa + \mathcal{A}\ln\beta_{t-1} + \mathcal{B}\ln\beta_{t-2} + \mathcal{C}d_{t-1} + \mathcal{D}d_{t-2} \tag{97}$$

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):

$$\begin{bmatrix}\ln\beta_t \\ d_t\end{bmatrix} = \begin{bmatrix}\kappa \\ (1-\rho)\mu\end{bmatrix} + \begin{bmatrix}\mathcal{A} & \mathcal{C} \\ 0 & \rho\end{bmatrix}\begin{bmatrix}\ln\beta_{t-1} \\ d_{t-1}\end{bmatrix} + \begin{bmatrix}\mathcal{B} & \mathcal{D} \\ 0 & 0\end{bmatrix}\begin{bmatrix}\ln\beta_{t-2} \\ d_{t-2}\end{bmatrix} + \begin{bmatrix}0 \\ \varepsilon_t^d\end{bmatrix} \tag{98}$$

Transform it into a VAR(1):

$$\begin{bmatrix}\ln\beta_t \\ d_t \\ \ln\beta_{t-1} \\ d_{t-1}\end{bmatrix} = \begin{bmatrix}\kappa \\ (1-\rho)\mu \\ 0 \\ 0\end{bmatrix} + \begin{bmatrix}\mathcal{A} & \mathcal{C} & \mathcal{B} & \mathcal{D} \\ 0 & \rho & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}\ln\beta_{t-1} \\ d_{t-1} \\ \ln\beta_{t-2} \\ d_{t-2}\end{bmatrix} + \begin{bmatrix}0 \\ \varepsilon_t^d \\ 0 \\ 0\end{bmatrix} \tag{99}$$

or, in matrix form,

$$\boldsymbol{Z}_t = \tilde{\boldsymbol{\kappa}} + \Phi \boldsymbol{Z}_{t-1} + \boldsymbol{\eta}_t \tag{100}$$

with

$$\boldsymbol{\eta}_t \sim \mathcal{N}\left[\begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0 & 0 & 0 & 0\\ 0 & \sigma_d^2 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\right] \tag{101}$$

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

$$\Gamma = \Phi \Gamma \Phi^T + \tilde{\Omega} \tag{102}$$

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,

$$\text{vec}(\Gamma) = (\Phi \otimes \Phi)\text{vec}(\Gamma) + \text{vec}(\tilde{\Omega}) \tag{103}$$

Then,

$$\text{vec}(\Gamma) = (I - \Phi \otimes \Phi)^{-1}\text{vec}(\tilde{\Omega}) \tag{104}$$

Solving this equation, we get

$$\mathbb{V}(\ln\beta_t) = \underbrace{\frac{2\mathcal{C}\mathcal{D}(\mathcal{A} + \rho(1 - \mathcal{B}^2 + \mathcal{A}\mathcal{B}\rho)) + (\mathcal{C}^2 + \mathcal{D}^2)(1 + \mathcal{A}\rho + \mathcal{B}(\rho(\mathcal{A} + \rho(1 - \mathcal{B}))))}{((1 + \mathcal{B})(1 + \mathcal{A} - \mathcal{B})(1 - \mathcal{A} - \mathcal{B})(1 - \rho(\mathcal{A} + \mathcal{B}\rho))}}_{\equiv \nu(\tau)} \frac{\sigma_d^2}{(1 - \rho^2)} \tag{105}$$

$$\mathbb{C}ov(\ln\beta_t, d_t) = \underbrace{\frac{\rho(\mathcal{C} + \rho\mathcal{D})}{(1 - \rho(\mathcal{A} + \rho\mathcal{B}))}}_{\equiv \upsilon(\tau)} \frac{\sigma_d^2}{(1 - \rho^2)} \tag{106}$$

The conditions required for stationarity ensure $\mathbb{V}(\ln\beta_t) > 0$ and $\mathbb{C}ov(\ln\beta_t, d_t)$ [show it]. It can be shown that $\partial\mathbb{V}(\ln\beta_t)/\partial\tau < 0$ when g is small and $\pi\tau$ are not too large [show it]. Focus now on the effect of taxes on

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

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.

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 10 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.

Figure 7 plots the estimated taxable share from 1951:IV to 2018:IV. As observed, it displays a steady decline until the early 2000s, when stabilizes around 30%. In other words, there was a big structural change in the stock ownership, moving it away from taxable units.

Table 10: ***Taxable share estimation.*** *Steps to compute the taxable share.*

	Total CE HHNPI	
1.- Substract foreign equities	- RoW x (Total CE HHNPI /Total CE All Sectors)	
	= HHNPI domestic CE	
2.- Substract the stocks issued by the passthrough entities	-S Corporations – (ETF + CEF + REITs) x HH share of Mutual Funds	
S corporations, ETFs, CEF and REITs	= HHNPI domestic C-CE	
3.- Substract NPI holdings	-NPI 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)$_{1987}$
	= HH domestic C-CE	
4.- Substract IRAs and 529 savings plans holdings	- IRA C-CE - 529 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
	= Direct HH domestic C-CE	
5.- Add Indirect Holdings of C Corporation Equity	+ Indirect HH domestic C-CE	Indirect HH C-CE = (CE MF + (CE ETF + CE CEF) x HH share of MF) x Direct HH domestic CE / Total HHNPI CE
	= HH Taxable 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

Figure 7: ***Taxable share evolution.*** *The graph plots the taxable share of equity income, estimated following the procedure explained above. It uses data from 1951:IV to 2018:IV.*

Appendix E: Projection facility

The equilibrium PD ratio given by **??** faces a discontinuity. For this reason, simulation requires to set up the following modified belief updating equation to ensure non-negative prices

$$\beta_{t+1} = w\left(exp\left\{ln\beta_t(1-g) + gln\frac{P_t}{P_{t-1}}\right\}\right) \tag{108}$$

where

$$w(x) = \begin{cases} x & \text{if} \quad x \leq \beta_t^L \\ \beta_t^L + \frac{x-\beta_t^L}{x+\beta_t^U-2\beta_t^L}(\beta_t^U - \beta_t^L) & \text{if} \quad x > \beta_t^L \end{cases} \tag{109}$$

and

$$\beta_t^q = PD^q\left\{PD^q\xi\delta(1-\pi\tau_{t+1}^K)^2 + \chi\delta(1-\pi\tau_{t+1}^K)\left(\frac{W_{t+1}}{D_{t+1}} + 1 - \tau_{t+1}^D + \pi\tau_{t+1}^K\frac{P_t}{D_t}\frac{D_t}{D_{t+1}}\right)\right\}^{-1}$$

(110)

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, (109) is similar to the one used by Adam et al. (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.

Appendix F: Parameterized Expectations Algorithm

In the spirit of Hakansson (1970), the proposed approximating function ψ is

$$\frac{C_t^*}{D_t} = \bar{\mathcal{E}}(\boldsymbol{X}_t) \approx \psi(\boldsymbol{X}_t; \chi) = c_t^y Y_t + c_t^w \frac{P_t}{D_t} S_{t-1} \tag{111}$$

where $c_t^y \equiv 1 - \chi\delta(1-\tau_t^D)\beta^D$ is the time-varying propensity to consume out of income, Y_t collects all the income sources (wages, dividends, net transfers) normalized by dividends, $c_t^w \equiv 1 - \chi\delta(1-\tau_t^K)\beta_t$ is the propensity to consume out of wealth, and χ is a parameter of ψ to be estimated. To evaluate the performance of this approximating function, χ must be estimated. To do so, I resort to simulation and Montecarlo integration. The algorithm involves the following steps:

1. Draw a series of the exogenous processes for a large T.

2. For a given $\boldsymbol{\chi} \in \mathbb{R}^n$, recursively compute the series of the endogenous variables.

3. Minimize the Euler Equation square residuals using non-linear least squares

$$G(\boldsymbol{\chi}) = \underset{\boldsymbol{\xi}\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{(T-\underline{T})} \sum_{t=\underline{T}}^{T} \left[\phi\Big(z_{t+1}^{\mathcal{P}}(\boldsymbol{\chi}), \varepsilon_{t+1}, z_t(\boldsymbol{\chi})\Big) - \frac{\psi(X_t(\boldsymbol{\chi}); \xi)^{-\gamma}}{\delta} \right]^2$$

with $\underline{T}$ are some initial periods burned. ϕ is the interior of the conditional expectation $\bar{\mathcal{E}}(X_t)$, z are the endogenous variables and ϵ the exogenous shocks.

Note the interior of the expectation must be computed according to investor's beliefs. Since investors know the process for dividends and wage-dividends, the only problematic objects are next period prices and next period consumption. Using agents subjective price model

$$\beta_{t+1}^{\mathcal{P}} = \beta_t^i \nu_{t+1} \Rightarrow \left(\frac{P_{t+1}}{P_t}\right)^{\mathcal{P}} = \beta_t^i \nu_{t+1} \varepsilon_{t+1}^p \Rightarrow \left(\frac{P_{t+1}}{D_{t+1}}\right)^{\mathcal{P}} = \left(\frac{P_{t+1}}{P_t}\right)^{\mathcal{P}} \frac{D_t}{D_{t+1}} \frac{P_t}{D_t}$$

In turn, expected consumption reads

$$\frac{C^{\mathcal{P}}_{t+1}}{D_{t+1}} = (1-\chi\delta(1-\pi\tau^K_{t+1})\beta^{\mathcal{P}}_{t+1})\left(\left(\frac{P_{t+1}}{D_{t+1}}\right)^{\mathcal{P}} + 1 - \tau^D_{t+1} + \frac{W_{t+1}}{D_{t+1}} - \pi\tau^K_{t+1}\left[\left(\frac{P_{t+1}}{D_{t+1}}\right)^{\mathcal{P}} - \frac{P_t}{D_t}\frac{D_t}{D_{t+1}}\right]\right) \tag{112}$$

4. Find a fixed point $\chi = G(\chi)$. For that, update χ following

$$\chi^{j+1} = \chi^j + d(G(\chi^j) - \chi^j) \tag{113}$$

where j iteration number and d the dampening parameter.

To evaluate how good the approximation is, I explore the size of the errors in consumption terms. Approximation errors are given by

$$u_{t+1} = \delta\phi\Big(z_{t+1}, \varepsilon_{t+1}, z_t\Big)^{-1/\gamma} - \psi(\chi; x_t)$$

The criterion to determine the degree of accuracy is the Bounded Rationality Measure (Judd (1992)):

$$J = \log_{10}\left(\mathbb{E}_t\left|\frac{u_{t+1}}{\frac{C_t}{D_t}}\right|\right) \tag{114}$$

being J a dimension-free quantity that expresses that error as a fraction of current consumption. For the baseline model, J = -5.99. It is equivalent to a mistake of $1 out of a million. The Mean Square Error is 5.71e-06. Figure 8 plots the histogram of J for 10.000 simulations of the model.

Solving the model with the lock-in effect

Algorithm to compute m_t:

1. Approximate the conditional expectation determining μ_{t+1} via a function $\Psi(\boldsymbol{X}_t)$, where $\boldsymbol{X}_t$ is a vector of state variables, that is:

$$\mu_{t+1} = (\tau^K)^{1+\xi}M_{t+1} = (\tau^K)^{1+\xi}\mathcal{E}(\boldsymbol{X}_{t+1}) \approx (\tau^K)^{1+\xi}\Psi(\boldsymbol{X}_{t+1}) \tag{115}$$

In particular, I use this linear polynomial:

$$\Psi(\boldsymbol{X}_t) = \alpha_0 + \alpha_1\beta_t + \alpha_2 G_t + \alpha_3\frac{P_t}{D_t} \tag{116}$$

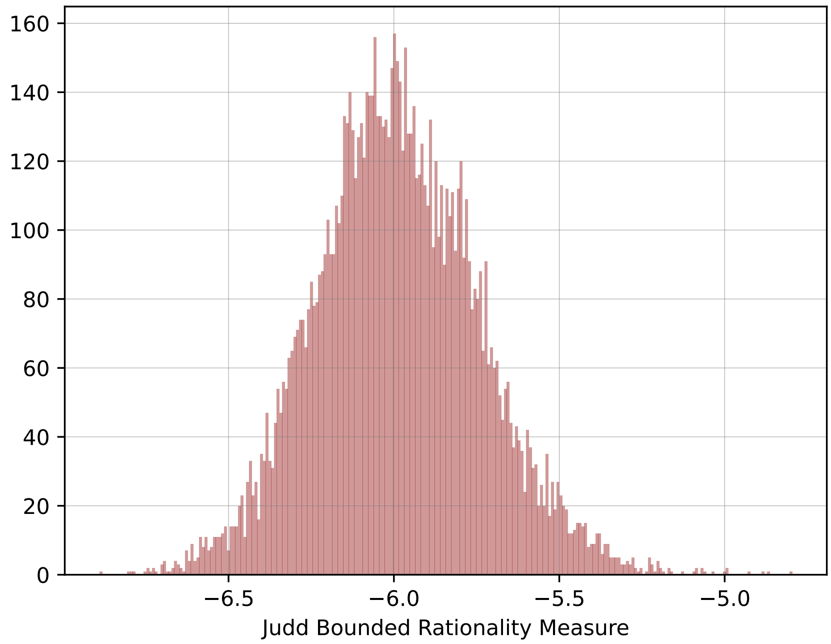

Figure 8: ***Histogram of the Judd Bounded Rationality Measure.*** *The histogram plots the Judd criterion as defined by equation (114) resulting from 10.000 simulations of the model.*

2. For a given vector $\boldsymbol{\alpha}$, compute the expectation of μ_t conditional on the information at time t, that is

$$m_t = \mathbb{E}_t^{\mathcal{P}}(\mu_{t+1}) = \mathbb{E}_t^{\mathcal{P}}(\mathcal{E}(\boldsymbol{X}_{t+1})) \approx \mathbb{E}_t^{\mathcal{P}}(\Psi(\boldsymbol{X}_{t+1})) \tag{117}$$

   (a) Future state variables depend on four shocks:

      i. β_{t+1} is predetermined: $\beta_{t+1} = \beta_t(1-g) + g\left(\frac{P_t}{D_t}\frac{D_t}{D_{t-1}}\frac{D_{t-1}}{P_{t-1}}\right)$

      ii. G_{t+1}/D_{t+1} depends on ε^D_{t+1}:

$$\frac{G_{t+1}}{D_{t+1}} = \frac{P_t/D_t - P_{t-1}/D_{t-1}D_{t-1}/D_t - \phi(\bar{\pi}_t)G_t/D_t}{\beta^D + \varepsilon^D_{t+1}}$$

      iii. P_{t+1}/D_{t+1} depends on (according to the subjective price model):

$$\frac{P_{t+1}}{D_{t+1}} = \frac{P_{t+1}}{P_t}\frac{P_t}{D_t}\frac{D_t}{D_{t+1}} = \frac{P_t}{D_t}\frac{\beta_t + u_t + \nu_{t+1} + \varepsilon^P_{t+1}}{\beta^D + \varepsilon^D_{t+1}}$$

   (b) Use the Gauss-Hermite quadrature rule extended to the multidimensional case:

      i. Let $\{q_i, \omega_i\}_{i=1}^I$ be a set of Hermite nodes and weights.

      ii. Note $\Psi(\boldsymbol{X}_{t+1}) = \Psi(\boldsymbol{X}_t, \varepsilon^D_{t+1}, u_t, \nu_{t+1}, \varepsilon^P_{t+1})$.

iii. Using the quadrature, each shock ε^{x}_{t+1} is replaced by $\sqrt{2}\sigma_x q_h$ (shocks are zero-mean Normally distributed) such that the expectation is computed as:

$$m_t \approx \mathbb{E}^{\mathcal{P}}_t(\Psi(\boldsymbol{X}_{t+1})) = \pi^{-N/2} \sum_i \sum_j \sum_k \sum_l \Psi(\boldsymbol{X}_t, q_i, q_j, q_k, q_l)\omega_i\omega_j\omega_k\omega_l$$

where N is the number of shocks.

3. Note that $m_t = m(P_t/D_t, \cdot)$ such that

$$\frac{P_t}{D_t} = \frac{\delta(1-\tau^D)\beta^D}{1-\delta\beta_t(1-(\tau^K)^{1+\xi}m(P_t/D_t,\cdot)) - \delta(\tau^K)^{1+\xi}m(P_t/D_t,\cdot)} \tag{118}$$

That is a nonlinear equation that can be solved numerically.

4. Estimate $\boldsymbol{\alpha}$ via PEA.