Capital Gains Taxation, Learning and Bubbles

PAU BELDA*

UAB · CREI · BSE

Latest version here.

Abstract

Why have there been more asset price boom-bust cycles since the 1980s despite the drop in macroeconomic risk? This paper argues that the fall of the Capital Gains Tax (CGT) has been a crucial factor. In a model of learning about prices, I show that a lower CGT make prices more responsive to changes in investors' beliefs, thereby elevating the likelihood of self-fulfilling booms and busts. The model can account for several hard-to-explain facts about the US stock market when using the observed path of tax cuts. In particular, it replicates 40% of the increase in excess volatility despite the decline in consumption growth volatility and 75% of the equity premium. Even with the drop in the safe real interest rate, the model predicts that the rise in volatility would have been largely avoided if tax cuts had not been implemented. Finally, I show that optimal policy prescribes a CGT that lean against market expectations, preventing belief-driven business cycles. Altogether, a CGT is identified as a unique macroprudential instrument that enhances the autonomy of monetary policy from financial stability considerations.

Keywords: Capital Taxation, Asset Pricing, Expectations, Macroprudential Policy, Equity Premium.
JEL codes: D83, D84, E32, E44, E62, G12, G14.

*PhD candidate, in the 2022-23 Job Market. First version: July 2020. This version: February 2023. I thank Albert Marcet for his guidance. Besides, I have benefited from comments from Klaus Adam, Ricardo Reis, Nicola Gennaioli, Dirk Krueger, Jordi Caballé, Luis E. Rojas, Wei Cui, Hugo Rodríguez, Alexander Ludwig, Abhay Abhyankar, Ramon Ruiz and participants in seminars at Universitat Autònoma de Barcelona, Barcelona School of Economics, Bocconi University, University of Oxford, University of Mannheim, Universitat de Barcelona, Universidad Carlos III, University of Bath, Universitat de les Illes Balears, 2022 Winter Meeting of the Econometric Society, SAEe2020, SAEe2022, EconMod2021 conference, ENTER Jamboree. This research has 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 usual disclaimers apply.

1.- Introduction

There was no Great Moderation in the stock or housing markets. While many macroeconomic variables exhibited lower fluctuations, the main asset markets followed the opposite path. Larger booms and busts occurred, such as the Dotcom episode in the late 1990s, the housing bubble in the early 2000s, or the post-Great Recession joint stock and housing price boom.[1]

Is the increase in price fluctuations in the context of lower macroeconomic risk consistent with the theory? Many models would contend otherwise. Following lower consumption growth volatility, models based on macroeconomic fundamentals would predict more stable prices following 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 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)). Thus, the negative covariance between macroeconomic and asset price fluctuations appears as a troubling observation.[2]

The central hypothesis of this study is that the increase in stock price volatility observed since the 1980s can be partially attributed to the reduction of the Capital Gains Tax (CGT). Indeed, substantial CGT cuts, such as Clinton's Taxpayer Relief Act (1997), preceded the emergence of boom-bust cycles since the 1980s. That pattern echoes the Roaring Twenties experience, where substantial tax cuts preceded the rise and fall of Wall Street.[3]

But why would a lower CGT boost price cycles? The dominant view would suggest that lower CGT would reduce price fluctuations through the lock-in effect, which states that a decline in taxes during a boom period would increase the supply of assets on the market and thus temper price pressures (e.g., Stiglitz (1983)).[4] Nonetheless, a CGT can also trigger a demand effect: in a boom, lower taxes would fuel expected payoffs, boosting stock demand and prices further. Through quantitative analysis, the study shows that this demand channel dominates the lock-in

[1]Understanding these events is critical since they appear to be tied to macroeconomic instability, resource misallocation, or wealth inequality (e.g., Hall (2017), Gopinath et al. (2017), Kuhn et al. (2020)).

[2]Using full-information Bayesian techniques, Chen et al. (2019) shows that long-run risks account for less than 25% of the variance of Price-Dividend ratio and that habit's contribution is negligible. The non-explained residual is particularly large since the 1990s, showing the difficulty of these models to account for the larger variance of the PD despite more stable aggregate consumption growth.

[3]In particular, the top CGT rate was cut in 1921 from 73% to 12.5%.

[4]The lock-in effect refers to asset holders not selling securities that they would have sold without taxes. The tax liability depresses the value of selling, below the value of keeping the asset in some cases, preventing portfolio rebalancing.

effect, accounting for a significant portion of the observed increase in stock price volatility.[5]

To study the effects of a CGT on price movements, I employ a model that displays price booms and busts and is consistent with microevidence about investors' expectations. In the model, price cycles are born from a feedback loop between expectations and prices that appears when agents learn about stock prices. Following Adam and Marcet (2011), learning is microfounded via an information friction: investors have imperfect knowledge about other agents' expectations. This friction prevents agents from deducing the equilibrium pricing function; instead, it forces rational investors to forecast prices using a subjective model of prices in an otherwise standard Lucas (1978) setup. In this framework, agents resort to Bayes' law to update their expectations. As documented in surveys (e.g., Greenwood and Shleifer (2014), Coibion and Gorodnichenko (2015)), investors extrapolate from past returns and underreact to news.

In this model, a CGT dampens the feedback loop that triggers these belief-driven cycles. Consider a good news shock. Investors would become more optimistic, and demand and prices would rise. However, the translation of higher optimism on a higher stock demand depends on the tax level; the higher the tax, the lower the net expected returns and the lower the increase in demand and prices. Thus, taxes reduce the response of demand and prices to a change in beliefs.[6] By Bayesian updating, a lower price increase would lead to a weaker upwards revision of beliefs and a smaller boom. In other words, higher taxes dampen the propagation of shocks such that momentum is shorter and mean reversion is faster. Altogether, a CGT reduces the beliefs elasticity of prices, anchoring beliefs around their fundamental value.

The model is estimated to replicate a list of facts about the US stock market for the 1946-2018 period, split into two halves to emphasize the changes occurring since the 1980s. On top of the increase in the Price-Dividend (PD) ratio volatility along with the fall in macroeconomic risk, the list includes a set of standard facts such as the equity premium, the procyclicality of survey expectations and the mean level of the PD ratio.

I decompose the PD ratio variance using the Campbell and Shiller (1988)'s equation extended with capital taxes. It turns out tax changes play a non-negligible role, explaining about 40% of the variance typically attributed to expected returns. According to the decomposition, the higher volatility in the PD ratio since the 1980s is the result of i) a higher discount factor, in part due to lower tax levels (10%); ii) tax changes (30%); iii) a drop in the correlation between stock returns

[5] This effect of a CGT on demand is sometimes called the *capitalization effect*. See Dai et al. (2008). While this channel is used to relate CGT to the price level, this paper explores its impact on price volatility.

[6] This mechanism is consistent with the empirical evidence in Giglio et al. (2021), which points out that the elasticity of stock holdings to beliefs decreases with taxes.

and dividend growth (60%), related to the surge in capital gains.

Computationally, the model is solved using a novel application of the Parameterized Expectations Algorithm with a theory-based approximating function that allows for closed-form solutions. Given the observed path of capital tax cuts and the empirical dividend process, I estimate the remaining structural parameters using the Simulated Method of Moments. Then, I formally test the hypothesis that model statistics differ from their empirical counterpart.

The model statistics pass most of the t-tests. For instance, the baseline model (without the lock-in effect) reproduces 60% of the increase in the unconditional variance of the PD ratio despite incorporating the observed reduction in the volatility of both dividend and consumption growth since the 1980s. It also matches the variance decomposition; the increase is driven by tax changes and the lower correlation between returns and dividend growth. The latter is related to the core model's mechanism: the decline in CGT amplify the beliefs-price spiral, boosting capital gains. Consequently, returns became more connected to capital gains and less to dividends, in line with the data. In this vein, the model predicts that if tax cuts had been avoided, the Great Moderation would have reduced stock market volatility by 20%. In other words, according to the model, tax cuts more than offset the lower macroeconomic risk.

Then, the lock-in effect is considered by allowing investors to decide the timing of capital gains realization. In particular, each investor manages a stock of unrealized capital gains G_t facing portfolio management costs in line with Gavin et al. (2015). When the realization of capital gains is deferred, the cost function penalizes investors with extra unrealized capital gains. These higher stock of gains increase the future tax liability of households. Altogether, investors face an additional trade off: realize capital gains and pay a taxes today or defer the realization to the future and have an additional tax liability in the future. With lock-in, the model produces an increase in the PD variance of about 40% of the one observed. Thus, the lock-in effect counteracts the effect of taxes through demand, but the latter dominates.[7]

In addition, the model replicates 30% of the rise in the mean PD ratio, related to the fact that lower taxes raise net asset payoffs, demand, and then prices. This result adds additional evidence to the literature highlight the negative relationship between capital taxes and stock prices such as McGrattan and Prescott (2005) or Brun and González (2017). Besides, in line with the models of learning about stock prices, the model produces the positive and high comovement between expectations and prices documented in surveys.

[7]This result is aligned with the estimations of the tax elasticity of capital gains realization estimated in the literature, that ranges from -0.3 to -0.7 (Zidar (2019), Agersnap and Zidar (2021)).

Finally, the model generates a high portion of the equity premium along with a low and stable risk-free rate, realistic consumption and dividend growth processes, a non-negative discount factor, and low risk aversion. The reason is twofold. First, the learning model generates high volatility from beliefs, increasing average returns by Jensen's inequality.[8] Additionally, the inclusion of taxes imparts a trend on the PD ratio that helps in getting high returns without exaggerating its volatility. Crucially, these two factors do not affect the risk-free rate.

An alternative explanation of the rise in volatility qualitatively suggested in the literature is the reduction in the safe real interest rate since the 1980s (e.g., Martin and Ventura (2018), Adam (2020)). I include this possibility by introducing a time-varying discount factor calibrated to replicate the risk-free rate's historical path. I show that, despite the drop in interest rates (i.e., the rise in the discount factor), the absence of tax cuts would have largely avoided the increase in volatility. Conversely, the lack of a decline in interest rates would have reduced but not entirely removed the destabilizing effects of lower taxes. This suggests an important message for monetary policy: with the appropriate CGT in place, Central Banks can use the interest rate policy to completely focus on price stability, without worrying much about its financial stability implications.

I further validate the model by testing predictions related to its main mechanism. First, I test the hypothesis that lower taxes would increase the elasticity of prices to beliefs using survey expectations. In line with the prediction, I show that this elasticity increased over the studied period and, like in the model, reacts persistently to a tax shock. Another prediction that relies just on prices and dividends is that the PD ratio would respond more to shocks in low-tax environments. Indeed, the empirical response of the PD ratio to an equivalent price growth shock is 60% larger in the period with lower taxes.

The last part of the paper digs into the normative side of capital gains taxation. In this family of models, asset markets are informationally inefficient (Adam et al. (2017)). The extra volatility arising from the learning process can be interpreted as a pecuniary externality.[9] For this pecuniary externality to have significant welfare consequences, excess volatility in asset prices is connected to aggregate consumption fluctuations. I use a tractable two-sector growth model with investment adjustment costs and learning about capital prices. The model links the capital market price to investment decisions, in line with the Q-theory (Tobin (1969)). With subjective beliefs,

[8] The fact that high beliefs volatility helps to get a high stock return was already exploited by Adam et al. (2016). This volatility coming from subjective beliefs avoids using a too volatile income process or a too high risk aversion.

[9] Since excess volatility emerges from the inability of agents to internalize the equilibrium price formation due to imperfect information about other market participants.

two feedback loops operate: the first, the one between the stock and price of capital, which is self-correcting; the second, the price-expectations loop, which is self-reinforcing. Their interaction gives rise to large and persistent cycles of over- and under-accumulation of capital.

In such a world, a Social Planner is asked to deliver the best possible competitive equilibrium by choosing a tax on unrealized capital gains and lump-sum transfers. She is endowed with all the relevant information, including investors' beliefs. The optimal policy prescribes using the CGT to counteract too optimistic/pessimistic beliefs about capital gains. In this way, the planner closes the gap between the market price and the shadow price of capital, restoring the First Best allocations.

The optimal CGT is a nonlinear function of the deviations of subjective expectations from its Rational Expectations counterpart (call it β^*). A shortcoming is that the optimal CGT is unbounded, inherits the dynamic properties of subjective beliefs and it is informationally demanding. Since tax volatility might not be desirable and mean subjective beliefs can be difficult to measure, an alternative implementation is suggested.[10] On the one hand, the CGT is set equal to 100% to eliminate the influence of subjective price beliefs (i.e., the source of the externality) on market prices. Since such a high tax depresses the market price a little, a subsidy for capital rents is introduced to restore efficient prices. The subsidy only depends on β^* and then is fairly stable and only requires a notion of fundamental value.

Related literature. The paper speaks to different literature on capital taxation, asset pricing, learning, business cycles and macroprudential policy. In what follows, I highlight the main contributions to each field.

This paper is the first to propose a theory of how a CGT regulates excess price volatility in a general equilibrium model. According to the dominant view, a CGT increases volatility due to the supply-side lock-in effect (e.g., Somers (1948), Somers (1960), Stiglitz (1983)). An alternative viewpoint contending that a CGT stabilizes prices through the demand-side capitalization effect is on Haugen and Heins (1969) and Haugen and Wichern (1973).[11] My work builds upon these last two papers; while they use exogenous beliefs, I highlight the critical role of endogenous beliefs.[12]

This novel theory is embedded in a quantitative model to show that tax cuts explain a significant portion of the rise in excess price volatility in the US. This evidence complements the model-based

[10]This implementation uses a decomposition between fundamental and non-fundamental price volatility in the spirit of the trading decomposition used by Dávila (2020).

[11]See Dai et al. (2008) for an explanation of the lock-in and capitalization effect and a literature review about their empirical relevance.

[12]In fact, applying Rational Expectations to Haugen and Heins (1969) and Haugen and Wichern (1973)'s model delivers a constant PD ratio so that taxes play no role in excess volatility whatsoever.

works by McGrattan and Prescott (2005) and Brun and González (2017) and the econometric approach of Sialm (2009), that pointed out that lower taxes can account for the rise is asset valuation levels. Besides, Dai et al. (2013) document a new statistical negative relationship between taxes and returns volatility exploiting the cross-sectional variations in accrued capital gains and dividend distributions of stocks around the CGT cuts of 1978 and 1997.[13] My paper suggests a theory that can rationalize this evidence and present time-series evidence.

The paper focuses on the excess volatility puzzle, highlighting its time-varying nature. Theoretically, the model deals with the puzzle by bringing up an additional source of variation (beliefs) in line with the Adaptive Learning literature (e.g., Timmermann (1993), Bullard and Duffy (2001), Cogley and Sargent (2008)). In particular, I follow the Internal Rationality framework, a micro-foundation of learning proposed by Adam and Marcet (2011). Adam et al. (2016) and Adam et al. (2017) presented quantitative versions, accounting for many asset pricing facts. The inclusion of taxes in this kind of models solves some of their shortcoming and allows to address new facts as the rise in excess volatility. Empirically, I present a version of the Campbell and Shiller (1988)'s PD variance decomposition with taxes. Additionally, I argue that tax cuts help explain the equity premium puzzle. Thus, the paper provides quantitative evidence backing McGrattan and Prescott (2003).

The model used for optimal policy analysis relates to papers dealing with joint stock market and business cycles (e.g., Boldrin et al. (2001)). In particular, recently some papers have considered the impact of learning about capital prices on business cycle through labour demand (Adam and Merkel (2019)), collateral constraints (Winkler (2020)) and wealth effects (Ifrim (2021)). This paper considers another possibility: learning about capital prices directly affect physical investment when there are investment adjustment costs.

While a CGT has been studied as a tool to raise revenue (e.g., Agersnap and Zidar (2021), Sarin et al. (2022)), this paper looks at it from a macroprudential standpoint. The recent literature on macroprudential policy has dealt chiefly with collateral constraints and taxes on borrowing (e.g., Lorenzoni (2008), Jeanne and Korinek (2010), Dávila and Korinek (2018), Jeanne and Korinek (2019)) and nominal rigidities (Farhi and Werning (2016)). This paper shares the emphasis on pecuniary externalities with most of the literature but departs from its origin (i.e., information rather than financial frictions).[14] Besides, while this recent literature has focused on constrained

[13]Building upon the idea that CGT are a risk-sharing device with the government that affect the level of stock returns (e.g., Lerner (1943), Stiglitz (1975), Sikes and Verrecchia (2012)), Dai et al. (2013) suggests CGT cuts reduce the risk-sharing, rising volatility. This point is comparable to that of Gemmill (1956)'s. The theory I propose can accommodate this income effect, but rely primarily on a substitution effect.

[14]Di Tella (2019) and Kurlat (2018) focus on the inefficiencies arising from financial frictions when the planner is

efficiency, Benigno et al. (2019) showed that a superior allocation is attainable with the same instruments. Following them, I study the optimal use of a CGT to restore unconstrained efficiency.

The findings of this study have significant policy implications. Firstly, it highlights the potential of a CGT to serve as a viable alternative to the widely debated Financial Transactions Tax (or "Tobin tax").[15] Secondly, the research demonstrates that, instead of relying solely on monetary policy to regulate asset prices, the implementation of an appropriate CGT can facilitate the disentanglement of interest rate policy from financial stability considerations.[16]

The rest of the paper proceeds as follows. Section 2 explores how Capital Gains Taxes can stabilize asset prices using a model with learning about prices. Section 3 presents a quantitative application of the theory to the US stock market. Section 4 derives an optimal CGT in a two-sector growth model. Section 5 concludes, pointing out some avenues for future research.

2.- Theory: a Capital Gains Tax to stabilize asset price cycles

This section is a theoretical exploration of the role of capital gains taxes in asset price cycles in an asset pricing model with Internal Rationality. It shows the main theoretical proposition of the paper, that is, the variance of the PD ratio is decreasing on the CGT level under some conditions. Section 2.1. sets up the basic model. In Section 2.2., the effects of capital gains taxes on the asset price level are analysed. Section 2.3. establishes the main proposition, relating the capital gains tax level to the PD ratio volatility.

2.1.- The model

In this section, a consumption-based asset pricing model with taxes on realized capital gains and dividends is set up. Its basic layer is the Lucas (1978)'s tree model with a general probability measure, as Adam and Marcet (2011), and capital taxes.

Demographics. The economy is populated by a unit mass of infinitely lived identical investors.

Goods and assets. There is a single perishable good in the economy. Furthermore, there exist a single risky asset in the form of a contract that delivers goods (called "dividends") each period and is marketable at an uncertain future price, giving rise to capital gains and losses.

informationally constrained. Farhi and Werning (2020) incorporates extrapolation into the optimal macroprudential analysis, showing that it plays a rather secondary role. In contrast, the analysis in the paper puts it at the forefront.

[15] See, for instance, Buss et al. (2016), Buss and Dumas (2019) or Dávila (2020) for theoretical analysis and Umlauf (1993) or Cappelletti et al. (2017) for empirical results that challenge the ability of the Tobin tax to stabilize asset prices.

[16] Asset pricing targetting has been shown to be appropriate in Nisticò (2012), Gambacorta and Signoretti (2014) or Ifrim (2021).

Resource processes. This is a pure exchange economy. When the time starts, each investor is endowed with one unit of stock ($S^i_{-1} = 1$). Dividends D_t are exogenous, following a random walk with drift process

$$\ln D_t = \ln a + \ln D_{t-1} + \ln \varepsilon^d_t \tag{1}$$

with a being the permanent component and $\varepsilon^d_t \sim \log \mathcal{N}(1, ^{s^2_d} - 1)$ an i.i.d. innovation.

Markets. Financial markets are competitive. Short selling is not allowed and there is no other form of borrowing. The goods market behaves also competitively.

Fiscal System. There is a linear tax $\tau^D \in [0, 1)$ on dividends. Capital gains are taxed in linearly at a rate $\tau^K \in [0, 1)$ on a realization basis. Every period, investors face some risk of being hit by a very bad shock. $z^i_t \sim Bernoulli(\pi)$ is a random variable indicating that possibility. If the event materializes ($z^i_t = 1$), investor i sells all her stock holdings and pay the corresponding taxes on capital gains (or, symmetrically transfers on capital losses). The probability of that catastrophic event occurring is π.[17] All the revenues (outflows), call them T_t, are transferred to (taxed from) the individuals in a lump-sum way.[18]

Investors' information set. Investors take τ^D, τ^K and T as given and know the stochastic process followed by D_t and z^i_t. Besides, they might face information limits about other investors' type. This possible friction is captured in a general way by introducing a subjective probability measure $\mathcal{P}^i$ that reflects investors' views about dividends and prices, allowing them to forecast future values of the relevant variables despite not having all the information. Thus, the underlying probability space is given by $(\Omega, \mathcal{B}, \mathcal{P}^i)$ where Ω is the state space with $\omega = \{D_t, P_t\}^\infty_{t=0}$ as a typical element, $\mathcal{B}$ denotes the σ-algebra of Borel subsets of Ω and $\mathcal{P}^i$ agent's i subjective probability measure over $(\Omega, \mathcal{B})$.

Optimizing behaviour. 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 an optimization program using a discount factor $\delta \in (0, 1)$ and her subjective probability measure $\mathcal{P}^i$, in line with the Internal Rationality framework setup by Adam and Marcet (2011). Thus, each

[17] This assumption is a shortcut that captures the lock-in effect pointed out in the literature (e.g. Somers (1960), Stiglitz (1983)) without having to deal with optimal trading strategies. It is equivalent to assume that investors anticipate that a fixed proportion of the expected capital gains are realized as in Sialm (2009). While an endogenous π will be richer, it would reduce model's tractability severely. I leave that for future research.

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

investor maximize its expected lifetime welfare

$$\max_{\{C_t^i, S_t^i, X_t^i\}_{t=0}^{\infty}} \mathbb{E}_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t U(C_t^i) \tag{2}$$

subject to the budget constraint

$$C_t^i + P_t X_t^i \leq (1-\tau^D) D_t S_{t-1}^i + z_t^i S_{t-1}^i \Big(P_t - \tau^K (P_t - P_{t-1}) \Big) + T_t \tag{3}$$

the stock holdings law of motion

$$S_t^i = (1 - z_t^i) S_{t-1}^i + X_t^i \tag{4}$$

and stock holdings bounds

$$0 \leq S_t^i \leq \bar{S}, \text{ given } S_{-1}^i = 1 \tag{5}$$

The utility function is a time-separable continuous, increasing in consumption $U'(C_t^i) > 0$ but concave $U''(C_t^i) \leq 0$ function. In this section, I assume risk-neutrality.[19] Lower and upper bounds on S_t^i are assumed for convenience; mathematically, these bounds ensure that the feasibility set is compact; economically, the lower bound rules out short-selling strategies aimed at avoiding taxes.

Model Equilibrium. The investor's program consists of maximizing a bounded continuous function over a compact non-empty feasible set.[20] By the Weierstrass extreme value theorem, these are sufficient conditions for the existence of a maximum. Moreover, the convexity of the feasible set implies the first order conditions are necessary and sufficient for the optimum by the Kuhn-Tucker (KT) theorem. Given the fiscal system, investor i's optimality conditions for an interior solution boil down to the following Euler Equation

$$P_t = \delta \mathbb{E}_t^{\mathcal{P}^i} \Big[(1-\tau^D) D_{t+1} + P_{t+1} - z_{t+1}^i \tau^K (P_{t+1} - P_t) \Big] \tag{6}$$

along with a transversality condition, and the sequence of budget constraints and market clearing conditions. As the main novelty, the expected discounted payoffs in the Euler Equation include the possibility of a bad shock occurrence next period, leading to realize gains and paying taxes. When the expectation about a catastrophe is positive, τ^K matters for equilibrium prices.[21] Given

[19] This assumption is just for simplifying the analytical derivations. Risk aversion is introduced in Section 3.

[20] See Adam et al. (2017) for a proof in a similar setup.

[21] Note that in this setup, the one-period ahead Euler Equation rather than the discounted sum of dividends is the relevant condition. The reason is that agent's i knowledge of the Law of Iterated Expectations is not useful because she ignores the probability measure of the marginal agent in future periods is an unknown $\mathcal{P}_j$, with j potentially

investors homogeneity, the equilibrium PD ratio can be derived directly from agent's i one-period ahead Euler Equation. It can be rewritten as

$$P_t = \delta(1-\tau^D)\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{D_{t+1}}{D_t}\right]D_t + \delta\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{P_{t+1}}{P_t}\right]P_t - \delta\tau^K\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{P_{t+1}}{P_t}z_{t+1}^i\right]P_t + \delta\tau^K\mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i\right]P_t \tag{7}$$

Solving the previous expression for P_t and dividing both sides by D_t, the PD ratio reads as

$$\frac{P_t}{D_t} = \frac{\delta(1-\tau^D)\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{D_{t+1}}{D_t}\right]}{1-\delta\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{P_{t+1}}{P_t}\right]+\delta\tau^K\left(\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{P_{t+1}}{P_t}z_{t+1}^i\right]-\mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i\right]\right)} \tag{8}$$

This formula can accommodate different processes for dividends and different assumptions for expectations, including Rational Expectations. Since agents are aware of the true stochastic processes for D_t and z_t^i, $\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{D_{t+1}}{D_t}\right] = \mathbb{E}_t\left[\frac{D_{t+1}}{D_t}\right] = \beta^D$ and $\mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i\right] = \pi$. To determine expectations involving future prices, I consider two possibilities.

First, suppose agents have full information, including knowledge about other investors type. This is the case of Rational Expectations. Iterating forward on the Euler Equation and applying a transversality condition, the equilibrium PD ratio is given by

$$\frac{P_t^{RE}}{D_t} = \frac{\delta(1-\tau^D)\beta^D}{1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K} \tag{9}$$

and then, $\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{P_{t+1}}{P_t}\right] = \mathbb{E}_t\left[\frac{P_{t+1}}{P_t}\right] = \beta^D$. Note that for the PD ratio to be positive $\delta\beta^D < \frac{1-\delta\pi\tau^K}{1-\pi\tau^K} > 1$, where the last inequality follows from $\delta < 1$. This is assumed throughout the paper.

Now suppose there is an information friction: agents' homogeneity is not common knowledge. Since the Law of Iterated Expectations is helpless to deduce equilibrium prices from the individual Euler Equation, agents need a model to forecast future prices (Adam and Marcet (2011)). The proposed subjective model is

$$\frac{P_t}{P_{t-1}} = b_t + \varepsilon_t^P \tag{10}$$

$$b_t = b_{t-1} + \vartheta_t \tag{11}$$

with $\varepsilon_t^P \sim i.i.\mathcal{N}(0, s_P^2)$ and $\vartheta_t \sim i.i.\mathcal{N}(0, s_b^2)$. The permanent component of price growth b_t is unobserved and has to be estimated from the history of prices. For that purpose, investors use a Kalman filter. The posterior (conditional on the observed price history) is given by $b_t|I_t \sim \mathcal{N}(\beta_t, \sigma^2)$

different from i. See Adam and Marcet (2011) for a discussion.

where σ^2 is the steady state Kalman filter uncertainty and the posterior mean evolves recursively following[22]

$$\beta_t = \beta_{t-1} + g\Big(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\Big) \tag{12}$$

where $g = \frac{\sigma^2+s_b^2}{\sigma^2+s_b^2+s_P^2}$ is the steady state Kalman gain. Hence, subjective price expectations are given by $\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}\right] = \beta_t$. The setup encompasses RE equilibrium beliefs as a special case; when $s_b^2 = 0$ and investors' initial prior is $b_0 = \beta^D$ with probability 1, $\beta_t = \beta^D \;\; \forall t$.

The last object to be determined is $\mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i \frac{P_{t+1}}{P_t}\right]$. Given the subjective model of prices,

$$\mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i \frac{P_{t+1}}{P_t}\right] = \mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i(b_t + \vartheta_{t+1} + \varepsilon_{t+1}^P)\right] = \mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i b_t\right] \tag{13}$$

Since $\mathbb{E}_t^{\mathcal{P}^i}\left[b_t\right] = \beta_t$, $\beta_t \perp\!\!\!\perp P_t$ and z_{t+1}^i is an i.i.d. Bernoulli trial, $\mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i b_t\right] = \mathbb{E}_t^{\mathcal{P}^i}\left[z_{t+1}^i\right]\mathbb{E}_t^{\mathcal{P}^i}\left[b_t\right] = \pi\beta_t$.

Plugging the subjective expectations in the pricing function (8), the equilibrium PD ratio under learning reads as

$$\frac{P_t^L}{D_t} = \frac{\delta(1-\tau^D)\beta^D}{1-\delta(1-\pi\tau^K)\beta_t - \delta\pi\tau^K} \tag{14}$$

Hence, the dynamics of the PD ratio are completely determined by $\{\beta_t\}$ and the latter is completely determined by the following 2nd order nonlinear difference equation:

$$\beta_t = \beta_{t-1}(1-g) + g\left(\frac{1-\delta(1-\pi\tau^K)\beta_{t-2} - \delta\pi\tau^K}{1-\delta(1-\pi\tau^K)\beta_{t-1} - \delta\pi\tau^K}\right)\beta^D\varepsilon_{t-1}^D \tag{15}$$

The steady state of the equation correspond to the RE expectations ($\beta_t = \beta^D$ for all t). Out of the steady state, beliefs orbit around the RE equilibrium, giving rise to beliefs booms and busts (see Adam et al. (2016)). In the next sections I show the implication of taxes for the beliefs dynamics.

2.2.- Capital Gains Taxes and the PD ratio level

In this section, the impact of taxes on the PD ratio level is analyzed under both Rational Expectations and learning. Given the pricing formula for RE (9),

$$\frac{dP_t^{RE}/D_t}{d\tau^D} = -\frac{\delta\beta^D}{1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K} < 0 \tag{16}$$

[22] As it is standard in the literature, equation (12) contains lagged price growth. The reason is that it is assumed agents observe in period t information about the lagged transitory component ε_{t-1}^P. The main advantage of this assumption is that it avoids multiplicity of equilibria. Besides, it turns out to perform better. See Adam et al. (2017) for a discussion.

and

$$\frac{dP_t^{RE}/D_t}{d\tau^K} = -\frac{\delta(1-\tau^D)\beta^D\delta\pi(\beta^D-1)}{(1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K)^2} < 0 \tag{17}$$

if $\beta^D > 1$. Thus, the model predicts a negative relationship between capital taxes and stock market valuations under RE, in line with McGrattan and Prescott (2005) or Sialm (2009).[23]

As for learning, hile the effects of τ^D on the PD ratio level are analogous to the case of Rational Expectations, τ^K has an extra effect through subjective beliefs dynamics. Thus,

$$\frac{dP_t^L/D_t}{d\tau^K} = \frac{\partial P_t^L/D_t}{\partial\tau^K} + \underbrace{\frac{\partial P_t^L/D_t}{\partial\beta_t}\frac{d\beta_t}{d\tau^K}}_{\text{Learning Amplification}} \tag{18}$$

Like under RE, $\frac{\partial P_t^L/D_t}{\partial\tau^K} < 0$ when agents expect positive capital gains (i.e., $\beta_t > 1$). Moreover,

$$\frac{d\beta_t}{d\tau^K} = \beta^D\varepsilon_{t-1}^D\frac{-\delta\pi\Delta\beta_{t-1}(1-\delta(1-\pi\tau^K)\beta_{t-1}-\delta\pi\tau^K)-\delta(1-\pi\tau^K)\Delta\beta_{t-1}\delta\pi(\beta_{t-1}-1)}{(1-\delta(1-\pi\tau^K)\beta_{t-1}-\delta\pi\tau^K)^2} \tag{19}$$

which turns out to be negative provided $\Delta\beta_{t-1} > 0$ and $\beta_{t-1} > 1$. Since $\frac{\partial P_t^L/D_t}{\partial\beta_t} > 0$, τ^K dampens the PD ratio during booms. Hence, with learning, the negative effect of τ^K on the PD ratio level is reinforced.[24]

2.3.- Capital Gains Taxes and the variance of the PD ratio

The results in the previous section provides some intuition about the role of a CGT on PD cycles. Under Rational Expectations, the PD ratio is a constant and then taxes play no role.[25] However, when there are information frictions and learning, a CGT dampens price fluctuations coming from beliefs; it depresses prices during booms and increases prices during bursts. The following proposition explores this ability formally. For notational convenience, I denote $P_t/D_t = P_t^L/D_t$ from now on.

Assumption 1. *Assume: i)* $g \in (0,\bar{g})$ *with* $\bar{g} = \frac{1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K}{\delta(1-\pi\tau^K)\beta^D}$*; ii)* $\pi\tau^K < \bar{\tau} = \frac{2\delta\beta^D-1}{2\delta\beta^D-\delta}$.

[23] This contrast with the results in Sialm (2006), according to which the tax level is irrelevant to the PD ratio level under CRRA utility. His result is driven by a consumption tax on the purchasing of new stocks. On the contrary, capital income taxes deliver results more aligned with the empirical observations.

[24] This suggests that capital gains taxes are particularly important to explain the rise in stock market valuations, complementing the role of dividend taxes pointed out by McGrattan and Prescott (2005).

[25] This is the result of the i.i.d. dividends growth assumption. In Appendix C I explore a case with persistent growth, in the spirit of Bansal and Yaron (2004). Then, a CGT affects negatively the variance of the PD ratio only by reducing ω (i.e., it has no effect on the variance of growth process that originates the movements in the PD ratio). Thus, learning is sufficient but not necessary for a CGT to reduce the PD variance; it is an amplification mechanism.

The first condition says that agents update their expectations in the direction of the difference between current price growth and expectations, but in a slow way. Thus, agents learn but the learning process is sluggish.[26] The second condition puts an upper bound on effective taxes, which is close to 1 for δ close 1, as typically assumed, and then it is a rather lax condition. Under this reasonable assumption, the following result holds:

Proposition. *Up to a linear approximation around Rational Expectations, the variance of the PD ratio is decreasing on the CGT level, that is,*

$$\frac{d\mathbb{V}ar[P_t/D_t]}{d\tau^K} < 0 \tag{20}$$

Proof. Appendix B.

The proposition shows that when subjective beliefs are close to Rational Expectations, the variance of the PD ratio is given by the sensitivity of prices to beliefs times the variance of subjective beliefs

$$\mathbb{V}ar\left[\frac{P_t}{D_t}\right] \approx \omega^2 \times \mathbb{V}ar(\beta_t) \tag{21}$$

with $\omega = \frac{\partial P_t/D_t}{\partial \beta_t}\Big|_{\beta_t=\beta^D}$. It turns out a CGT reduces both the transmission of belief fluctuations to prices (ω) and the variance of beliefs:

$$\frac{d\mathbb{V}ar[P_t/D_t]}{d\tau^K} \approx \underbrace{\frac{\partial\mathbb{V}ar[P_t/D_t]}{\partial\omega^2}}_{>0}\underbrace{\frac{d\omega^2}{d\tau^K}}_{<0} + \underbrace{\frac{\partial\mathbb{V}ar[P_t/D_t]}{\partial\mathbb{V}ar[\beta_t]}}_{>0}\underbrace{\frac{d\mathbb{V}ar[\beta_t]}{d\tau^K}}_{<0} < 0 \tag{22}$$

The proposition can be illustrated by plotting the beliefs dynamics, that fully characterize the PD dynamics, at different tax levels. The second order difference equation (15) can be represented in a two-dimensional phase diagram on the (β_{t-1}, β_t) plane (keeping the dividend shock at its mean values). Starting from Rational Expectations, figure 1 show the dynamic response of beliefs to a shock (to dividends). When a positive news shock hits the market, prices become higher than expected and investors review their beliefs upwards following Bayes' Law. If this revision is strong enough, prices would rise further feeding into even higher beliefs. Thus, for some periods, there is momentum in the sense of a rise in optimism that is self-reinforcing. At some point, prices do not

[26] A small g is consistent with the empirical evidence on underreaction pointed out by Coibion and Gorodnichenko (2015).

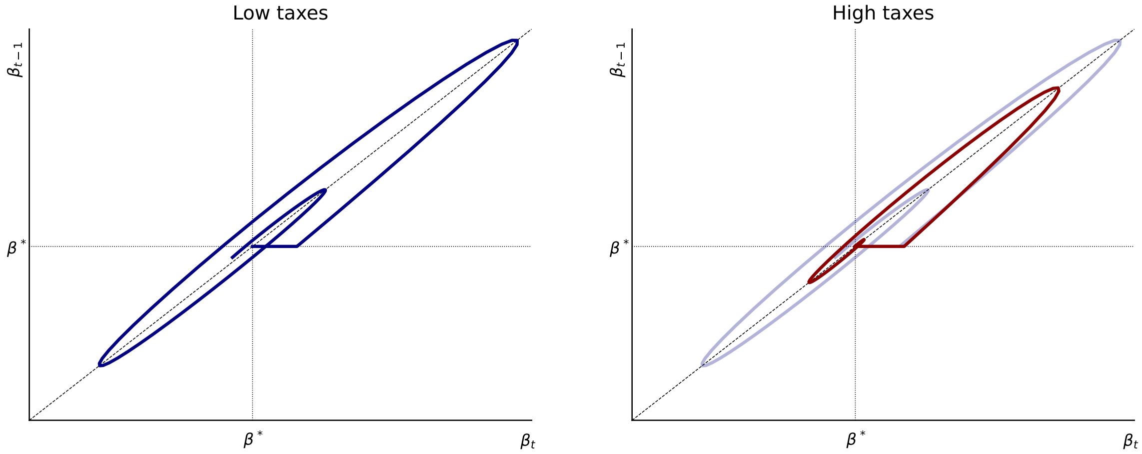

Figure 1: ***Expectations Dynamics and Capital Gains Taxes.*** *The graph illustrates the dynamics of subjective capital gains expectations (β) given by the 2nd Order Difference Equation (15). It considers $\Delta\beta_{t+1} = f(\beta_t, \beta_{t-1})$ with $\varepsilon^D_{t-1} = 1$. The left (right) panel uses lower (higher) taxes.*

grow as much as expected, so investors start correcting their beliefs downwards; this is a bust. It is through a sequence of booms and busts that beliefs revert to their fundamental value. In line with the proposition, these oscillations around the RE value are smaller the higher the CGT. In other words, momentum is shorter and mean reversion faster the higher the CGT such that a CGT anchors expectations closer to its fundamental value.

3.- Quantitative Analysis: Changes in the US Stock Market

This section uses the theory to understand important changes in the US stock market since the 1980s due to the decline in capital taxes. In Section 3.1., I document asset pricing facts for the 1946-2018 period, split into two equal halves. An adaptation of the Campbell and Shiller (1988) PD variance decomposition with capital taxes is presented. Section 3.2. offers a quantitative model, defines the equilibrium concept, and introduces a novel application of the Parameterized Expectations Algorithm to solve it. Section 3.3. describes the parameterization procedure involving structural estimation via the Simulated Method of Moments. Section 3.4. and Section 3.5. reports the estimation results and its robustness to several alternative choices. In Section 3.6., the model is extended to include alternative candidates to explain the rise in volatility. Section 3.7. focuses on why the model can generate a significant equity premium. Finally, Section 3.8. tests the primary model's mechanism using survey expectations.

3.1.- Facts

This section documents asset pricing facts using US data from 1946 to 2018. It splits the observations into two halves to highlight the changes that have occurred since the 1980s. I report statistics involving a range of capital taxes, the PD ratio, stock returns, bond returns and price, dividend, and consumption growth. I also present an extended version of the Campbell and Shiller (1988) variance decomposition that includes capital taxes.

Fact 1: Decline in capital taxes. It is well known that personal taxes on investment income went down the last decades (e.g., McGrattan and Prescott (2005), Sialm (2009)). Investment income is affected by taxes on dividends, capital gains and interests. As it is customary in the literature, I measure them using effective average marginal rates, that is, a value-weighted mean of the marginal tax rates of investors in the various income brackets once adjusting for the features of the tax code (as maximum and minimum taxes, partial inclusion of social security or phaseouts of the standard deduction).[27] Thus, the dividends tax rate is

$$\tau_t^D = \tau_t^d(1 - \eta_t) \tag{23}$$

the capital gains tax

$$\tau_t^K = (\phi\tau_t^{skg} + (1 - \phi)\tau_t^{lkg})(1 - \eta_t) \tag{24}$$

and finally, the interest tax

$$\tau_t^B = \tau_t^b(1 - \eta_t) \tag{25}$$

In the previous expressions, τ_t^d, τ_t^{skg}, τ_t^{lkg} and τ_t^b are the 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 non-taxable share. Data sources are in Appendix A; computation details on the non-taxable share are in Appendix C. As illustrated in figure 2, taxes exhibited a substantial decline which, although with different timing, represented a movement towards a general lower tax environment. This overall tax decline was the result of the joint action of tax reforms along with regulatory changes involving pensions savings vehicles that led to a massive change in asset holdings from taxable to non-taxable accounts (see McGrattan and

[27] 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 are taken from Sialm (2009). See Appendix A for details.

Prescott (2005)).[28,29]

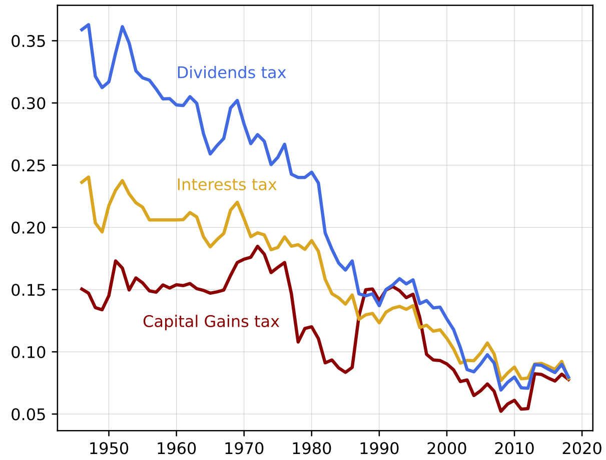

Figure 2: ***Capital taxes rates along the postwar period****. The graph plots the capital taxes on dividends (blue), interest income (yellow) and capital gains (red) as defined by equation 23, 24 and 25. Annual series 1946-2018. See Appendix A for data sources and Appendix C for details on the computations.*

Fact 2: Rise in stock valuations. Aggregate stock market valuation, measured by the PD ratio, has skyrocketed, almost doubling its mean level. 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 3. The accounting reason is that the increase in price growth (from a quarter average of 0.48% to 1.48%) has exceed by far a slightly higher dividend growth (from 0.49% to 0.75%). Thus, a higher PD ratio is a result of the sharp rise of capital gains. A related observation is that mean returns have mildly decreased, giving rise to some reduction in the equity premium.

Fact 3: Rise in the volatility of the PD ratio.[30] The rise in levels have gone hand in hand with larger fluctuations of the PD ratio. Indeed, the PD ratio standard deviation turns out to be two and half times higher after 1982 than before.[31] To understand the variation in the PD ratio, the literature has resorted to a so-called "dynamic accounting equation", first derived by Campbell

[28] Important reforms were the reduction of capital gains by Carter in 1978 and Clinton in 1997, partially counteracted by Reagan in 1986. When it comes to dividends, Reagan 1982 and Bush 2001 and 2003 represented substantial tax cuts. See xxxx for a history of tax reforms in the US.

[29] According to my estimates, the share of equity income paying taxes drop from 87% in 1946 to just 30% in 2018. This sharp decline is in line with the literature estimations (McGrattan and Prescott (2005), Sialm (2009), Rosenthal and Austin (2016))

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

[31] In log terms, it doubled from 0.07 to 0.14. In turn, the PD level went up by a factor of 1.2. Without logs, it went from 6.48 to 16.43 (x2.5 times) while the level went from 25.48 to 47.09 (x 1.8 times).

and Shiller (1988) to analyze the PD variations. The CS Equation (after Campbell-Shiller) reads as follows:

$$p_t - d_t \approx \text{constant} + \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \tag{26}$$

where lower case letters mean log-variables ($x_t = lnX_t$), $\rho = PD/(1+PD)$, with PD being the mean PD ratio in the sample. Thus, they point out that the log PD ratio is approximately equal to the difference between the discounted sum of future dividend growth and stock returns escalated by a constant, as an accounting fact. However, this dynamic equation misses the role of taxes in shaping returns. A version of the CS Equation with taxes is derived to overcome this absence. It decomposes the contribution of pre-tax returns between after-tax returns and taxes. Here I sketch the derivation. First, start with this identity

$$1 = \hat{r}_{t+1}^{-1} \hat{r}_{t+1} \tag{27}$$

with

$$\hat{r}_{t+1} = (1 - \pi\tau_{t+1}^K)\Big(\frac{P_{t+1} - P_t}{P_t}\Big) + (1 - \tau_{t+1}^D)\frac{D_{t+1}}{P_t} \tag{28}$$

being the after-tax net stock return. Manipulating the identity a bit, it becomes

$$\frac{P_t}{D_t} = \tilde{R}_{t+1}^{-1} \frac{D_{t+1}}{D_t}\Big((1 - \pi\tau_{t+1}^K)\Big(\frac{P_{t+1}}{D_{t+1}}\Big) + 1 - \tau_{t+1}^D\Big) \tag{29}$$

with $\tilde{R}_{t+1} = 1 + \hat{r}_{t+1} - \pi\tau_{t+1}^K$ as the after-tax gross return. Now, log-linearize it such that

$$p_t - d_t = -\tilde{r}_{t+1}^{-1} + \Delta d_{t+1} + \ln\Big(e^{ln(1-\pi\tau_{t+1}^K)} e^{p_{t+1} - d_{t+1}} + e^{ln(1-\pi\tau_{t+1}^D)}\Big) \tag{30}$$

Approximating the last object in the right hand side with a first order Taylor polynomial, iterating forward and imposing a transversality condition, a version of the CS Equation with taxes is obtained:

$$\begin{aligned} p_t - d_t \approx &\text{constant} + \overbrace{\sum_{j=1}^{\infty} \rho^{j-1} (\Delta d_{t+j})}^{\equiv \bar{d}_t} - \overbrace{\sum_{j=1}^{\infty} \rho^{j-1} (\tilde{r}_{t+j})}^{\equiv \bar{r}_t} \\ &+ \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \rho \ln(1 - \pi\tau_{t+j}^K)}_{\equiv \bar{\tau}_t^K} + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \tilde{\rho} \ln(1 - \tau_{t+j}^D)}_{\equiv \bar{\tau}_t^D} \end{aligned} \tag{31}$$

with $\rho = \frac{(1-\pi\tau^K)\frac{P}{D}}{(1-\pi\tau^K)\frac{P}{D}+1-\tau^D}$ and $\tilde{\rho} = \rho\frac{1-\tau^D}{(1-\pi\tau^K)\frac{P}{D}}$, all the variables being evaluated at their means. Thus, a high PD ratio must come from either higher dividends, lower after-tax returns or lower taxes in the future. Following a standard computation, the variance of the log PD ratio can expressed as follows:

$$\begin{aligned}\mathbb{V}ar(p_t - d_t) \approx \quad & \mathbb{C}ov\Big(p_t - d_t, \bar{d}_t\Big) - \mathbb{C}ov\Big(p_t - d_t, \bar{r}_t\Big) \\ & + \mathbb{C}ov\Big(p_t - d_t, \bar{\tau}_t{}^K\Big) + \mathbb{C}ov\Big(p_t - d_t, \bar{\tau}_t{}^D\Big)\end{aligned} \tag{32}$$

Thus, this version enriches the CS Equation: the PD ratio would fluctuate because changes not only in expected dividends or returns (first line) but also in expected capital taxes (second line). Furthermore, tax changes influence the discount factor ρ.

The decomposition shows an important part of the PD ratio fluctuations typically attributed to movements in discount rates seems related to changes in capital taxes instead. Thus, the share of returns in total PD variance falls from 98.85% in the tax-free version to just 53.69%, with taxes accounting for the difference.[32]

This decomposition is helpful to figure out the sources of the increase in the PD ratio volatility. Basically, it is due to three factors. On the one hand, the covariance between future dividend growth and the PD ratio moved from negative to positive territory (in terms of correlations, from -0.56 to 0.83). This signals that returns and dividend growth varied in opposite directions since the 1980s. The second factor is the change in capital taxes, mostly CGT.[33] Finally, a part of these effects is driven by a greater discount factor due, in part, to lower taxes.[34] These statistics are also collected in table 2, documenting Fact 3.

The increase in the volatility of the PD ratio and the almost constant returns volatility contrast with the decline in dividend and consumption growth risk. As it is well-known, since the 1980s, the macroeconomic risk went down, with dividend and consumption growth volatility declining by 20% and 46%, respectively.

Relation to the theory. Can the theory exposed in Section 2 help explain these facts? Tax cuts would increase net payoffs and then prices, for a given level of dividends. Moreover, tax cuts

[32]To empirically implement the previous equation one has to deal with infinite sums, which are not observable. For that end, I follow the VAR approach first outlined by Campbell and Shiller (1988) with short-run restrictions, with the variable ordering being dividends, taxes, returns and the PD ratio.

[33]Notice it does not mean that investors anticipated tax cuts necessarily. The present value of taxes is a projection from past data through the lens of the VAR. It turns out the correlation between the present value of taxes at time t $\bar{\tau^K}_t$ and time t taxes τ_t^K is -0.91 and -0.92 for each sample respectively; for dividends, it is -0.99 in both samples. In other words, the present value of taxes approximately follows the path of current taxes.

[34]Keeping ρ at their 1946-1982 value, the PD variance for the 1982-2018 period would be down by about 10%.

Table 1: ***Variance Decomposition of the Price-Dividend ratio.*** *The table reports* $\mathbb{C}ov\Big(p_t - d_t, \bar{x}_t\Big)$ *with* $\bar{x}$ *being the present value of dividend growth, stock returns, a capital gains tax factor and a dividend tax factor as specified in equation (32). The smaller gray values shows* $\frac{\mathbb{C}ov(p_t - d_t, \bar{x}_t)}{\mathbb{V}ar(p_t - d_t)} \times 100$ *for the same variables. Present values are computed using a VAR, estimated separately for each subsample; see the main text for more details.*

	1946-2018		1946-1982		1982-2018	
Returns	-18.89	-10.26	-8.53	-9.24	-13.18	-9.68
	98.85%	53.69%	118.54%	129.33%	93.67%	69.30%
Dividend growth	1.58	2.75	-1.01	-1.98	1.65	2.37
	8.27%	14.39%	-13.98%	-27.77%	11.73%	16.98%
Capital Gains tax	-	3.7	-	-0.29	-	1.73
	-	19.36%	-	-4.09%	-	12.39%
Dividend tax	-	3.92	-	0.36	-	1.00
	-	20.51%	-	5.00%	-	7.15%
Total Approximation	20.47	20.63	7.52	7.32	14.83	14.79
	107.12%	107.95%	104.56%	102.47%	105.40%	105.83%
$\mathbb{V}ar(p_t - d_t)$	19.11		7.15		13.98	
Discount factor ρ	0.9784		0.9726		0.9815	

would have increased the influence of beliefs on prices. This would increase capital gains and their influence on stock returns, reducing the correlation between returns and dividends. In addition, more influence of beliefs changes on prices would boost volatility, counteracting the reduction in macroeconomic risk. The next section explores whether this hypothesis can quantitatively account for the facts.

3.2.- Extended model and a new solution algorithm

This section extends the model set up in Section 2 to better equip it to replicate the stylized facts reported in the previous section. Moreover, it defines the concept of equilibrium and introduces a new algorithm to solve the model based on the Parameterized Expectations Algorithm.

The model is modified along four dimensions. First, the assumption of risk neutrality is abandoned. In this version, investors are allowed to dislike risk in a Constant Relative Risk Aversion (CRRA) sense, with γ regulating its risk aversion level. Second, an additional source of exogenous income is introduced, to avoid a too high correlation between dividends and consumption at odds

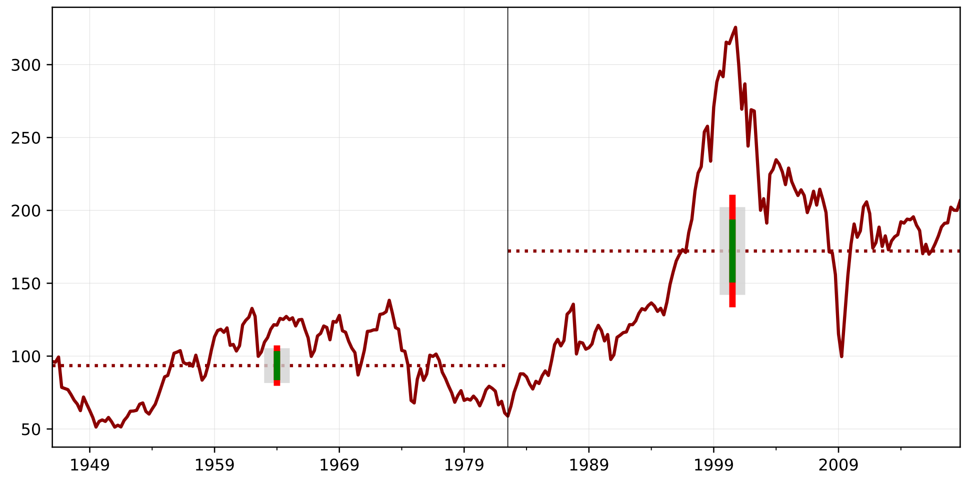

Figure 3: ***Change in the mean and standard deviation of the Price/Dividend ratio.*** *The graph plots the evolution of the PD ratio in the 1946:I-2018:IV period. The dotted lines plot the mean of each subperiod 1946:I-1982:II and 1982:III - 2018:IV. The gray box depicts the standard deviation. The green (red) bar shows the lower (upper) bound of the standard deviation confidence interval. These confidence intervals are computed using Newey-West standard errors.*

with the data. Following Adam et al. (2017), it is assumed agents get a wage endowment W_t each period, following this process:

$$\ln\left(1+\frac{W_t}{D_t}\right) = (1-p)\ln(1+\rho) + p\ln\left(1+\frac{W_{t-1}}{D_{t-1}}\right) + \ln\varepsilon_t^w \tag{33}$$

D_t are aggregate dividends, $1+\rho$ is the average wage-dividend ratio and $p \in [0,1)$ its quarterly persistence. The innovations are jointly distributed with dividend shocks following

$$\begin{pmatrix} \ln\varepsilon_t^D \\ \ln\varepsilon_t^W \end{pmatrix} \sim ii\mathcal{N}\left(-\frac{1}{2}\begin{pmatrix} \sigma_D^2 \\ \sigma_W^2 \end{pmatrix}, \begin{pmatrix} \sigma_D^2 & \sigma_{DW} \\ \sigma_{DW} & \sigma_W^2 \end{pmatrix} \right)$$

Third, stochastic taxes are introduced. It is assumed that every of the taxes on investment income follows a unit root process, that is

$$\tau_t^j = \tau_{t-1}^j + \epsilon_t^{\tau^j} \qquad \text{for j = K, D, B} \tag{34}$$

where $\epsilon_t^\tau \sim \; ii\mathcal{N}(0, s_\tau^2)$.[35] Tax shocks are assumed to be orthogonal to dividend and consumption shocks.

[35] When the observed tax time series is 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 a Gaussian white noise. Normality has been tested via the Shapiro-Wilk Normality test.

Table 2: ***Facts. US Stock Market changes: 1946-1982 vs. 1982-2018.*** *This table reports U.S. stock market moments using the data sources described in Appendix A. Growth rates and returns are annualized.*

		1946-1982	1982-2018
		Fact 1: Decline in capital taxes	
Capital Gains tax	$\mathbb{E}(\tau_t^K)$	0.15	0.09
Dividends tax	$\mathbb{E}(\tau_t^D)$	0.29	0.12
Received Interest tax	$\mathbb{E}(\tau_t^B)$	0.20	0.11
		Fact 2: Rise in asset price levels	
PD level	$\mathbb{E}(PD_t)$	25.48	47.09
Dividend growth	$\mathbb{E}(D_t/D_{t-1}-1)$	0.49	0.75
Stock price growth	$\mathbb{E}(P_t/P_{t-1}-1)$	0.48	1.84
Quarterly real bond returns	$\mathbb{E}(r_t^b)$	0.42	0.38
Quarterly real stock returns	$\mathbb{E}(r_t^s)$	4.73	4.34
		Fact 3: Rise in PD volatility	
PD volatility	$\mathbb{V}ar(p_t-d_t)$	7.15	13.98
Comovement PD - dividends	$\mathbb{C}ov(p_t-d_t,\bar{d}_t)$	-1.98	2.37
Comovement PD - returns	$\mathbb{C}ov(p_t-d_t,\bar{r}_t)$	-9.24	-9.68
Comovement PD - Capital Gains tax	$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^K)$	-0.29	1.73
Comovement PD - Dividends tax	$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^D)$	0.36	1.00
Stock returns volatility	$\sigma(r_t^s)$	7.87	7.41
Dividend growth volatility	$\sigma(D_t/D_{t-1}-1)$	2.52	1.97
Consumption growth volatility	$\sigma(C_t/C_{t-1}-1)$	1.13	0.60

Finally, risk-free bonds and taxes on bond interest are introduced in order to deal with the equity premium. The informational assumptions are: i) agents know the fundamental processes (dividends, wages, taxes) but ii) investors' homogeneity is not common knowledge. Then, prices cannot be deduced from individual optimal computations and investors use the subjective price model determined by equation (10) and (11), with the updating equation (12).

Competitive Equilibrium. Given initial endowments $S_{-1}^i = 1$, and the probability measure $\{\mathcal{P}_i\}_{i=1}^I$ (involving the dividends, wage and tax processes given by expressions (1), (33), (34)), the price model (10) and the updating equation (12)), a Competitive Equilibrium consists of sequences of allocations $\{\{C_t^i, S_t^i, B_t^i\}_{t=0}^\infty\}_{i=1}^I$ and prices $\{P_t, r_t^b\}_{t=0}^\infty$ such that:

1. Investors behave optimally, satisfying:

   (a) KKT First Order Conditions. They boil down to the following Euler Equations[36]

$$(C_t^i)^{-\gamma} = \delta \mathbb{E}_t^{\mathcal{P}_i}\left([P_{t+1} + D_{t+1} - \tau_{t+1}^D D_{t+1} - \pi \tau_{t+1}^K (P_{t+1} - P_t)] P_t^{-1} (C_{t+1}^i)^{-\gamma} \right) \quad (35)$$

[36] Since Inada conditions hold, we can ignore consumption lower corner. By concavity, the budget constraint will always bind. Assets lower and upper bounds are large enough to never bind.

$$(C_t^i)^{-\gamma} = \delta(1 + (1 - \tau_t^B) r_t^b)\mathbb{E}_t^{\mathcal{P}_i}\left((C_{t+1}^i)^{-\gamma}\right) \tag{36}$$

and the budget constraint:

$$C_t^i + P_t S_t^i + B_t^i \leq W_t + (P_t + D_t) S_{t-1}^i + (1 + (1 - \tau_t^B) r_{t-1}^b) B_{t-1}^i + T_t - (\tau_t^D D_t + \pi \tau_t^K (P_t - P_{t-1})) S_{t-1}^i \tag{37}$$

(b) A transversality condition:

$$\lim_{j \to \infty} \left(\frac{\delta}{1 - \delta \pi \tau_t^K}\right)^j \mathbb{E}_t^{\mathcal{P}_i}\left[\frac{C_{t+j}^i}{C_t^i}(1 - \pi \tau_{t+j}^K) P_{t+j}\right] = 0 \tag{38}$$

2. Markets clear:

$$\text{Equities:} \quad \int_0^1 S_t^i di = 1 \tag{39}$$

$$\text{Bonds:} \quad \int_0^1 B_t^i di = 0 \tag{40}$$

$$\text{Goods:} \quad \int_0^1 C_t^i di = D_t + W_t \tag{41}$$

State variables. The state space is made of income sources and taxes/transfers, previous stock holdings and current aggregate stock supply and, due to information incompleteness, current price and price growth beliefs, that is, $\boldsymbol{X}_t = (D_t, W_t, \tau_t, T_t, S_{t-1}, P_t, \beta_t)$.[37] Given the homogeneity property of the CRRA function, the state vector can be reduced to $\boldsymbol{X}_t = (\frac{W_t}{D_t}, \tau_t, T_t, S_{t-1}, \frac{P_t}{D_t}, \beta_t)$. In this way, the model is rewritten in terms of non-explosive ratios that allows it to be solved.

Recursive Solution via the Parameterized Expectations Algorithm. A recursive solution boils down to a time-invariant stock demand function $S_t = S(\boldsymbol{X}_t)$.[38] The main difficulty to derive such invariant function is that optimality conditions includes an unknown subjective conditional expectation. Thus, the optimal Consumption-Dividends ratio must satisfy

$$\frac{C_t}{D_t} = \left\{\delta \mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{P_{t+1}}{D_{t+1}} + 1 - \tau_{t+1}^D - \pi \tau_{t+1}^K\left(\frac{P_{t+1}}{D_{t+1}} - \frac{P_t}{D_t}\frac{D_t}{D_{t+1}}\right)\right)\frac{D_t}{P_t}\left(\frac{C_{t+1}}{D_{t+1}}\right)^{-\gamma}\left(\frac{D_{t+1}}{D_t}\right)^{1-\gamma}\right]\right\}^{-1/\gamma} \tag{42}$$

[37] From now on, I disregard the use of the superindex i to flag individual control variables to save notation.

[38] See Adam et al. (2017) for a proof of existence of a recursive equilibrium in the same model without taxes. It continuous to hold with taxes.

In equilibrium, the previous conditional expectation is a function $\mathcal{E}$ of the states, hence

$$\frac{C_t}{D_t} = \left(\delta\mathcal{E}(\boldsymbol{X}_t)\right)^{-\frac{1}{\gamma}} \equiv \bar{\mathcal{E}}(\boldsymbol{X}_t) \tag{43}$$

To solve the model, $\bar{\mathcal{E}}(\boldsymbol{X}_t)$ must be computed somehow. The Parameterized Expectations Algorithm (PEA) is one of the alternatives.[39] PEA consists of replacing the conditional expectation $\mathcal{E}(\boldsymbol{X}_t)$ by some parametric function ψ. ψ is not unique; popular possibilities are polynomials, splines or neural networks. In this model, there is no practical difference between approximating the conditional expectation $\mathcal{E}(\boldsymbol{X}_t)$ and approximating the policy function $\bar{\mathcal{E}}(\boldsymbol{X}_t)$. Exploiting that, I propose an approximating function rooted in economic theory. The idea is that of homotopy: start with a version of the model that has analytical solution (e.g., a simpler version of the model with Rational Expectations) and keep the structure of the policy function as an approximating function. See Appendix D for a detailed description. In particular, I propose the following ψ:

$$\frac{C_t^*}{D_t} = \bar{\mathcal{E}}(\boldsymbol{X}_t) \approx \psi(\boldsymbol{X}_t;\chi) = c_t Z_t \tag{44}$$

where $c_t \equiv (1-\chi\delta(1-\pi\tau_t^K)\beta_t)$ is the time-varying propensity to consume, Z_t stands for agent's current resources (i.e., the right hand side of the budget constraint) and χ is a parameter of ψ to be estimated. Then, the stock policy can be obtained using the budget constraint

$$S_t^* = S(\boldsymbol{X_t}) \approx (1-c_t)Z_t\frac{D_t}{P_t} \tag{45}$$

The stock policy indicates that investors save a time-varying fraction of their current wealth, driven by discounted expectations. Thus, a rise in optimism would increase demand, but such an increase would be lower the higher the tax. As a result, the magnitude of the price change decrease with the tax level. This is an illustration of the *Proposition 2* that explicitly uses investors' demand as mediator. Figure 4 illustrates these mechanics.

Finally, equilibrium prices can be obtained using the equity market clearing condition:[40]

$$\frac{P_t}{D_t} = \frac{(1-c_t)\left(\frac{W_t}{D_t} + 1 - \tau_t^D + \pi\tau_t^K\frac{P_{t-1}}{D_{t-1}}\frac{D_{t-1}}{D_t}\right)}{1-(1-c_t)(1-\pi\tau_t^K)} \tag{46}$$

[39]The first use of this approach was due to Wright and Williams (1982a), Wright and Williams (1982b), Wright and Williams (1984). My application builds on the version outlined by Marcet (1988).

[40]It assumes there are no tax rebates, that are included in in the robustness analysis.

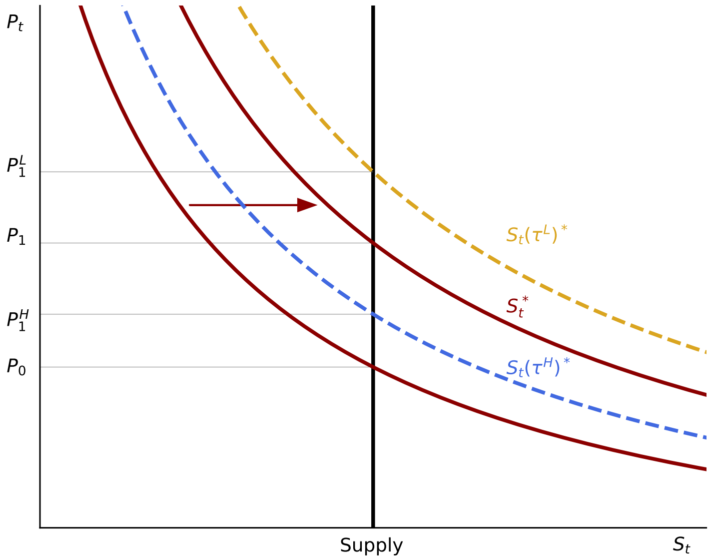

Figure 4: ***Response of stock demand to an increase in optimism at different tax levels.*** *The graph plots the stock policy function (equation (45)) keeping everything constant except prices. Then, as β_t^P increases, the curve moves rightwards. This displacement is shown at three different tax levels: low (blue), moderate (the baseline, in red) and high (yellow).*

The use of this simple approximating function has some advantages. First, we are left with a single parameter to estimate as opposed to the potentially large number of parameters of alternative approximating functions. As a result, multicollinearity problems typically associated with PEA are avoided.[41] Moreover, the procedure delivers a closed-form solution for equilibrium prices. Of course, a potential cost is that the function is quite rigid; however, it turns out to perform very well, with Euler Equation errors equivalent to $1 out of a million. See Appendix D for a detailed explanation of the algorithm and its accuracy.

3.3.- SMM estimation

This section explains the simulation strategy. It has to deal with two issues: a discontinuity in the pricing formula and the parameterization of the model. The former is solved by introducing a projection facility. The model is parameterized following a mix strategy, with some parameters calibrated from the US data and the rest being estimated via the Simulated Method of Moments.

As is standard in the learning literature, I employ a projection facility that restricts beliefs to ensure non-negative and non-explosive prices. Following Adam et al. (2016), the projection facility

[41] Although that can also be solved in other ways (e.g., using Chebyshev polynomials. See Christiano and Fisher (2000)).

starts to dampen belief coefficients that imply a PD ratio equal to PD^L and sets an effective upper bound at PD^U. It can be understood as an approximate implementation of a Bayesian updating scheme where agents have a truncated prior that puts probability zero on beliefs that imply a too high PD ratio. Appendix D contains the details.

On the other hand, the parameterization strategy is twofold. The model has a total of 12 parameters; a subset of them is picked from US data and the rest is estimated. Specifically, the vector $\widetilde{\theta} = \{\beta^D, \sigma_D, \sigma_W, \sigma_{WD}, p, \pi\}$ is picked directly from US data. I calibrate $\beta^D, \sigma_D, \sigma_W, \sigma_{WD}$ distinguishing between the two studied subperiods such that the reduction in macroeconomic volatility is captured. Parameter values are specified in the panel a) of Table 3 and data sources are reported in Appendix A. Here I only specify the strategy to calibrate π, which is not directly observable. To find empirical counterparts, I decompose it into three ratios:

$$\pi \equiv \frac{\text{Equity RKG}}{\text{Equity KG}} = \frac{\text{Equity RKG}}{\text{RKG}} \frac{\text{RKG}}{\text{KG}} \frac{\text{KG}}{\text{Equity KG}} \tag{47}$$

where (R)KG stands for (Realized) Capital Gains. The composition of both realized capital gains and total capital gains can be found in the SOI Tax Stats and the US Financial Accounts, respectively.[42] Ratios are employed because capital gains from different sources are not directly comparable and does not have the same time coverage. On average, 34% of capital gains came from equities; only about 10% of total capital gains were realized; and 27% of realized capital gains resulted from selling stocks. Altogether, the estimated π is equal to 8%, that is, only 8% of equity capital gains were realized and then, taxable.[43,44]

The remaining parameters, collected in the vector $\theta = \{\delta, g, \gamma, \rho, PD^L, PD^U\}$, are estimated via an extension of the Simulated Method of Moments.[45] Aiming at testing the power of taxes to explain the various observed changes, estimated parameters are kept fixed throughout the sample. Hence, a total of n=7 parameters are estimated to match a subset of M moments from the ones reported in table 2. In the baseline specification, I target M=18 moments and M=20 when including the risk-free rate. The vector θ is chosen as to minimize the distance between model $\widetilde{\mathcal{S}}(\theta)$ and data

[42] The time coverage is unequal. IRS data only covers the 1997-2012 period along with the year 1985, whereas the US Financial Accounts cover the whole postwar period. See Section 4.5.for an alternative calibration.

[43] Hence, after taking into account all deductions, the movement towards non-taxable accounts and the fact that only a small fraction gets realized, the average marginal tax on capital gains moved from a maximum of 1.48% in 1972 to a minimum of 0.42% in 2008, far away from statutory rates as high as 40% and never lower than 15% for the top brackets.

[44] This fact suggests that studies assuming that capital gains are taxed on accrual without adjusting for the fact that only a small fraction of gains are realized (e.g., Gourio and Miao (2011), Anagnostopoulos et al. (2012) or Brun and González (2017)) would heavily overstate the effects of capital gains taxes.

[45] I includes functions of moments, instead of pure moments. Then, the procedure is adapted via the delta method. See Adam et al. (2016).

$\hat{\mathcal{S}}$ statistics, that is,

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[\hat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta)\right]' \hat{\Sigma}_{\mathcal{S}}^{-1} \left[\hat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta)\right] \tag{48}$$

where $\hat{\Sigma}_{\mathcal{S}}$ is the weighting matrix, which determines the relative importance of each statistic deviations from its target. A diagonal weighting matrix whose diagonal is composed of the inverse of the estimated variances of the data statistics is used. Model-implied statistics are generated through a Montecarlo experiment with 1000 realizations. I formally test the hypothesis that any of the individual model statistics differ from its empirical counterpart. Finally, the model is fed with the empirical time series for capital taxes and dividend growth.[46]

*Table 3: **Calibrated parameters.** This table reports the values of the parameters calibrated directly from US data, using various data sources specified in Appendix A.*

Calibrated parameters		1946-1982	1982-2018
Mean dividend growth	β^D	1.0049	1.0075
Dividends growth standard deviation	σ_D	0.0252	0.0197
Wage shocks standard deviation	σ_W	0.0261	0.0196
Covariance (wage, dividend)	σ_{WD}	-0.0006	-0.0004
Persistence wage-dividend ratio	p	0.99	
Fraction of realized capital gains	π	0.08	

3.4.- Estimation results

In this section, the estimation results are reported. Table 5 contains the statistics from the US data and the baseline estimation. The upper (lower) part details the statistics (non) included in the SMM estimation. Standard errors and t-statistics testing the null of equality between data and model statistics are also shown. The estimated parameter vector $\hat{\theta}$ is reported in the table 4.

The model statistics perform quantitatively well, passing most of the t-tests. The model reproduces 60% of the increase in PD volatility despite the reduction in the standard deviation of both dividends and wage shocks. Besides, it also matches decently the variance decomposition; as in the data, the increase in the variance is driven by the change to procyclical dividend growth and tax changes rather than by more volatile returns. Results show a good fit for the other statistics. The model generates 1/3 of the increase in the PD level, a high positive correlation between beliefs

[46] In this way the simulated series can potentially exhibit the same trajectories and trends as the observed ones. Instead, if I use a constant tax on each subsample, trends appearing in the PD ratio would not show up, distorting then the comparison between the observed and simulated data. In other words, by introducing the empirical time series, I compute the possible transition from a high to low taxes, instead of just simulate two long lasting regimes.

Table 4: ***Estimated parameters.*** *This table reports the vector of estimated parameters $\hat{\theta}$ for different specifications of the model.*

SMM estimated parameters $\hat{\theta}$		Baseline	With r_t^b	Tax Rebates	Tax Foresight	Change in π
Discount factor	δ	1.00	1.00	1.00	1.00	1.00
Kalman gain	g	0.0261	0.0269	0.0300	0.0220	0.0274
Risk aversion	γ	1.29	0.70	0.53	1.13	1.04
Projection facility start	PD^L	390.17	539.04	467.78	358.35	372.02
Projection facility upper bound	PD^U	523.05	753.84	538.08	610.36	568.29
Average wage-dividend ratio	ρ	4.74	4.76	5.92	5.40	5.20

and the PD ratio and 3/4 of the equity premium, although the risk-free is not included in the estimation.

Results are very similar when the risk-free rate is included in the estimation, as reported in table 6. The main difference is that matching the risk-free rate requires lower risk-aversion, which would reduce stock returns, other things equal. To avoid it, the algorithm estimates a higher gain and larger bounds for the projection facility such that volatility increases. In other words, matching the risk-free rate is compatible with not lowering stock returns too much by generating extra volatility through subjective beliefs. This issue is discussed further in Section 3.7..

Table 7 reports the statistics using the baseline parameterization in the same model but with Rational Expectations. This version fails in many dimensions. As already pointed out by Adam et al. (2017), RE are capable to generate neither enough volatility nor the equity premium. Moreover, RE fails to generate a sufficient increase in the volatility of the PD ratio. This last fact points out that the effect of a CGT on the feedback loop between prices and beliefs is crucial to explain the observed patterns. Under RE, taxes reduce the influence of expectations on prices but, without the feedback from prices to beliefs, this effect is quantitatively too weak.

3.5.- Robustness tests

This section explores the robustness of the results along some dimensions. On the one hand, I try alternative choices for taxes, introducing tax rebates and tax foresight. Moreover, I estimate a version in which the decline in CGT is partially offset by an increase in the realization of capital gains (captured by π) in the second half of the sample.

The first four columns of table 8 show the model statistics when agents get back the full amount paid in taxes. In this case, only a substitution effect is operating. By and large, the model displays

*Table 5: **SMM baseline estimation results.** This table reports the data and model moments, together with standard errors and t-statistics. The first four columns report the observed statistics along with their Newey-West standard error for the US data. The next four columns reports model-implied statistics and its t-statistics. The model uses the parameterization described in Table 4. Rates of growth, variance and convariance have been multiplied by 100. The top panel reports moments included in the SMM estimation; the bottom panel non-included moments.*

	US data				Model			
	1946-1982		1982-2018		1946-1982		1982-2018	
	$\hat{\mathcal{S}}_i$	$\hat{\sigma}_{\hat{\mathcal{S}}_i}$	$\hat{\mathcal{S}}_i$	$\hat{\sigma}_{\hat{\mathcal{S}}_i}$	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat
Included in SMM estimation								
$\mathbb{E}(PD_t)$	25.48	1.55	47.09	4.04	27.48	-1.28	34.06	3.23
$\mathbb{E}(P_t/P_{t-1}-1)$	0.48	0.64	1.84	0.63	0.62	-0.22	1.28	0.89
$\mathbb{E}(r_t^s)$	4.73	0.76	4.33	0.76	4.62	0.13	4.64	-0.40
$\mathbb{V}ar(p_t-d_t)$	7.15	1.35	13.98	3.56	6.09	0.78	10.08	1.09
$\mathbb{C}ov(p_t-d_t,\bar{d}_t)$	-1.98	0.60	2.37	0.56	-0.26	-2.87	2.04	0.59
$\mathbb{C}ov(p_t-d_t,\bar{r}_t)$	-9.24	1.51	-9.68	3.04	-6.84	-1.59	-7.63	-0.68
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^K)$	-0.29	0.30	1.73	0.31	-0.19	-0.36	0.84	2.89
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^D)$	0.36	0.33	1.00	0.33	-0.18	1.61	0.56	1.36
Non-included in SMM estimation								
$\mathbb{E}(D_t/D_{t-1}-1)$	0.49	0.35	0.75	0.34	0.49	0.00	0.75	0.00
$\sigma(D_t/D_{t-1})$	2.52	0.46	1.97	0.32	2.52	0.00	1.97	0.00
$\mathbb{E}(r_t^b)$	0.42	0.02	0.38	0.03	0.50	-5.42	1.17	-25.23
$\sigma(r_t^s)$	7.87	0.72	7.41	0.81	5.28	3.58	6.72	0.85
$corr(PD_t, PD_{t-1})$	0.96	0.13	0.98	0.07	0.98	-0.18	0.97	0.06
$corr(PD_t,\beta_t)$	0.84	0.09	0.84	0.08	0.86	-0.21	0.88	-0.49

the same properties than in the baseline case. The main difference is that the model originates a bit more volatility. The reason is that having more resources boost stock demand and prices up which leads to more optimism that feeds back that into more demand and further price rises. Note, though, the extra volatility comes from an increase in returns volatility that is produced by more volatile beliefs, at odds with the data.

Under tax foresight, agents would include known tax changes in their model for prices. In particular, equation (10) now becomes

$$\frac{P_t}{P_{t-1}} = \frac{1-\tau_t^K}{1-\tau_{t-1}^K} b_t + \varepsilon_t^P \tag{49}$$

Consequently, subjective beliefs on capital gains become $\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}\right] = \frac{1-\tau_{t+1}^K}{1-\tau_t^K}\beta_t$. As column 4 to 8 of table 8 illustrates, the basic properties of the model are largely unchanged. If anything, the

*Table 6: **SMM estimation results including the risk-free interest rate.** This table reports the data and model moments, together with standard errors and t-statistics. The first four columns report the observed statistics along with their Newey-West standard error for the US data. The next four columns reports model-implied statistics and its t-statistics. The model uses the parameterization described in Table 4. Rates of growth, variance and convariance have been multiplied by 100. The top panel reports moments included in the SMM estimation; the bottom panel non-included moments.*

	US data		Model			
	1946-1982	1982-2018	1946-1982		1982-2018	
	$\hat{\mathcal{S}}_i$	$\hat{\mathcal{S}}_i$	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat
	Included in SMM estimation					
$\mathbb{E}(PD_t)$	25.48	47.09	27.84	-1.52	34.99	3.00
$\mathbb{E}(P_t/P_{t-1} - 1)$	0.48	1.84	0.63	-0.25	1.39	0.71
$\mathbb{E}(r_t^s)$	4.73	4.33	4.63	0.12	4.80	-0.62
$\mathbb{E}(r_t^b)$	0.42	0.38	0.26	9.40	0.61	-8.14
$\mathbb{V}ar(p_t - d_t)$	7.15	13.98	7.05	0.07	13.55	0.12
$\mathbb{C}ov(p_t - d_t, \bar{d}_t)$	-1.98	2.37	-0.24	-2.91	1.02	2.40
$\mathbb{C}ov(p_t - d_t, \bar{r}_t)$	-9.24	-9.68	-7.74	-0.99	-11.92	0.74
$\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^K)$	-0.29	1.73	-0.18	-0.38	0.56	3.79
$\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^D)$	0.36	1.00	-0.21	1.69	0.52	1.48
	Non-included in SMM estimation					
$\mathbb{E}(D_t/D_{t-1} - 1)$	0.49	0.75	0.49	0.00	0.75	0.00
$\sigma(D_t/D_{t-1})$	2.52	1.97	2.52	0.00	1.97	0.00
$\sigma(r_t^s)$	7.87	7.41	5.54	3.22	8.45	-1.28
$corr(PD_t, PD_{t-1})$	0.96	0.98	0.98	-0.17	0.96	0.30
$corr(PD_t, \beta_t)$	0.84	0.84	0.87	-0.30	0.83	0.31

model produces a bit less of volatility.

Finally, an increase in the realization of capital gains is introduced. The idea is that lower CGT would diminish the lock-in effect. This possibility is captured by allowing for a change in π in the 1982-2018 period. The data shows that the realized to total capital gains ratio modestly went up from 10% to 11% while the share of gains coming from equities over total gains went down from 45% to 25% since the 1980s. In terms of equation (47), that implies that the average realization rate was 11% in 1946-1982 and 12% afterwards. The model is parameterized accordingly. The last columns of table 8 shows that this increase in π barely changes the results.

3.6.- The lock-in effect

In the model in Section 2, I assumed that each investor faces some risk of being hit by a very bad shock $z_t^i \sim Bernoulli(\pi)$ that forces her to sell all her stocks. In equilibrium, a fraction π of agents

Table 7: ***RE model statistics.*** *This table reports the data and model moments, together with standard errors and t-statistics. In particular, the first two columns report the observed statistics for the US data. The next four columns reports model-implied statistics and its t-statistics. The model uses the parameterization described in Table 4, baseline column. Rates of growth, variance and convariance have been multiplied by 100.*

	US data		Model			
	1946-1982	1982-2018	1946-1982		1982-2018	
	$\hat{\mathcal{S}}_i$	$\hat{\mathcal{S}}_i$	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat
$\mathbb{E}(PD_t)$	25.48	47.09	73.64	-31.06	90.71	-10.80
$\mathbb{E}(P_t/P_{t-1}-1)$	0.48	1.84	0.70	-0.35	0.77	1.69
$\mathbb{E}(r_t^s)$	4.73	4.33	2.13	3.43	1.94	3.16
$\mathbb{V}ar(p_t-d_t)$	7.15	13.98	2.18	3.68	3.67	2.89
$\mathbb{C}ov(p_t-d_t,\bar{d}_t)$	-1.98	2.37	3.05	-8.40	2.09	0.50
$\mathbb{C}ov(p_t-d_t,\bar{r}_t)$	-9.24	-9.68	3.84	-8.64	0.40	-3.32
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^K)$	-0.29	1.73	2.43	-9.19	1.80	-0.24
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^D)$	0.36	1.00	0.22	0.41	0.50	1.53
$\mathbb{E}(D_t/D_{t-1}-1)$	0.49	0.75	0.49	0.00	0.75	0.00
$\sigma(D_t/D_{t-1})$	2.52	1.97	2.52	0.00	1.97	0.00
$\mathbb{E}(r_t^b)$	0.42	0.38	0.80	-23.98	1.18	-25.45
$\sigma(r_t^s)$	7.87	7.41	3.39	6.19	3.37	4.98
$corr(PD_t,PD_{t-1})$	0.96	0.98	0.98	-0.13	0.98	-0.10
$corr(PD_t,\beta_t)$	0.84	0.84	-1.00	-22.69	-1.00	-20.74

sells their assets and pay taxes; hence, the effective rate on total capital gains is $\pi\tau^K$. In this note, π is endogeneized following the approach used in Gavin et al. (2007) and Gavin et al. (2015).

Consider the same representative investor from Section 2. As before, she only pay capital gains taxes on realized gains but now she is allowed to decide the timing of realization. In particular, she manages a stock of unrealized capital gains G_t facing portfolio management costs. The stock of unrealized capital gains follows this law of motion

$$G_{t+1} = G_t + (P_t - P_{t-1})S_{t-1} - g_t + AC_t \tag{50}$$

with g_t are the realized capital gains and AC_t stands for adjustment costs. It is assumed $AC_t = g_t - \phi(\pi_t)G_t$ for $\pi_t = g_t/G_t$, with $\phi'(\cdot) > 0$, $\phi''(\cdot) < 0$. When the realization of capital gains is deferred, the cost function penalizes investors with extra unrealized capital gains. These higher stock of gains increase the future tax liability of households. Altogether, investors face an additional trade off: realize g_t capital gains and pay a taxes $\tau^K g_t$ today or defer the realization to the future and pay a cost.

Table 8: ***Results with alternative tax assumptions.*** *This table reports the data and model moments, together with standard errors and t-statistics for different models. The first four columns shows the results for the model with tax rebates. The next four columns for the model with tax foresight. The last columns for the model when π goes up in the second subsample. Estimated parameters are in table 4. Rates of growth, variance and convariance have been multiplied by 100. The top panel reports moments included in the SMM estimation; the bottom panel non-included moments.*

	Model with Tax Rebates				Model with Tax Foresight				Model with change in π			
	1946-1982		1982-2018		1946-1982		1982-2018		1946-1982		1982-2018	
	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat
Included in SMM estimation												
$\mathbb{E}(PD_t)$	27.88	-1.55	36.25	2.68	31.05	-3.59	38.21	2.20	28.00	-1.62	33.64	3.33
$\mathbb{E}(P_t/P_{t-1}-1)$	0.70	-0.35	1.29	0.87	0.73	-0.40	1.03	1.29	0.59	-0.17	1.22	0.99
$\mathbb{E}(r_t^s)$	4.58	0.19	4.50	-0.22	4.24	0.64	3.94	0.53	4.51	0.29	4.59	-0.34
$\mathbb{V}ar(p_t-d_t)$	4.73	1.79	11.77	0.62	4.87	1.68	7.70	1.76	5.81	0.99	9.61	1.22
$\mathbb{C}ov(p_t-d_t,\bar{d}_t)$	-0.38	-2.68	0.01	4.20	-2.11	0.21	-0.74	5.54	0.29	-3.80	2.28	0.16
$\mathbb{C}ov(p_t-d_t,\bar{r}_t)$	-6.68	-1.69	-10.05	0.12	-7.45	-1.18	-7.32	-0.78	-6.07	-2.09	-6.37	-1.09
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^K)$	-0.85	1.87	1.25	1.55	-0.46	0.57	0.91	2.65	-0.19	-0.36	1.97	-0.76
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^D)$	-0.05	1.23	0.56	1.34	0.09	0.81	0.51	1.49	-0.34	2.09	0.61	1.20
Non-included in SMM estimation												
$\mathbb{E}(D_t/D_{t-1}-1)$	0.49	0.00	0.75	0.00	0.49	0.00	0.75	0.00	0.49	0.00	0.75	0.00
$\sigma(D_t/D_{t-1})$	2.52	0.00	1.97	0.00	2.52	0.00	1.97	0.00	2.52	0.00	1.97	0.00
$\mathbb{E}(r_t^b)$	0.17	14.67	0.33	0.57	0.46	-2.92	1.06	-21.76	0.34	4.03	0.92	-17.62
$\sigma(r_t^s)$	5.48	3.29	8.14	-0.90	5.22	3.66	6.06	1.66	4.98	3.98	6.35	1.31
$corr(PD_t,PD_{t-1})$	0.98	-0.14	0.96	0.26	0.98	-0.13	0.98	0.00	0.98	-0.18	0.98	0.02
$corr(PD_t,\beta_t)$	0.82	0.52	0.86	0.56	0.80	-0.74	0.85	0.44	0.82	-0.52	0.85	0.44

With this additional decision, the investor's problem consists in choosing sequences of consumption, stock holdings, realized and unrealized capital gains as to maximize their lifetime welfare:

$$\max_{\{C_t,S_t,g_t,G_{t+1}\}_{t=0}^{\infty}} \mathbb{E}_0^{\mathcal{P}} \sum_{t=0}^{\infty} \delta^t U(C_t) \tag{51}$$

subject to the budget constraint

$$C_t + P_t S_t \leq ((1-\tau^D)D_t + P_t)S_{t-1} - \tau^K g_t + T_t \tag{52}$$

to the unrealized capital gains law of motion

$$G_{t+1} = G_t + (P_t - P_{t-1})S_{t-1} - \phi\Big(\frac{g_t}{G_t}\Big)G_t \tag{53}$$

and stock holdings bounds

$$0 \leq S_t^i \leq \bar{S}, \text{ given } S_{-1}^i = 1 \tag{54}$$

Let λ_t and μ_t be the Lagrange multipliers associated with constraints (52) and (53). The first order conditions are given by

$$U_t' = \lambda_t \tag{55}$$

$$P_t = \delta \mathbb{E}_t^{\mathcal{P}} \left[\frac{\lambda_{t+1}}{\lambda_t} \Big((1-\tau^D) D_{t+1} + P_{t+1} - \frac{\mu_{t+1}}{\lambda_{t+1}} (P_{t+1} - P_t) \Big) \right] \tag{56}$$

$$\tau^K \lambda_t = \mu_t \phi_t' \tag{57}$$

$$\mu_t = \delta \mathbb{E}_t^{\mathcal{P}} \left[\mu_{t+1} \Big(1 - \phi_{t+1}' G_{t+1} - \phi_{t+1} \Big) \right] \tag{58}$$

where x_t' stands for the first derivative of function x with respect to its argument at time t. For the application, let $\phi(\pi_t) = ln(\pi_t)$. Hence, $\phi_t' = 1/\pi_t$. Using that and combining FOCs, three optimality conditions come up. First, expression (57) determines the optimal realization of capital gains:

$$\pi_t = \frac{\mu_t}{\lambda_t \tau^K} \tag{59}$$

The optimal realized-to-total gains fraction depends on three terms: i) positively on the shadow price of unrealized gains μ_t indicating that the more costly non-realizing gains is, the more agents realize them; ii) negatively on the marginal value of wealth at period λ_t as investors would avoid realizing gains to save taxes if today's resources are scarcer; iii) negatively on the tax level τ^K, reflecting the lock-in effect.

Table 9 shows the model performance when the lock-in effect is endogenous. As expected, the increase in PD volatility is lower, about 40% of the observed. Thus, the lock-in effect partially counteracts the effect through demand that has been previously quantified. In net terms, though, the demand effect dominates the lock-in effect.

3.7.- An alternative hypothesis: the drop in the safe real interest rate

In this section, I use the model to evaluate a possible alternative explanation: the fall in safe real interest rates. The connection between interest rates and asset prices booms and busts have been suggested in some papers (e.g., Taylor (2007), Martin and Ventura (2018)) and explicitly explored in models of learning (Adam et al. (2012), Adam and Merkel (2019)). Although the safe rate can affect asset prices through a variety of channels, I limit the discussions to its effects on the discount

*Table 9: **Model with lock-in effect.** This table reports the data and model moments, together with standard errors and t-statistics. In particular, the first two columns report the observed statistics for the US data. The next four columns reports model-implied statistics and its t-statistics. The model uses the parameterization described in Table 4, baseline column. Rates of growth, variance and convariance have been multiplied by 100.*

	US data		Model			
	1946-1982	1982-2018	1946-1982		1982-2018	
	$\hat{\mathcal{S}}_i$	$\hat{\mathcal{S}}_i$	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat	$\widetilde{\mathcal{S}}_i(\hat{\theta})$	t-stat
$\mathbb{E}(PD_t)$	25.48	47.09	25.58	-0.06	31.50	3.86
$\mathbb{E}(P_t/P_{t-1}-1)$	0.48	1.84	0.59	-0.19	1.21	0.99
$\mathbb{E}(r_t^s)$	4.73	4.33	4.86	-0.18	4.77	-0.57
$\mathbb{V}ar(p_t-d_t)$	7.15	13.98	5.18	1.45	8.27	1.60
$\mathbb{C}ov(p_t-d_t,\bar{d}_t)$	-1.98	2.37	-0.03	-2.79	-0.00	4.26
$\mathbb{C}ov(p_t-d_t,\bar{r}_t)$	-9.24	-9.68	-5.96	-2.16	-6.78	-0.96
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^K)$	-0.29	1.73	-0.17	-0.41	1.02	-2.29
$\mathbb{C}ov(p_t-d_t,\bar{\tau}_t^D)$	0.36	1.00	-0.16	1.56	0.56	1.32
$\mathbb{E}(D_t/D_{t-1}-1)$	0.49	0.75	0.49	0.00	0.75	0.00
$\sigma(D_t/D_{t-1})$	2.52	1.97	2.52	0.00	1.97	0.00
$\mathbb{E}(r_t^b)$	0.42	0.38	0.38	0.50	1.04	-10.02
$\sigma(r_t^s)$	7.87	7.41	4.96	3.53	5.77	2.67
$corr(PD_t,PD_{t-1})$	0.96	0.98	0.98	-0.13	0.98	-0.10
$corr(PD_t,\beta_t)$	0.84	0.84	0.84	-0.10	0.89	-0.63

factor, which is the case made by Adam and Merkel (2019).[47] Other things equal, lower interest rates would increase δ and then the sensitivity of prices to expectations. In this sense, interest rates and taxes are alike.

The drop in interest rates is included in the model through the discount factor δ. I recover a series of $\{\delta_t\}$ such that the model perfectly replicate the observed path of the risk-free rate. For that purpose, I use the Euler Equation for bonds (36) with the empirical risk-free time series, that is,

$$\delta_t = \left\{(1+(1-\tau_t^b)r_t^b)\mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+1}/D_{t+1}}{C_t/D_t}\frac{D_{t+1}}{D_t}\right)^{-\gamma}\right]\right\}^{-1} \tag{60}$$

where τ_t^b and r_t^b are the observed tax and interest rate. However, this expression is not enough to recover δ_t as consumption is endogenous and its expectation depends on δ. To overcome this

[47] In particular, low interest rate can boost indebtedness, which interacts with beliefs-driven booms via collateral constraints.

problem, I resort to the following approximation

$$\mathbb{E}_t^{\mathcal{P}}\Big[\Big(\frac{C_{t+1}/D_{t+1}}{C_t/D_t}\frac{D_{t+1}}{D_t}\Big)^{-\gamma}\Big] \approx \mathbb{E}_t\Big[\Big(\frac{C_{t+1}/D_{t+1}}{C_t/D_t}\frac{D_{t+1}}{D_t}\Big)^{-\gamma}\Big] = \mathbb{E}_t\Big[\Big(\frac{W_{t+1}/D_{t+1}+1}{W_t/D_t+1}\frac{D_{t+1}}{D_t}\Big)^{-\gamma}\Big] \quad (61)$$

The expected consumption growth according to the subjective probability is approximately equal to the expected equilibrium growth. It can be shown that

$$\mathbb{E}_t\Big[\Big(\frac{W_{t+1}/D_{t+1}+1}{W_t/D_t+1}\frac{D_{t+1}}{D_t}\Big)^{-\gamma}\Big] = \Big[(1+\rho)^{1-p}\beta^D\Big(1+\frac{W_t}{D_t}\Big)^{p-1}\Big]^{-\gamma}\mathbb{E}_t\Big[(\varepsilon^W_{t+1}\varepsilon^D_{t+1})^{-\gamma}\Big] \quad (62)$$

and

$$\mathbb{E}_t\Big[(\varepsilon^W_{t+1}\varepsilon^D_{t+1})^{-\gamma}\Big] = exp\Big\{-\frac{\gamma}{2}(\sigma^2_W+\sigma^2_D)+\frac{\gamma^2}{2}(\sigma^2_W+\sigma^2_D-2\sigma_{DW})\Big\}$$

Thus, $\{\delta_t\}$ is contingent on wage-dividend realizations. With this time-varying discount factor, the model is reestimated. The results are remarkably similar to the baseline; the same 15 moments pass their t-test.[48] Hence, the inclusion of a time-varying discount factor does not significantly improve the ability to generate a high enough volatility rise.

Would the increase in volatility have occurred in the absence of tax cuts? In other words, was it primarily driven by the drop in the real interest rate? To answer this question, I simulate the model without tax cuts, that is, $\tau^K_t = \tau^K_{1976.IV}$ for $t > 1976.IV$, as the tax rate began to fall after 1976. In this case, the model predicts that the mean PD ratio in the 1982-2018 would have been roughly the same as before the 1980s, and the volatility would have fallen by about 40% (instead of increasing by 140%). Hence, higher taxes would have prevented the increase in both the level and fluctuations of the PD ratio, despite the drop in real interest rates. An alternative exercise is to run the model without the drop in the interest rate, that is, keeping $r^b_t = r^b_{1981.III}$ for $t > 1981.III$, when it achieved its highest value. Would tax cuts still disturb the market? In this case, mean and volatility go up considerably less; for instance, the volatility increases by 40% instead of by 140%. Altogether, this exercise suggests that, even after accounting for the fall in interest rates, tax cuts continue to be relevant in explaining the rise in volatility.

Finally, I present a complementary exercise involving calibration. Using the model's estimation, I calibrate a series of stock supply shocks as to achieve a perfect match between the model-implied and the empirical PD ratio. In this way, the model's PD ratio is generated by a known combination of taxes, interest rates, dividends, wages and a residual force. Then, I run a similar experiment: instead of the decline in taxes observed after 1976, I introduce a moderate tax rise, $\tau^K_t = 0.25$ for

[48] Not reported here due its similarity with table 5.

Table 10: ***Price-dividend mean and variance; estimated model and counterfactuals.*** *The table reports statistics for the estimated model using the observed path of the risk-free rate and two counterfactuals that show the consequences without the decline either of taxes or interest rates. Numbers in gray show the increase in the statistic in the 1982-2018 period with respect to its level in the 1946-1982 period.*

	Model		$\tau_t^K = \tau_{1976.IV}^K$ for $t > 1976.IV$	$r_t^b = r_{1981.III}^b$ for $t > 1981.III$
	1946-1982	1982-2018	1982-2018	1982-2018
$\mathbb{E}(PD_t)$	27.43	35.19	26.47	32.05
		+28.29%	-3.50%	+16.84%
$\mathbb{V}ar(p_t - d_t)$	4.9	11.51	3.08	7.02
		+134.90%	-37.14%	+43.27%

$t > 1976.IV$. Graph 5 pictures the exercise. With higher taxes, the PD ratio would have evolved more smoothly, avoiding the increase in both level and volatility, despite the decline in safe real interest rates.

3.8.- The Equity Premium

This section explores the reasons behind the relatively good equity premium generated by the learning model using a realistic consumption and dividend growth processes, a positive discount factor and low risk-aversion. First, it analyzes the drivers behind mean stock returns. Second, it explores its relation to the drivers of the risk-free rate.

To articulate the discussion, I use the following decomposition of the stock return geometric mean[49]

$$\left(\prod_{t=1}^{N} \frac{P_t + D_t}{P_{t-1}}\right)^{\frac{1}{N}} = \underbrace{\left(\prod_{t=1}^{N} \frac{D_t}{D_{t-1}}\right)^{\frac{1}{N}}}_{R_1} \underbrace{\left(\frac{PD_N + 1}{PD_0}\right)^{\frac{1}{N}}}_{R_2} \underbrace{\left(\prod_{t=1}^{N-1} \frac{PD_t + 1}{PD_t}\right)^{\frac{1}{N}}}_{R_3} \tag{63}$$

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 over the initial PD ratio value, which might 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 11 reports the decomposition using empirical and simulated data. Since the dividend growth process has been parameterized directly from the data, the models replicate R_1 fairly well. Regarding R_2, on average both models show certain increase in the PD ratio in both periods while in reality the R_2 is only above 1 in the second period. This is a mismatch of the model. An

[49] It was first suggested by Adam et al. (2016).

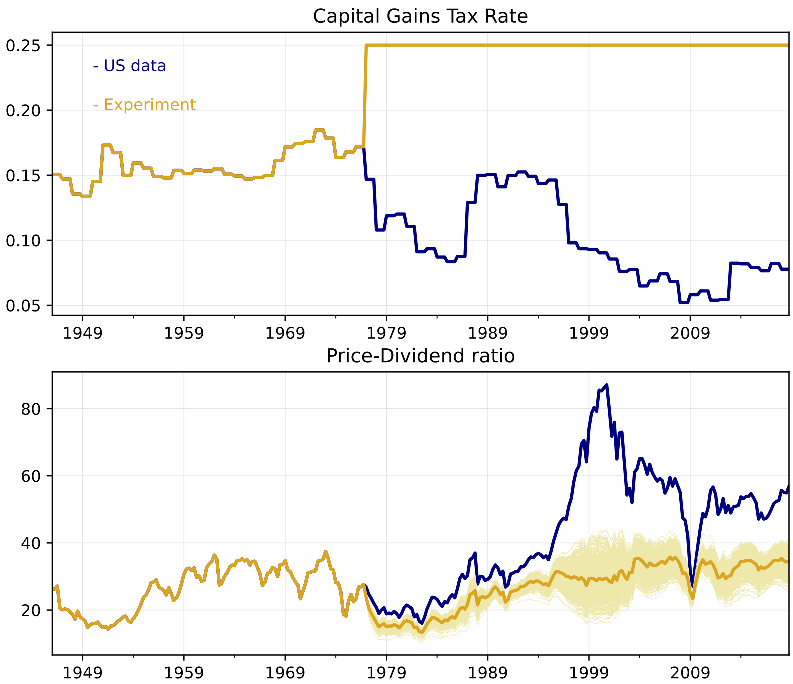

Figure 5: ***Effects of an alternative tax policy in the context of falling interest rates.*** *The graphs shows that observed PD ratio that is perfectly matched by the model that uses the observed path of taxes and interest rates, and a calibrated series of stock supply shocks. Yellow lines show the experiment, a moderate increase in the tax rate, and its consequences for the PD ratio. Light yellow lines in the bottom panel represent different simulations. A total of 1000 simulations were run.*

important part of it is due to the capital gains tax cut from the early 1980s, which tend to bring the PD ratio up during the final quarters 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)[50]. Since R_1 and R_2 look similar for learning and RE the difference must come from R_3. Indeed, the RE values for R_3 are about one half of the right ones while for learning they are a bit closer. In fact, the too low R_2 is compensated by a too high R_3 for the 1982-2018 period. Altogether, and besides this trading between R_2 and R_3, the learning model matches the mean stock return rather well because it gets the mean and volatility of the PD ratio correctly using the calibrated dividend process[51].

Table 11: ***Decomposition of the stock return geometric mean.*** *The table shows the stock returns mean decomposition according to expression (63). The first column uses U.S. data; the second, simulated data using the learning model; the third, simulated data using the RE model. The last row is the stock return geometric mean. Simulated data uses the parameterization shown in table 3.*

	US data		Learning model		RE model	
	1946-1982	1982-2018	1946-1982	1982-2018	1946-1982	1982-2018
R_1	0.46	0.74	0.46	0.74	0.46	0.74
R_2	-0.27	0.88	0.14	0.36	0.24	-0.01
R_3	4.10	2.37	4.28	3.43	1.44	1.60
$\mathbb{E}(r^s)$	4.30	4.03	4.91	4.57	2.15	2.34

The second part of the equity premium is the risk-free rate. Although it has no closed-form solution in the quantitative model, it does so for RE when $p = 1$. This benchmark is useful to understand why it is not too high. It is given by

$$r_t^b = \left(\delta^{-1}\mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]^{-1} - 1\right)\frac{1}{1-\tau_t^b} \tag{64}$$

where $\mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right] = a^{-\gamma}exp\Big\{\gamma(\sigma_W^2 + \sigma_D^2)(1+\gamma)/2\Big\}exp\{\sigma_{DW}\gamma^2\}$.

Thus, r_t^b depends essentially on the mean and volatility of the income processes and the level of risk aversion. As a result, getting a high stock return and a low bond rate is complicated in many models; the reason is that either high risk aversion or income volatility is needed. However, the required levels appear unrealistic (Mehra and Prescott (1985)) plus high risk aversion would also lead to a too high risk-free rate (Weil (1989)). Contrarily, this paper resort to alternative

[50] If R_2 is computed burning some periods before and after 1982:II, observed and simulated data look much more alike (1.0049 vs 1.0041 for the 1982-2018 period).

[51] When mean dividend growth is estimated instead of simply calibrated from the data, it tends to be too high which ends up depressing stock returns.

forces that make returns high enough. The main driver is non-fundamental volatility coming from beliefs, which makes compatible realistic income processes with volatile enough, and then high, stock returns. However, beliefs volatility is unable to do all the job (Adam et al. (2016), Adam et al. (2017)). The second driver is the increase in the PD ratio brought about by stronger income growth and lower taxes. Thus, $R_3 > 1$ helps to increase mean stock returns. In other words, relying on beliefs alone would either be insufficient (as in Adam et al. (2017)) or require a too high beliefs volatility while introducing a trend in the PD ratio (as the one coming from taxes) helps sorting out this problem.

The previous reasoning explains why the model does a decent job at matching the equity premium level. Additionally, its decline is captured by the model too. In reality as well as in the model, the fall in stock returns is mostly due to the reduction in R_3 as a result of a higher PD ratio, which overcomes the opposite effect via R_2. Besides, the mean risk-free rate is also declining, mostly due to the fall in τ^b.[52] In other words, the fact the model produces an increase in the PD ratio helps to explain both the level and trajectory of the equity premium.

3.9.- Additional tests about the model's main mechanism

This section presents time series evidence of the model predictions. An important mechanism of the the model is that tax cuts increase the elasticity of the PD ratio to beliefs. I use survey expectations to test it. Additionally, the model predicts a larger reaction of the PD ratio to shocks in low tax regimes. The data is aligned with these two model's predictions.

Consider the model with learning from Section 2. Assuming $\pi = 1$ for simplicity, the sensitivity of prices to beliefs can be measured in terms of a time-varying elasticity that reads as

$$\epsilon_t \equiv \frac{\partial P_t/D_t}{\partial \beta_t^p} \frac{\beta_t^p}{P_t/D_t} = \frac{\delta \beta_t^p (1-\tau^K)}{1-\delta(1-\tau^K)\beta_t^p - \delta\tau^K} \tag{65}$$

This elasticity depends on the capital gains tax level. In particular,

$$\frac{\partial \epsilon_t}{\partial \tau^K} < 0 \text{ if } \delta < 1 \tag{66}$$

Besides, from the difference equation (**??**) governing expectations dynamics, it can be shown

[52]In this case, there is no feedback loop affecting the bond price due to its one-period maturity. Hence, τ^b level is neutral and only tax changes have an impact on bond prices. These impact is very small so that the risk-free rate is very stable.

that

$$\frac{\partial \beta_t}{\partial \tau^K} < 0 \text{ if } \delta^2(1-\tau^K)(\beta_{t-1}-1) < 1 \tag{67}$$

The condition is satisfied for standard values of the parameters. These two partial derivatives give us the sign of the reaction of the variables of interest to a tax change. To get the dynamic response, I resort to simulations. I run a VAR with capital gains taxes, the PD ratio, expectations and the time-varying elasticity using both empirical and simulated data coming from the baseline model. Empirical expectations are taken from surveys.[53] The time-varying elasticity is obtained using a state-space model as

$$\begin{aligned} \ln\frac{P_t}{D_t} &= \mu + \tilde{\epsilon}_t \ln\beta_t^S + v_t^p \\ \tilde{\epsilon}_t &= \tilde{\epsilon}_{t-1} + w_t^{\tilde{\epsilon}} \end{aligned} \tag{68}$$

where $\tilde{\epsilon}_t$ is the desired elasticity, β_t are expectations, either from surveys or from the model and v_t^p and $w_t^{\tilde{\epsilon}}$ are jointly Normal i.i.d zero-mean perturbations with constant variance and uncorrelated. In other words, it is a Kalman algorithm used to estimate the unknown time-varying parameter $\tilde{\epsilon}$. The model parameters are estimated via Maximum Likelihood. Model details and estimation results are in Appendix E. With the parameters and the time series for the PD ratio and survey expectations I can pin down a series $\{\tilde{\epsilon}_t\}$. Figure 6 plots the elasticity.

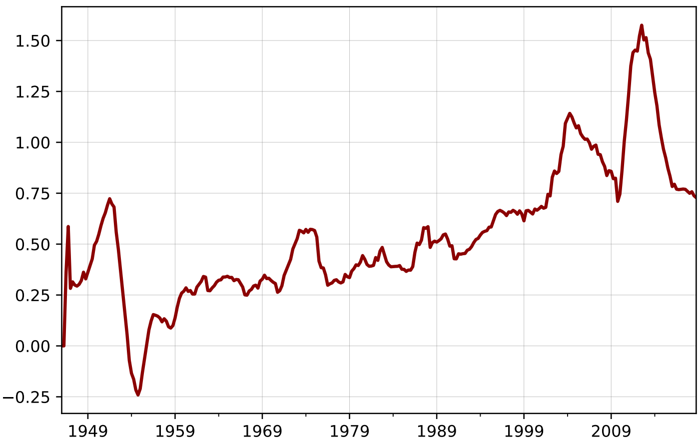

Figure 6: ***Estimated expectations elasticity of the PD ratio.*** *The line plots the time series of* $\{\tilde{\epsilon}_t\}$ *as defined in equation 68, using the parameters from the MLE (see Appendix E). The time series cover the 1946:I-2018:IV period.*

[53] I use UBS Gallup survey data as baseline, which is the UBS Index of Investor Optimism. It has monthly data from 1998:M5 to 2007:M10, with 702 responses per month on average and has thereafter been suspended. Some adjustments have been made to convert monthly stock returns expectations into quarterly risk-adjusted capital gains ones, in line with the model. See Appendix A. Furthermore, the survey data is extended to cover the whole postwar period, following the approach in Adam et al. (2017).

Hence, the following model is fitted into both simulated and observed data:

$$\boldsymbol{Y}_t = A + B\boldsymbol{Y}_{t-1} + \boldsymbol{u}_t \tag{69}$$

with $\boldsymbol{u}_t \sim \mathcal{N}(0, \Omega)$ and $\boldsymbol{Y}_t = [\tau_t^K, \beta_t^S, P_t/D_t, \tilde{\epsilon}_t]$. I impose $\boldsymbol{u}_t = S\bar{\boldsymbol{u}}_t$ with S being the Cholesky factor of the matrix Ω and $\bar{\boldsymbol{u}}_t \sim \mathcal{N}(0, I)$.[54]

Figure 7 shows the response of the variables to a persistent reduction in taxes of 1pp. As expected, the tax shock has persistent effects on taxes and also on the PD ratio, in line with the tax capitalization hypothesis. Besides, the data seems to be consistent with the suggested mechanism: both expectations and expectations elasticity of the PD ratio go up following a negative tax shock.[55]

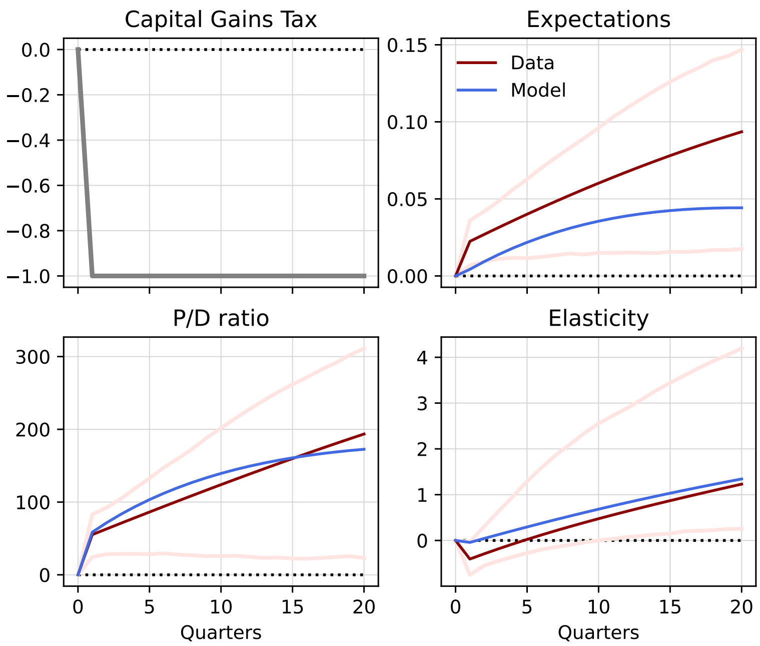

Figure 7: ***Responses to a persistent reduction in*** τ^K***.*** *The figure plots the response of different variables to a 1pp persistent reduction in the capital gains tax. Red lines are the responses using the observed data for the US. Blue lines are the responses implied by the quantitative model with the baseline parameterization of table 3. Bands at 68% confidence level have been computed via data simulations using 1000 repetitions.*

In addition, the model predicts that the response of the PD ratio to shocks would be larger and more durable in lower tax regimes. The underlying reason is that shocks can get amplified via expectations and high taxes would lean against this amplification. To test it, a minimalist

[54] I use short run restrictions for identification. Following the theory and how the variables have been computed, the ordering is: taxes, expectations, the PD ratio and the elasticity. Results are robust the switching the PD ratio and the elasticity.

[55] This result offers an explanation for an issue pointed out by Adam et al. (2017): the correlation between prices and beliefs became stronger after the year 2000. They suggested it had to do with lower interest rates. It seems, though, the elasticity began raising before, so tax cuts seems to play a role.

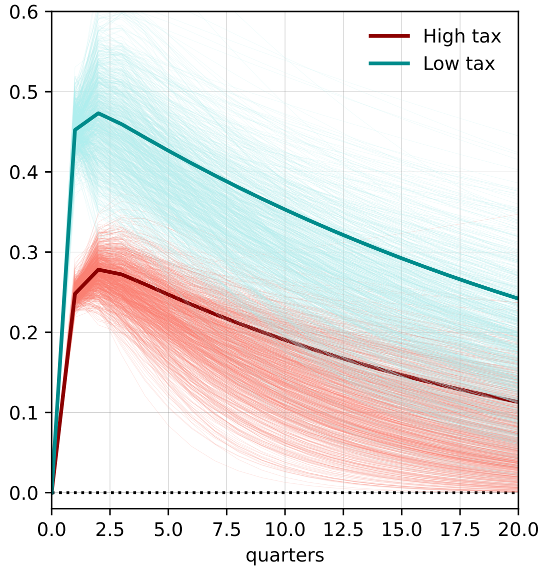

Figure 8: ***Empirical response of the PD ratio to a price shock in different tax regimes.*** *The graph plots the response of the PD ratio to an equivalent price shock in two different subsamples: red lines signals the 1946-1982 sample characterized by higher taxes; blue lines are for the 1982-2018, characterized by lower taxes. Thinner lines represents 1000 repetitions using bootstrapped shocks.*

VAR model with price growth and the PD ratio is estimated for each sample. Indeed, it turns out the PD ratio increased almost twice as much as in the low tax sample than in the high one when stimulated by a price shock of the same magnitude, as figure 8 illustrates.

4.- Optimal Capital Gains Taxation

In this section, the normative use of a tax on unrealized capital gains for macrofinancial stability purposes is studied. Thus, whereas Section 2 pointed out the particular role of capital gains taxation and Section 3 analyzed the ability of taxes, as historically given, to explain certain transformations in the US stock market, this section asks: What is the appropriate capital gains tax to get the best of imperfect capital markets?

The problem in hand is the excess volatility in capital prices which can be read as a pecuniary externality. The reason is that excess volatility emerges from the inability of agents to internalize the equilibrium price formation due to information frictions.[56] Thus, the lack of knowledge of the true determinants of prices pushes agents to make decisions using forecasts derived from their subjective models that ignore the effect of their own forecasts on market prices and everybody else's

[56] This inability is due to the fact investors ignore other investors' characteristics such that the standard derivation of equilibrium prices combining individual and aggregate optimality conditions is not possible.

predictions. In short, rational individuals trying to make the best prediction about future prices but missing general equilibrium effects end up causing excessive volatility.

A precondition to exploring an optimal tax is establishing a connection between asset prices and consumption fluctuations that is missing from the previous endowment model. To that end, I set up a tractable two-sector growth model with investment adjustment costs and learning about capital prices. The model links the capital market price to investment decisions, in line with the Q-theory. As a result, cycles of over- and under-accumulation of capital emerge, driven by excessive asset price fluctuations.

The section is structured as follows. Section 4.1. sets up a centralized two-sector growth model with investment adjustment costs. Section 4.2. decentralizes the economy by introducing efficient capital markets. Contrarily, Section 4.3. decentralized the economy when investors have imperfect market knowledge and learn about prices. Section 4.4. studies an optimal taxation problem. Finally, Section 4.5. proposes an alternative implementation of the optimal policy that avoids too volatile taxes.

4.1.- The First Best economy

In this section, a model with endogenous consumption is introduced. It consists of a two-sector growth model with investment adjustment costs. The model is highly simplified, reduced to the minimum ingredients needed to connect capital prices to output. The model structure is described next.

Demographics. The economy is populated by a continuum of measure 1 of infinitely living identical agents.

Goods. There is a perishable consumption good (or simply "good") and a non-perishable capital good (or simply "capital") that depreciates at a constant rate d each period. Goods deliver utility whereas capital is used to produce goods.

Production technology. There is a goods production function that uses capital K with an inelastically supplied 1 unit of labour in a particular technological environment given by Z to deliver goods $F : (Z, K) \to \mathbb{R}_+$. F has neoclassical properties; the technology level Z is exogenous and stochastic. In addition, capital is produced via a linear function that converts $I_t + G(I_t)$ units of goods into I_t units of capital; $G(I_t)$ represents investment adjustment costs, a convex function, symmetric, with $G(0) = 0$, $G'(0) = 0$ and $G''(\cdot) > 0$[57].

[57]This specification implicitly assumes diminishing returns to scale in adjustment costs. In this way, Hayashi (1982)'s theorem does not hold. The violation of the theorem can be avoided by assuming G also depends negatively

Welfare. The utility function U is time-separable, continuous, at least twice-differentiable function with $U'(C_t) > 0$ and $U''(C_t) < 0$, with Inada properties.

Social Planner's problem. In the previous economy, the Social Planner faces a dynamic allocation problem consisting on distribute goods between investment and consumption to maximize the lifetime social welfare:

$$\max_{\{C_t, I_t, K_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \delta^t U(C_t) \tag{70}$$

s.t. i) Consumption-goods resource constraint:

$$C_t + I_t + G(I_t) \leq F(Z_t, K_{t-1}) \tag{71}$$

ii) Capital-goods resource constraint:

$$K_t \leq I_t + (1-d)K_{t-1} \tag{72}$$

iii) Non-negative consumption:

$$C_t \geq 0 \tag{73}$$

First Best (optimal growth path). Given initial capital K_{-1} and an exogenous productivity process $\{Z_t\}_{t=0}^{\infty}$, the Social Planner equilibrium consists of sequences of allocations $\{C_t, I_t, K_t\}_{t=0}^{\infty}$ such that:

1. Resource constraints (71)-(72) are satisfied.

2. First order conditions:

$$u_t^c = \lambda_t \tag{74}$$

$$1 + G_t^I = q_t/\lambda_t \tag{75}$$

$$\frac{q_t}{\lambda_t} = \delta \mathbb{E}_t \left[\frac{\lambda_{t+1}}{\lambda_t} \left(F_{t+1}^k + (1-d) \frac{q_{t+1}}{\lambda_{t+1}} \right) \right] \tag{76}$$

   where $G_t^I = \frac{\partial G(I_t)}{\partial I_t}$, $u_t^c = \frac{\partial U(C_t)}{\partial C_t}$, $F_{t+1}^k = \frac{\partial F(Z_{t+1}, K_t)}{\partial K_t}$; λ_t is the Lagrange multiplier of the goods resource constraint, reflecting the marginal value of goods; q_t is the Lagrange multiplier of the capital resource constraint and then, q_t/λ_t reflects the marginal value of capital (in terms of goods).

on K and it is homogeneous of degree one. However, it complicates the analysis a bit without adding any crucial insight to the question in hand. See Romer and Romer (2010) for a discussion.

3. A transversality condition.

$$\lim_{j\to\infty} \delta^j \mathbb{E}_t \left[\frac{u^c_{t+j}}{u^c_t} \frac{q_{t+j}}{\lambda_{t+j}} K_{t+j} \right] = 0 \tag{77}$$

Altogether, the model equilibrium is characterized by choices about capital accumulation. It is determined by the intersection of two functions that relate capital stock to its shadow price. First, combining equations (72) and (75), a positive relationship between K and $\bar{q} \equiv q/\lambda$ arises. Besides, by the properties of F, the Euler Equation (76) gives rise to a negative $K - \bar{q}$ relationship. These two curves pin down a unique equilibrium, from which investment, consumption and output follows. As it is well-known, in dynamic terms the model is described by two difference equations characterizing the evolution of K and $\bar{q}$.[58]

4.2.- Efficient Markets

In this section, investment is decentralized. Thus, on top of the previous elements, markets are introduced and with them the information atomistic investors possess is specified.

Markets. The economy consists of two markets for capital and goods. In the former, capital producers and capital users meet to sell and buy new and old capital at price Q_t[59]. In the latter, capital producers acquire the inputs they need to produce new capital and households meet their consumption demand. Goods price acts as the unit of account of the economy and as such, it is normalized to 1. Markets are competitive.

Information set. Agents have all the structural knowledge about the economy. In particular, homogeneity is common knowledge and households are aware of firms' problem.

Then, we must characterize the problems of the two group of agents: capital producers and producing households.

Capital producers. They maximize profits by choosing investment on new capital. Then,

[58] See Romer and Romer (2010) for a textbook treatment.

[59] A mapping between capital and stock price can be established along Adam and Merkel (2019)'s lines. Assume that capital K_t can be securitized via equities S_t without any cost. In equilibrium, arbitrage is not possible and then, the ex-dividend equity price must be equal to the market value of capital net of dividends. Thus, consider that a fraction $x \in (0,1)$ of profits is distributed such that dividends $D_t = xK_{t-1}F^k_t$. Assume that the rest is reinvested in new capital $(1-x)K_{t-1}F^k_t/Q_t$. Hence, the market value of capital per share after dividends payments is $P_t = Q_t((1-d)K_{t-1} + (1-x)K_{t-1}F^k_t Q_t)$. It follows that the PD ratio is given by

$$\frac{P_t}{D_t} = \frac{(1-d)}{x}\frac{Q_t}{F^k_t} + \frac{1-x}{x} \tag{78}$$

For reasonable x (not too small), the PD is basically a proportion of the Capital-Rent ratio. Therefore, the connection with the stock market model is that learning about stock prices would be an implicit way of learning about the market value of capital.

they acquire goods to produce capital (facing the adjustment costs $G(I_t)$) that will be sold at price Q_t in capital markets as to maximize their profits Π_t. Their static problem can be stated as

$$\max_{\{I_t\}_{t=0}^{\infty}} \Pi_t = Q_t I_t - I_t - G(I_t) \tag{79}$$

Producing households. In this economy, households buy goods to satisfy their consumption demand and capital to produce goods. Each of them supply a unit of labour inelastically. Hence, their problem can be written as

$$\max_{\{C_t,K_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \delta^t U(C_t) \tag{80}$$

s.t. i) Budget constraint:

$$C_t + Q_t K_t \leq F(Z_t, K_{t-1}) + (1-d) Q_t K_{t-1} + \Pi_t \tag{81}$$

ii) Non-negative consumption:

$$C_t \geq 0 \tag{82}$$

Competitive Equilibrium. Given K_{-1}, a Competitive Equilibrium consists of sequences of allocations $\{C_t, I_t, K_t\}_{t=0}^{\infty}$ and prices $\{Q_t\}_{t=0}^{\infty}$ such that:

1. Capital producers behave optimally, satisfying

$$Q_t = 1 + G_t^I \tag{83}$$

2. Households behave optimally, satisfying:

   (a) The sequence of budget constraints (81).

   (b) The sequence of Euler Equation

$$Q_t = \delta \mathbb{E}_t \left[\frac{u_{t+1}^c}{u_t^c} \left(F_{t+1}^k + (1-d) Q_{t+1} \right) \right] \tag{84}$$

   (c) A transversality condition

$$\lim_{j \to \infty} \delta^j \mathbb{E}_t \left[\frac{u_{t+j}^c}{u_t^c} Q_{t+j} K_{t+j} \right] = 0 \tag{85}$$

3. Markets clear:

$$\text{Goods: } C_t + I_t + G(I_t) = F(Z_t, K_{t-1}) \tag{86}$$

$$\text{Capital: } K_t = I_t + (1-d)K_{t-1} \tag{87}$$

First Welfare Theorem. It is clear that both institutions, the planner and markets, have to satisfy the same aggregate resource constraints. Besides, in equilibrium, market and planner's Euler Equation reads exactly the same

$$1 + G_t^I = \delta \mathbb{E}_t \left[\frac{\lambda_{t+1}}{\lambda_t} \left(F_{t+1}^k + (1-d)(1 + G_{t+1}^I) \right) \right] \tag{88}$$

which implies

$$Q_t = \bar{q}_t \tag{89}$$

Thus, the market capital price is equal to its shadow price. By the arguments in the previous section, it follows that quantities will be those of the First Best.

4.3.- Inefficient Markets

In this section, the full information assumption is relaxed. This departure from Rational Expectations gives rise to an additional uncertainty source, price formation, that adds new dynamics to the model. First, the new information set is specified:

Information set. Households have structural knowledge about the economy except they ignore they all are equal. This incomplete information makes them unable to derive current capital prices from their optimality conditions since they cannot neither use market clearing conditions ex-ante nor apply the Law of Iterated Expectations. This friction is formalized by introducing a subjective probability measure $\mathcal{P}^i$ that reflects investors' views about productivity, capital and prices.Thus, the underlying probability space is given by $(\Omega, \mathcal{B}, \mathcal{P}^i)$ with $\mathcal{B}$ denoting the corresponding σ-algebra of Borel subsets of Ω and $\mathcal{P}^i$ agent's i subjective probability measure over $(\Omega, \mathcal{B})$. For generality, we include prices in the the state space Ω, with $\omega = \{Z_t, K_t, P_t\}_{t=0}^{\infty}$ as a typical element.

In this world, the problems agents face are the same as in the efficient market case except now households use their subjective probability measure, that is

$$\max_{\{C_t, K_t\}_{t=0}^{\infty}} \mathbb{E}_0^{\mathcal{P}} \sum_{t=0}^{\infty} \delta^t U(C_t) \tag{90}$$

Hence, the Euler Equation reads as

$$Q_t = \delta \mathbb{E}_t^{\mathcal{P}} \left[\frac{u_{t+1}^c}{u_t^c} \Big(F_{t+1}^k + (1-d) Q_{t+1} \Big) \right] \tag{91}$$

To fully characterize equilibrium, the following subjective price model is assumed:

$$\frac{Q_{t+1}}{Q_t} \frac{u_{t+1}^c}{u_t^c} = \theta_t + \varepsilon_t^P \tag{92}$$

$$\theta_t = \theta_{t-1} + \nu_t \tag{93}$$

with i.i.d. normally distributed innovations. The posterior of the unobserved component θ follows a Normal distribution

$$\theta_t \sim \mathcal{N}(\ln \beta_t, \sigma_\theta^2)$$

where σ_θ^2 is the steady state Kalman estimate uncertainty and the posterior mean evolves recursively following

$$\beta_{t+1} = \beta_t + g \Big(\frac{Q_t}{Q_{t-1}} \frac{u_t^c}{u_{t-1}^c} - \beta_t \Big) \tag{94}$$

Hence, $\mathbb{E}_t^{\mathcal{P}} \left(\frac{Q_{t+1}}{Q_t} \frac{u_{t+1}^c}{u_t^c} \right) = \beta_t$.

Investors use this model to forecast capital gains and learn from new information, responding to their uncertainty about equilibrium price formation. This learning process adds an additional source of fluctuations to the model. In particular, the model equilibrium dynamics are now described by three difference equations: the capital law of motion (87), the Euler Equation (91) and the expectations updating equation (94). Then, two feedback loops operate in learning markets. First, the one between the stock and price of capital, which is self-correcting. Second, the price-expectations loop described throughout the paper, which is reinforcing and can drive the economy in waves of over and under capital accumulation.

The expectations loop amplifies the dynamics emerging from the efficient model. To illustrate it, figure 9 plots the response of both the capital stock and price to a transitory productivity shock in the (K_t, Q_t) diagram, starting from the steady state. With efficient markets, an increase in productivity would move the price and stock of capital up for one period, surprising the agents. However, since the displacement is known to be temporary, they find no reason to revise expectations so that the only force at play are lower returns from a higher stock of capital that brings prices down; then, with prices below and the capital stock above their steady state levels, the economy enters a path of gradual disinvestment until reaching the steady state. With learning, the initial price surprise

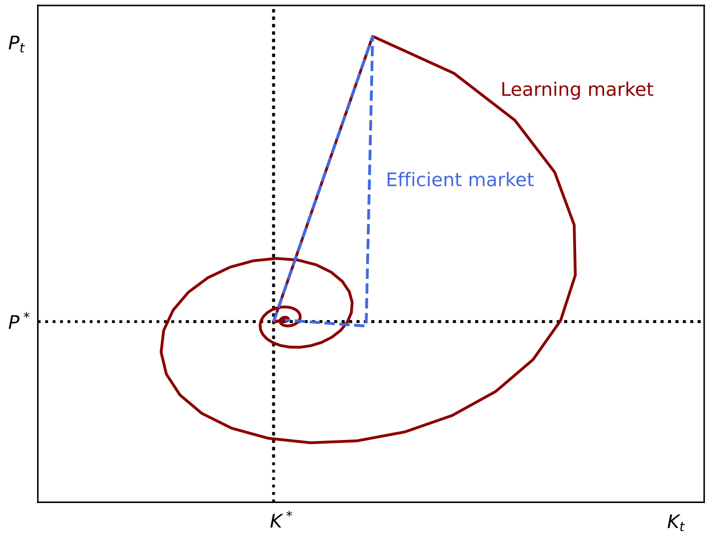

Figure 9: ***Response of capital and capital price to a transitory productivity shock under Rational Expectations and Learning.*** *The graph uses a plane with the capital stock on the x-axis and the capital price on the y-axis. Starting in the steady state, the economy is perturbed by a one-off productivity shock. The blue line shows the response of capital stock and price under efficient pricing (Rational Expectations). The red line shows that response when agents learn.*

leads agents to review their forecast upwards which in turn, raises prices and capital feeding back into a new upward revision. However, there is a counteracting force: as the price boom leads to accumulate capital, returns decline which pushes prices downwards. Eventually, declining capital rents overcome the effect of more optimistic expectations, which are defeated. At that point, the process revert in the form of a bust. It is throughout a sequence of boom and busts, rather than following an smooth saddle path, that the economy goes back to the steady state.

4.4.- A capital gains tax to stabilize inefficient markets

In this section, a tax on unrealized capital gains is introduced and an optimal taxation problem is analyzed. In line with Benigno et al. (2019), the optimal tax is derived to implement the First Best[60].

Capital gains taxation in a production economy. I first modify the household's budget constraint by introducing taxes on capital gains (τ^K):

$$C_t + Q_t K_t \leq F(Z_t, K_{t-1}) + (1-d)Q_t K_{t-1} + \Pi_t - \tau_t^K (Q_t - Q_{t-1})(1-d)K_{t-1} + T_t \qquad (95)$$

[60] Benigno et al. (2019) argue that while the literature on pecuniary externalities focuses on setting the right taxes to implement constrained efficiency, it is possible to use better these instruments and implement the First Best.

Besides, it is assumed the government simply transfers the revenues back in a lump-sum manner, that is,

$$\tau_t^K(Q_t - Q_{t-1})(1-d)K_{t-1} = T_t \tag{96}$$

Then, the Euler Equations becomes

$$Q_t = \delta \mathbb{E}_t^{\mathcal{P}}\left[\frac{u_{t+1}^c}{u_t^c}\left(F_{t+1}^k + (1-\tau_{t+1}^K)(1-d)Q_{t+1} + \tau_{t+1}^K(1-d)Q_t\right)\right] \tag{97}$$

The tax distort the intertemporal incentives by influencing the present value of future payoffs and then the capital price and equilibrium allocations.

Optimal taxation problem. Given K_{-1} and an exogenous productivity process $\{Z_t\}_{t=0}^{\infty}$, the paternalistic planner's problem is to choose both capital gains and lump-sum taxes to deliver the best competitive equilibrium with learning, that is,

$$\max_{\{\tau_t^K, T_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \delta^t U(C_t)$$

s.t. households budget constraint (95); the government budget constraint (96); the capital producers' profits equation; goods and capital market clearing conditions (86, 87); the investment function (83); the households' Euler Equation (97); and the beliefs updating equation (94).

Solution. To replicate the efficient allocations, it is sufficient for the planner to set taxes as to equalize the Euler Equation under Rational Expectations and learning and to transfer the proceeds back to households in a lump-sum manner. If that is possible, prices in the learning world would be the same as under Rational Expectations. In turn, lump-sum taxes would undo the income effect triggered by the capital gains tax, leaving the budget constraint unchanged. Altogether, with prices at the right level and unchanged resources, the allocations will be the efficient ones as the remainder optimality conditions are exactly the same in both worlds. In other words, the only difference with efficient markets is that now there is learning to respond to deal with limited information so that the planner would like to use taxes to undo the effects of that friction. Agents will continue to have imperfect knowledge and learn, but that process would not generate excess price volatility anymore because taxes would avoid the transmission of beliefs deviations from RE to prices and quantities.

The Rational Expectations' Euler Equation can be rewritten as:

$$Q_t^{RE} = \frac{\delta \mathbb{E}_t\left[\frac{u_{t+1}^c}{u_t^c} F_{t+1}^k\right]}{1 - \delta(1-d)\beta_t^*} \tag{98}$$

where $\beta_t^* \equiv \mathbb{E}_t\left[\frac{u_{t+1}^c}{u_t^c}\frac{Q_{t+1}}{Q_t}\right]$. The learning counterpart with taxes reads as

$$Q_t^L = \frac{\delta \mathbb{E}_t^{\mathcal{P}}\left[\frac{u_{t+1}^c}{u_t^c} F_{t+1}^k\right]}{1 - \tau_{t+1}^K \delta(1-d)\mathbb{E}_t^{\mathcal{P}}\left[\frac{u_{t+1}^c}{u_t^c}\right] - (1-\tau_{t+1}^K)\delta(1-d)\beta_t} \tag{99}$$

Then, the market inefficiency under learning, call it X, boils down to the distance between the efficient and the learning price, that is, $X_t = Q_t^{RE} - Q_t^L(\tau_{t+1}^K)$. The optimal taxation problem amounts to find the root of X. In other words, a tax level τ^* is optimal if and only if

$$X_t(\tau^*) = 0 \tag{100}$$

The root of X_t can be written as

$$\tau_{t+1}^* = 1 - \frac{\mathbb{E}_t\left[u_{t+1}^c\Big(P_{t+1} - P_t\Big)\right]}{\mathbb{E}_t^{\mathcal{P}}\left[u_{t+1}^c\Big(P_{t+1} - P_t\Big)\right]} \tag{101}$$

Thus, the optimal tax is a nonlinear function of the deviation of subjective from objective expectations about capital gains (adjusted by wealth's marginal value). Consider the limit case with vanishing risk aversion to derive clear intuition. Then, the previous formula simplifies to

$$\tau_{t+1}^* = 1 - \frac{\beta_t^* - 1}{\beta_t - 1} \tag{102}$$

There are two limit cases that can be derived. First, when subjective beliefs tend toward the objective ones, the optimal tax is zero:

$$\lim_{\beta_t \to \beta_t^*} \tau_{t+1}^* = 0$$

Second, when objective expectations tend to 1, the optimal tax is simply one:

$$\lim_{\beta_t^* \to 1} \tau_{t+1}^* = 1$$

Beyond these cases, the sign of the tax can be defined by parts:[61]

$$\tau_{t+1}^* = \begin{cases} > 0 \text{ if } \begin{cases} \beta_t > \beta_t^* \mid \beta_t > 1 \ (A) \\ \beta_t < \beta_t^* \mid \beta_t < 1 \ (B) \end{cases} \\ < 0 \text{ if } \begin{cases} \beta_t > \beta_t^* \mid \beta_t < 1 \ (C) \\ \beta_t < \beta_t^* \mid \beta_t > 1 \ (D) \end{cases} \end{cases} \tag{103}$$

Intuitively, case A shows that when investors are too optimistic, meaning they expect prices to rise more than justified by fundamentals, capital gains should be taxed. That can be the situation in a typical a boom. Taxes should also be positive when investors are too pessimistic, meaning they expect prices to decrease more than justified by fundamentals (case B). In this case, typical of a burst, taxes on negative capital gains are actually subsidizing capital losses. Hence, in A (B), taxes dampen the upwards (downwards) hike in beliefs.

The formula recommends a negative tax in two scenarios. In case C, investors are not optimistic enough, meaning they expect only a moderate increase in price growth, below what would be reasonable based on fundamentals. Then, investors would be actually subsidized to boost their optimism. In case D, investors are not pessimistic enough, meaning they expect only a soft reduction in price growth, below what Rational Expectations investors would forecast. Then, a negative tax on negative expected capital losses would take resources from investors, aiming at making them expecting more losses until anchoring their beliefs at their fundamental value. Figure 10 illustrates these four cases.

The optimal tax inherits the subjective expectations dynamics. By the learning updating rule, $\beta_t = \beta(\beta_{t-1}, \beta_{t-2}, \tau_{t-1}, \tau_{t-2}, \cdot)$ shows high serial correlation (for small gains). In turn, β_t^* is a function of the states (Z_t, K_{t-1}), both obeying AR(1) process. It follows that optimal taxes would display high serial correlation. Yet, subjective beliefs β_t deviates from RE quite substantially which may generate big movements in taxes. In fact, the optimal tax is unbounded and then, in some cases, the tax might reach values well beyond $\pm$ 1. From a policy standpoint, that is probably an important shortcoming; next section deals with it.

[61] Cases in which $\beta_t^* = 1$ or $\beta_t = 1$ are ignored; the first because leads to a tax equal to 1 as already pointed out; the second because it yields an undefined fraction.

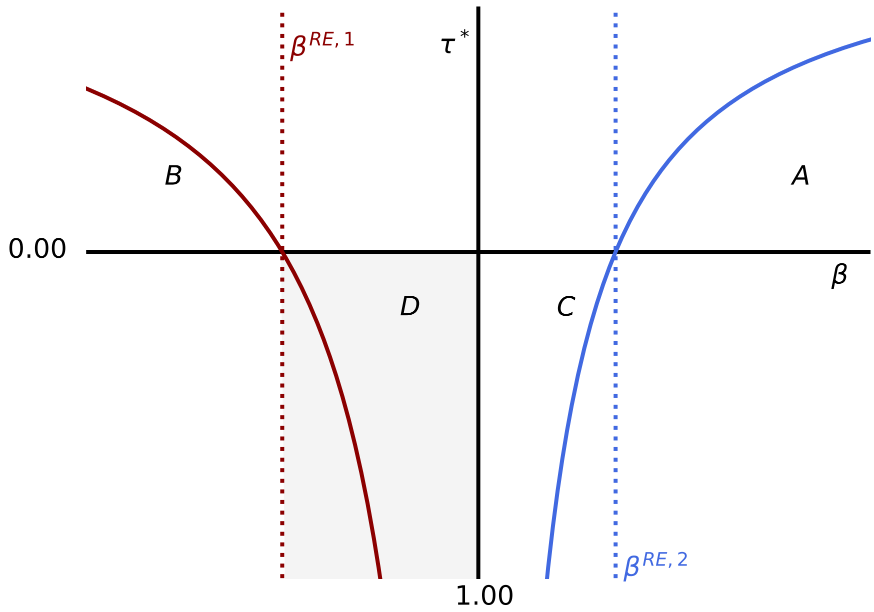

Figure 10: ***Optimal capital gains tax.*** *The figure shows the optimal capital gains tax τ^* as a function of subjective expectations β_t for two different values of objective expectations β^{RE}, one above and one below 1. Letters signal the 4 four cases highlighted in expression (103). Shadowed areas sets out the cases in which investors expect to pay taxes, when expected capital gains are positive (negative) and the tax is positive (negative).*

4.5.- An alternative implementation

In this section, an alternative implementation of the optimal policy is presented. It uses a CGT to eradicate the influence of subjective beliefs on prices and a subsidy on capital rents to avoid chronic under-investment. To great extent, this combination avoids tax volatility.

A CGT equal to 100% can eliminate the influence of beliefs on prices. To illustrate why, I use a decomposition of total volatility between fundamental and non-fundamental. Following the procedure in Section 3, the variance of the capital price can be approximated by

$$\mathbb{V}ar(Q_t) \approx \underbrace{z^2\mathbb{V}ar(Z_t) + k^2\mathbb{V}ar(K_{t-1})}_{\text{Fundamental}} \quad + \quad \underbrace{b^2\mathbb{V}ar(\beta_t)}_{\text{Non-Fundamental Volatility} \equiv \mathcal{V}} \tag{104}$$

where $x = \partial Q_t/\partial X_t$ evaluated at the approximation point for $x = z, k, b$. Then, the optimal tax must satisfy

$$\tau^* \iff \mathcal{V}(\tau^*) = 0 \tag{105}$$

Note that the two objects in $b\mathbb{V}ar(\beta_t)$ depend on taxes. Then, finding a τ that makes $b = 0$

would be a sufficient condition. Since

$$b = \frac{\partial Q_t}{\partial \beta_t}(Z^*, K^*, 1) = \frac{\delta^2 F(Z^*, K^*)(1-d)(1-\tau^K_{t+1})}{(1-\delta(1-d)\tau^K_{t+1} - \delta(1-d)(1-\tau^K_{t+1}))^2} \tag{106}$$

with $f(Z^*, K^*)$ being $\mathbb{E}_t(F^k_{t+1})$ evaluated at the approximation point, it turns out that a tax equal to 1 eliminate the externality

$$\tau^* = 1 \iff b = 0 \;\Rightarrow\; \mathcal{V} = 0$$

When $\tau^K_{t+1} = 1$, the equilibrium capital price becomes

$$Q^L_t = \frac{\delta \mathbb{E}_t\left[\frac{u^c_{t+1}}{u^c_t} F^k_{t+1}\right]}{1-\delta(1-d)} \tag{107}$$

which is exactly the price under Rational Expectations when $\beta^*_t = 1$. In other words, this derivation reaches the same conclusion as before but in the other direction: an optimal tax equal to 1 generates a price equivalent to the efficient when no capital gains are expected.

Importantly, $\tau^* = 1$ does not imply a trivial solution consisting of correcting non-fundamental volatility by killing also fundamental volatility. Thus,

$$\lim_{\tau^K \to 1} x = \tilde{x} > 0$$

for $x = z, k$. Put it differently, the volatility implementation of the optimal tax is in the spirit of the so-called "Principle of Targetting" of Pigouvian taxation (see Dixit (1985)), according to which a corrective tool has to tax directly the source of the externality. In this case, the direct source of the externality is the excessive volatility of capital gains expectations and thus, a tax on capital gains is directly related to it.

The main shortcoming of this approximation is that it might deliver a too low capital price and then, chronic sub-investment. The question is whether this can be compensated by a new instrument, since lump-sum taxes cannot affect the capital price. A tax on capital rents might be a natural alternative. Thus, suppose the government can tax capital profits with τ^r. With this new instrument, equation (107) becomes

$$Q_t = \frac{\delta(1-\tau^r_{t+1})\mathbb{E}_t\left[\frac{u^c_{t+1}}{u^c_t} F^k_{t+1}\right]}{1-\delta(1-d)} \tag{108}$$

Hence, by setting

$$(\tau^r_{t+1})^* = 1 - \frac{1-\delta(1-d)}{1-\delta(1-d)\beta^*_t} \tag{109}$$

Thus, if RE implies an almost constant capital gains expectations, $(\tau^r_{t+1})^*$ would be almost constant and then, the First Best can be implemented by a constant τ^K and a not-too volatile τ^r along with lump-sum taxes.[62] Altogether, the alternative implementation offers a way of stabilizing capital markets avoiding excessive volatility in taxes and relaxing the informational requirements.

5.- Conclusions

This paper has analyzed how a Capital Gains Tax influence asset price cycles. Differently than commonly thought, I show that a CGT stabilize price by its sensitivity to belief fluctuations, using an asset-pricing model with learning about prices.

The model provides an explanation for the simultaneous rise in stock price volatility despite the decline in macroeconomic risk. Precisely, tax cuts would increase the influence of subjective beliefs on stock prices, boosting self-fulfilling booms. Importantly, this observation is explained along with a long list of asset pricing facts, including part of the level and rise of the PD ratio, the equity premium and the comovement between expectations and prices. To the best of our knowledge, this is the first model that explains all of them.

Furthermore, the last part of the paper has explored the usage of taxes to correct excess price volatility stemming from investors' information limitations. While subjective beliefs are crucial in explaining stock market volatility, the excess volatility they cause can be seen as a pecuniary externality that can cause undesirable real fluctuations. In such a case, a tax on unrealized capital gains that corrects too optimistic/pessimistic beliefs proves able to restore the First Best.

Altogether, the arguments developed in the paper suggest that a CGT can be an effective tool to prevent asset price booms and the financial and macroeconomic fluctuations associated to them. Thus, while the ability of a Financial Transaction Tax to prevent excess price volatility has been widely questioned, a CGT emerges as a sound alternative.

The research has left some issues opened. On the empirical side, the analysis has focused on the US aggregate stock market leaving cross-sectional analysis unexplored. Although the effect of capital gains tax on the cross-section of stocks was analyzed by Dai et al. (2013) for two tax reforms it would be interesting to expand the analysis using larger time windows. Moreover, the decline in capital taxes since the 1980s was a global phenomenon. An international analysis of its

[62] $(\tau^r_{t+1})^*$ would be less volatile than τ^*_{t+1} as long as β^*_t is more stable than β_t.

effects on capital and real markets and its interaction with financial deregulation and capital flows liberalization appears as an interesting research avenue.

Theoretically, the model has largely abstracted from the lock-in effect and the influence of a CGT on trading. The effects of a CGT on a richer setup that integrates asset pricing and trading dynamics are to be done. Another important absence in the model is credit. Different instruments have been suggested to deal with excessive credit cycles. It remains unknown whether a CGT would be sufficient to prevent joint cycles of expectations, asset prices and credit. Finally, I have taken payout policies as given, ignoring the possible reaction of firms to tax changes and its impacts on investment, productivity or employment.

From a broader standpoint, the optimal capital taxation literature has not considered the use of capital gains taxes so far due to their focus on one-sector models.[63] Thus, the optimal use of capital gains taxes to fund government spending is to be explored. Finally, it is well known that capital gains have important redistributive implications.[64] The analysis in this paper would suggest that a CGT could help not only ex-post (i.e., redistributing capital gains) but even ex-ante (i.e. avoiding part of the wealth inequality that comes from asset price dynamics).

[63] Not even the recent work of Chari et al. (2020) that includes a tax rich system with taxes on dividends, capital rents or wealth.

[64] See Fagereng et al. (2022) for a recent analysis.

References

Adam, K. (2020). Monetary policy challenges from falling natural interest rates. Technical report, ECB conference.

Adam, K., Kuang, P., and Marcet, A. (2012). House price booms and the current account. *NBER Macroeconomics Annual*, 26(1):77–122.

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.

Adam, K. and Merkel, S. (2019). Stock price cycles and business cycles.

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

Anagnostopoulos, A., Cárceles-Poveda, E., and Lin, D. (2012). Dividend and capital gains taxation under incomplete markets. *Journal of Monetary Economics*, 59(7):599–611.

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.

Benigno, G., Chen, H., Otrok, C., Rebucci, A., and Young, E. R. (2019). Optimal policy for macro-financial stability. Technical report, National Bureau of Economic Research.

Boldrin, M., Christiano, L. J., and Fisher, J. D. (2001). Habit persistence, asset returns, and the business cycle. *American Economic Review*, 91(1):149–166.

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.

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.

Chari, V. V., Nicolini, J. P., and Teles, P. (2020). Optimal capital taxation revisited. *Journal of Monetary Economics*, 116:147–165.

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

Christiano, L. J. and Fisher, J. D. (2000). Algorithms for solving dynamic models with occasionally binding constraints. *Journal of Economic Dynamics and Control*, 24(8):1179–1232.

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.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78.

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. (2020). Optimal financial transaction taxes. Technical report, National Bureau of Economic Research.

Dávila, E. and Korinek, A. (2018). Pecuniary externalities in economies with financial frictions. *The Review of Economic Studies*, 85(1):352–395.

Di Tella, S. (2019). Optimal regulation of financial intermediaries. *American Economic Review*, 109(1):271–313.

Dixit, A. (1985). Tax policy in open economies, chapter 6 in a. auerbach and m. feldstein (edited): Handbook of public economics, vol. 1.

Engle, R. F., Ghysels, E., and Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95(3):776–797.

Engle, R. F. and Rangel, J. G. (2008). The spline-garch model for low-frequency volatility and its global macroeconomic causes. *The Review of Financial Studies*, 21(3):1187–1222.

Fagereng, A., Gomez, M., Gouin-Bonenfant, E., Holm, M., Moll, B., and Natvik, G. (2022). Asset-price redistribution. *Manuscript*.

Farhi, E. and Werning, I. (2016). A theory of macroprudential policies in the presence of nominal rigidities. *Econometrica*, 84(5):1645–1704.

Farhi, E. and Werning, I. (2020). Taming a minsky cycle. *Unpublished manuscript, Harvard University*.

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

Forni, M., Gambetti, L., Marco, M., and Sala, L. (2020). Common component structural vars.

Gambacorta, L. and Signoretti, F. M. (2014). Should monetary policy lean against the wind?: An analysis based on a dsge model with banking. *Journal of Economic Dynamics and Control*, 43:146–174.

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.

Gopinath, G., Kalemli-Özcan, Ş., Karabarbounis, L., and Villegas-Sanchez, C. (2017). Capital allocation and productivity in south europe. *The Quarterly Journal of Economics*, 132(4):1915–1967.

Gourio, F. and Miao, J. (2011). Transitional dynamics of dividend and capital gains tax cuts. *Review of Economic Dynamics*, 14(2):368–383.

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

Hall, R. E. (2017). High discounts and high unemployment. *American Economic Review*, 107(2):305–30.

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.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica: Journal of the Econometric Society*, pages 213–224.

Ifrim, A. (2021). The fed put and monetary policy: An imperfect knowledge approach. *Available at SSRN 3921072*.

Jeanne, O. and Korinek, A. (2010). Excessive volatility in capital flows: A pigouvian taxation approach. *American Economic Review*, 100(2):403–07.

Jeanne, O. and Korinek, A. (2019). Managing credit booms and busts: A pigouvian taxation approach. *Journal of Monetary Economics*, 107:2–17.

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

Kuhn, M., Schularick, M., and Steins, U. I. (2020). Income and wealth inequality in america, 1949–2016. *Journal of Political Economy*, 128(9):3469–3519.

Kurlat, P. (2018). How i learned to stop worrying and love fire sales. 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.

Lorenzoni, G. (2008). Inefficient credit booms. *The Review of Economic Studies*, 75(3):809–833.

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

Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in igarch (1, 1) and covariance stationary garch (1, 1) models. *Econometrica: Journal of the Econometric Society*, pages 575–596.

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.

Mehra, R. and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of monetary Economics*, 15(2):145–161.

Nisticò, S. (2012). Monetary policy and stock-price dynamics in a dsge framework. *Journal of Macroeconomics*, 34(1):126–146.

Romer, C. D. and Romer, D. H. (2010). The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. *American Economic Review*, 100(3):763–801.

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

Sarin, N., Summers, L., Zidar, O., and Zwick, E. (2022). Rethinking how we score capital gains tax reform. *Tax Policy and the Economy*, 36(1):1–33.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance*, 44(5):1115–1153.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436.

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. (1975). The effects of income, wealth, and capital gains taxation on risk-taking. *Stochastic Optimization Models in Finance*, pages 291–311.

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

Taylor, J. B. (2007). Housing and monetary policy. Technical report, National Bureau of Economic Research.

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. (1969). A general equilibrium approach to monetary theory. *Journal of money, credit and banking*, 1(1):15–29.

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

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of monetary economics*, 24(3):401–421.

Winkler, F. (2020). The role of learning for asset prices and business cycles. *Journal of Monetary Economics*, 114:42–58.

Wright, B. D. and Williams, J. C. (1982a). The economic role of commodity storage. *The Economic Journal*, 92(367):596–614.

Wright, B. D. and Williams, J. C. (1982b). The roles of public and private storage in managing oil import disruptions. *The Bell Journal of Economics*, pages 341–353.

Wright, B. D. and Williams, J. C. (1984). The welfare effects of the introduction of storage. *The Quarterly Journal of Economics*, 99(1):169–192.

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: Proof of the Proposition.

Take a linear approximation of equation (14) around the Rational Expectations value (i.e., $\beta_t = \beta^D$):

$$\frac{P_t}{D_t} \approx \frac{P_t^{RE}}{D_t} + \omega(\beta_t - \beta^D) \tag{110}$$

with $\frac{P_t^{RE}}{D_t} = \frac{\delta(1-\tau^D)\beta^D}{1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K}$ being a constant and ω evaluated at the approximation point. Taking the variance of both sides

$$\mathbb{V}ar\left[\frac{P_t}{D_t}\right] \approx \omega^2 \times \mathbb{V}ar(\beta_t) \tag{111}$$

as claimed in point i). Assume τ^D and $\pi\tau^K$ are both within the [0,1) interval. Then,

$$\omega = \frac{\partial P_t/D_t}{\partial \beta_t}\Big|_{\beta_t=\beta^D} = \frac{\delta^2(1-\tau^D)\beta^D(1-\pi\tau^K)}{(1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K)^2} > 0 \tag{112}$$

Differentiating (112) with respect to τ^K

$$\frac{d\omega}{d\tau^K} = \frac{-\delta^2(1-\tau^D)\beta^D\pi(1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K) - \delta^2(1-\tau^D)\beta^D(1-\pi\tau^K)\delta\pi(\beta^D-1)}{(1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K)^2} \tag{113}$$

The numerator boils down to $-\delta^2(1-\tau^D)\beta^D\pi(1-\delta)$. Since $\delta < 1$, the numerator is negative, $\frac{d\omega}{d\tau^K} < 0$ and $\frac{d\omega^2}{d\tau^K} < 0$ as claimed.

Applying the chain rule to expression (111),

$$\frac{d\mathbb{V}ar[P_t/D_t]}{d\tau^K} \approx \frac{\partial \mathbb{V}ar[P_t/D_t]}{\partial \mathbb{V}ar[\beta_t]} \frac{d\mathbb{V}ar[\beta_t]}{d\tau^K} + \frac{\partial \mathbb{V}ar[P_t/D_t]}{\partial \omega^2} \frac{d\omega^2}{d\tau^K}$$

Note $\frac{\partial \mathbb{V}ar[P_t/D_t]}{\partial \mathbb{V}ar[\beta_t]} = \omega^2 > 0$ and $\frac{\partial \mathbb{V}ar[P_t/D_t]}{\partial \omega^2} = \mathbb{V}ar[\beta_t] > 0$, provided the latter exists. Given, $\frac{d\omega^2}{d\tau^K} < 0$, the only thing left is to characterize $\mathbb{V}ar[\beta_t]$ and prove $\frac{d\mathbb{V}ar[\beta_t]}{d\tau^K} < 0$.

Start with the characterization of $\mathbb{V}ar[\beta_t]$. Equation (15) can be linearly approximated as an AR(2) process around $(\beta_{t-1}, \beta_{t-2}, \varepsilon^D_{t-1}) = (\beta^D, \beta^D, 1)$:

$$\beta_t \approx \beta^D + \mathcal{A}(\beta_{t-1} - \beta^D) + \mathcal{B}(\beta_{t-2} - \beta^D) + \mathcal{C}(\varepsilon^D_{t-1} - 1) \tag{114}$$

with

$$\mathcal{A} = \frac{\partial \beta_t}{\partial \beta_{t-1}}\Big|_{\substack{\beta_{t-1}=\beta^D \\ \beta_{t-2}=\beta^D \\ \varepsilon^D_{t-1}=1}} = 1 - g + \frac{g\delta\beta^D(1-\pi\tau^K)}{1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K} \tag{115}$$

$$\mathcal{B} = \frac{\partial \beta_t}{\partial \beta_{t-2}}\Big|_{\substack{\beta_{t-1}=\beta^D \\ \beta_{t-2}=\beta^D \\ \varepsilon^D_{t-1}=1}} = -\frac{g\delta\beta^D(1-\pi\tau^K)}{1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K} \tag{116}$$

$$\mathcal{C} = \frac{\partial \beta_t}{\partial \varepsilon^D_t}\Big|_{\substack{\beta_{t-1}=\beta^D \\ \beta_{t-2}=\beta^D \\ \varepsilon^D_{t-1}=1}} = g\beta^D \tag{117}$$

If $\{\beta_t\}$ is stationary, it would have the following variance

$$\mathbb{V}ar(\beta_t) \approx \frac{(1-\mathcal{B})\mathcal{C}^2\sigma^2_D}{(1+\mathcal{B})(1-\mathcal{A}-\mathcal{B})(1+\mathcal{A}-\mathcal{B})} \tag{118}$$

Now I verify the conditions that ensure $\{\beta_t\}$ is stationary. It is known that for the process to be stationary, parameters $\mathcal{A}$, $\mathcal{B}$ must lie within the region $-1 < \mathcal{B} < 1$, $\mathcal{A} + \mathcal{B} < 1$, $\mathcal{B} - \mathcal{A} < 1$. I proceed to verify these conditions. Note $\mathcal{B} < 0$ provided $\pi\tau^K < 1$. $\mathcal{B}$ must also satisfy $\mathcal{B} = -\frac{g\delta(1-\pi\tau^K)\beta^D}{1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K} > -1$. Rearranging the terms, $1 - \delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K > g\delta(1-\pi\tau^K)\beta^D$. The left hand side is positive by assumption A. The right hand side is also positive given $\pi\tau^K < 1$. Then, a positive but small enough gain is sufficient for the inequality to hold. In particular,

$$g < \frac{1-\delta(1-\pi\tau^K)\beta^D - \delta\pi\tau^K}{\delta(1-\pi\tau^K)\beta^D} \equiv \bar{g} \tag{119}$$

The next condition is $\mathcal{A} + \mathcal{B} < 1$. Using expression (115) and (116), the condition boils down to $\mathcal{A}+\mathcal{B} = 1-g+\frac{g\delta\beta^D(1-\pi\tau^K)}{1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K} - \frac{g\delta\beta^D(1-\pi\tau^K)}{1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K} = 1-g$. Thus, $\mathcal{A}+\mathcal{B}<1$ holds if

$$g > 0 \tag{120}$$

Conditions (119) and (120) requires there is some learning, but not too much. The final condition is $\mathcal{B}-\mathcal{A}<1$. Since $\mathcal{B}<0$, proving $\mathcal{A}>0$ is enough. It requires $1 > g\Big(1-\frac{\delta\beta^D(1-\pi\tau^K)}{1-\delta\beta^D(1-\pi\tau^K)-\delta\pi\tau^K}\Big)$. Since $g>0$, $\frac{\delta\beta^D(1-\pi\tau^K)}{1-\delta\beta^D(1-\pi\tau^K)-\delta\pi\tau^K} > 1$ is sufficient such that the element within the parenthesis is non-positive. Rearranging the previous inequality, $0 \geq 1-2\delta\beta^D(1-\pi\tau^K)-\delta\pi\tau^K$. The following expression ensures the last inequality to hold

$$\pi\tau^K \leq \frac{2\delta\beta^D-1}{2\delta\beta^D-\delta} \equiv \bar{\tau} \tag{121}$$

Since $\delta<1$, $\bar{\tau}<1$, compatible with $\pi\tau^K<1$. Thus, a not too high effective tax is enough to ensure $\mathcal{A}>0$.

The next step is to prove $\mathbb{V}ar[\beta_t]$ is decreasing on τ^K. Using the definition of $\mathcal{A},\mathcal{B}$ and the chain rule it turns out

$$\frac{d\mathbb{V}ar(\beta_t)}{d\tau^K} = \frac{\partial\mathbb{V}ar(\beta_t)}{\partial\mathcal{A}}\frac{d\mathcal{A}}{d\tau^K} + \frac{\partial\mathbb{V}ar(\beta_t)}{\partial\mathcal{B}}\frac{d\mathcal{B}}{d\tau^K} \tag{122}$$

Now I show $\frac{d\mathbb{V}ar(\beta_t)}{d\tau^K}<0$ holds. First,

$$\frac{\partial\mathbb{V}ar(\beta_t)}{\partial\mathcal{A}} = -\frac{2(\mathcal{B}-1)\mathcal{C}^2\sigma_D^2\mathcal{A}}{(\mathcal{B}+1)(\mathcal{A}-\mathcal{B}+1)^2(\mathcal{A}+\mathcal{B}-1)^2} \tag{123}$$

Given $-1<\mathcal{B}<1$ holds, $(\mathcal{B}-1)<0$ and $\mathcal{B}+1>0$. $\mathcal{A}$ is also positive. Then, this derivative has a positive sign.

Now, note $\mathcal{A}$ is decreasing on τ^K

$$\frac{d\mathcal{A}}{d\tau^K} = \frac{-g\delta\beta^D\pi(1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K)-g\delta\beta^D\pi(1-\pi\tau^K)\delta(\beta^D-1)}{(1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K)^2} < 0 \tag{124}$$

since the numerator boils down to $-g\delta\beta^D\pi(1-\delta)$ and $\delta<1$.

Now, check the effects of τ^K through $\mathcal{B}$. The derivative of the variance with respect to $\mathcal{B}$ is given by

$$\frac{\partial\mathbb{V}ar(\beta_t)}{\partial\mathcal{B}} = \frac{2\mathcal{C}^2\sigma_D^2(\mathcal{B}^3-2\mathcal{B}^2+\mathcal{B}+\mathcal{A}^2)}{(\mathcal{B}+1)^2(\mathcal{A}-\mathcal{B}+1)^2(\mathcal{A}+\mathcal{B}-1)^2} \tag{125}$$

Since the denominator and $2\mathcal{C}^2\sigma_D^2$ are both positive, the sign is determined by $(\mathcal{B}^3 - 2\mathcal{B}^2 + \mathcal{B} + \mathcal{A}^2)$. Since $\mathcal{B} - \mathcal{A} < 1$, $(\mathcal{B}-1)^2 < \mathcal{A}^2$. Then, $(\mathcal{B}^3 - 2\mathcal{B}^2 + \mathcal{B} + (\mathcal{B}-1)^2) > 0$ implies $(\mathcal{B}^3 - 2\mathcal{B}^2 + \mathcal{B} + \mathcal{A}^2) > 0$. Note $(\mathcal{B}^3 - 2\mathcal{B}^2 + \mathcal{B} + (\mathcal{B}-1)^2) > 0$ is equivalent to $\mathcal{B}^3 - \mathcal{B}^2 - \mathcal{B} + 1 > 0$ which holds for $-1 < \mathcal{B} < 1$ and $\mathcal{B} > 1$. Hence, within the stationarity region $-1 < \mathcal{B} < 1$, the expression is positive which makes the derivative positive as well.

The effect of τ^K on $\mathcal{B}$ is given by

$$\frac{d\mathcal{B}}{d\tau^K} = \frac{g\delta\beta^D\pi(1 - \delta(1 - \pi\tau^K)\beta^D - \delta\pi\tau^K) + g\delta\beta^D\pi(1 - \pi\tau^K)\delta(\beta^D - 1)}{(1 - \delta(1 - \pi\tau^K)\beta^D - \delta\pi\tau^K)^2} \tag{126}$$

which is positive for the same reasons that $\frac{d\mathcal{A}}{d\tau^K} < 0$.

Altogether, for $\frac{d\mathbb{V}ar(\beta_t)}{d\tau^K} < 0$ it must be

$$\underbrace{\frac{\partial \mathbb{V}ar(\beta_t)}{\partial \mathcal{A}}}_{>0} \underbrace{\frac{d\mathcal{A}}{d\tau^K}}_{<0} + \underbrace{\frac{\partial \mathbb{V}ar(\beta_t)}{\partial \mathcal{B}}}_{>0} \underbrace{\frac{d\mathcal{B}}{d\tau^K}}_{>0} < 0 \tag{127}$$

Since $\frac{\partial \mathcal{A}}{\partial \tau^K} = -\frac{\partial \mathcal{B}}{\partial \tau^K}$, inequality (127) boils down to $\frac{\partial \mathbb{V}ar(\beta_t)}{\partial \mathcal{A}} > \frac{\partial \mathbb{V}ar(\beta_t)}{\partial \mathcal{B}}$, that is, $-\frac{2(\mathcal{B}-1)\mathcal{C}^2\sigma_D^2\mathcal{A}}{(\mathcal{B}+1)(\mathcal{A}-\mathcal{B}+1)^2(\mathcal{A}+\mathcal{B}-1)^2} > \frac{2\mathcal{C}^2\sigma_D^2(\mathcal{B}^3-2\mathcal{B}^2+\mathcal{B}+\mathcal{A}^2)}{(\mathcal{B}+1)^2(\mathcal{A}-\mathcal{B}+1)^2(\mathcal{A}+\mathcal{B}-1)^2}$. Simplifying the previous expression, one gets to $-(\mathcal{B}-1)\mathcal{A} > \frac{\mathcal{B}^3-2\mathcal{B}^2+\mathcal{B}+\mathcal{A}^2}{\mathcal{B}+1}$. It can be shown this inequality holds for $-1 < \mathcal{B} < 1$ and $\mathcal{B} - \mathcal{B}^2 < \mathcal{A} < 1 - \mathcal{B}$. Note $-1 < \mathcal{B} < 1$ and $\mathcal{A} < 1 - \mathcal{B}$ are stationary conditions already proven; $\mathcal{B} - \mathcal{B}^2 < \mathcal{A}$ must be proven. Using the definition of $\mathcal{A}, \mathcal{B}$, it turns out $\mathcal{A} = 1 - g - \mathcal{B}$. Using this equality, $\mathcal{B} - \mathcal{B}^2 < 1 - g - \mathcal{B}$ or $0 < 1 - g + \mathcal{B}^2 - 2\mathcal{B}$. Intuitively, $\mathcal{B}^2$ and $-2\mathcal{B}$ are positive so that, the inequality would hold if g is not too high. In particular, given $g > 0$, the inequality holds if $\mathcal{B} < 1 - g^{0.5}$ which given $\mathcal{B} < 0$ is true if $g^{0.5} < 1$. To check this last inequality, use the upper bound $\bar{g} > g$, that is, check if $\bar{g}^{0.5} = \left(\frac{1-\delta(1-\pi\tau^K)\beta^D-\delta\pi\tau^K}{\delta(1-\pi\tau^K)\beta^D}\right)^{0.5} < 1$. It boils down to $\delta(1 - \pi\tau^K)\beta^D > 1 - \delta(1 - \pi\tau^K)\beta^D - \delta\pi\tau^K$ which has been already shown to hold provided $\pi\tau^K < \bar{\tau}$. Hence, $\frac{d\mathbb{V}ar(\beta_t)}{d\tau^K} < 0$ as claimed in the proposition.

Appendix C: Capital Gains Tax stabilization properties under RE

Assume that dividends growth contains a persistent component x_t in the spirit of Bansal and Yaron (2004):

$$\ln D_t = \ln\beta^D + \ln D_{t-1} + \varphi \ln x_t + \ln\varepsilon_t^D \tag{128}$$

$$\ln x_t = \ln x_{t-1} + \ln \vartheta_t^x \tag{129}$$

with $\ln\varepsilon_t^D \sim i.i.\mathcal{N}(-\frac{\sigma_D^2}{2}, \sigma_D^2)$ and $\ln\vartheta_t^x \sim i.i.\mathcal{N}(-\frac{\varphi s_x^2}{2}, s_x^2)$. In this setup, the following proposition holds:

Proposition: A Capital Gains Tax can stabilize the PD ratio under Rational Expectations. *Assume that investors have perfect information, including the dividends stochastic process (128) and (129), and agents' homogeneity. Then, the variance of the PD ratio is decreasing on the CGT level, that is,*

$$\frac{d\mathbb{V}ar\left[P_t^{RE}/D_t\right]}{d\tau^K} < 0 \tag{130}$$

Proof. In this setup, forward iteration on the Euler Equation (6), the Law of Iterated Expectations and a transversality condition delivers the following equilibrium PD ratio

$$\frac{P_t^{RE}}{D_t} = \frac{\delta(1-\tau^D)\beta^D x_t^{\varphi}}{1-\delta(1-\pi\tau^K)\beta^D x_t^{\varphi} - \delta\pi\tau^K} \tag{131}$$

since $\mathbb{E}_t\left[\frac{D_{t+j}}{D_t}\right] = (\beta^D x_t^{\varphi})^j$. Following the same steps as in the proposition in the main text, the unconditional variance of the PD ratio can by approximated around $x_t = 1$ as

$$\mathrm{v} = \mathbb{V}ar\left[\frac{P_t^{RE}}{D_t}\right] \approx \omega^2 \times \mathbb{V}ar(x_t) \tag{132}$$

It turns out

$$\frac{d\mathbb{V}ar\left[P_t^{RE}/D_t\right]}{d\tau^K} = \frac{\partial\mathbb{V}ar\left[P_t^{RE}/D_t\right]}{\partial\omega^2}\frac{d\omega^2}{d\tau^K} < 0$$

since $\frac{\partial\mathbb{V}ar\left[P_t^{RE}/D_t\right]}{\partial\omega^2} = \mathbb{V}ar(x_t) > 0$ and $\frac{d\omega^2}{d\tau^K} < 0$. ■

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 12 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 11 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.

Figure 11: ***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.*

Table 12: ***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 $(\text{NPI CE + MF})_{1987}$/ $(\text{HHNPI CE + MF})_{1987}$
	= 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

Appendix E: Projection facility

The equilibrium PD ratio given by 131 faces a discontinuity. For this reason, simulation requires to set up the following modified belief updating equation to ensure non-negative prices

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

where

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

and

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

(135)

for $q = L, U$. Thus, this projection facility starts to dampen belief coefficients that imply a price-dividend ratio equal to PD^L and sets an effective upper bound at PD^U. Projection facilities are usual devices in this sort of algorithms (see Ljung (1977)); particularly, (134) is similar to the one used by Adam 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

The proposed approximating function for the consumption policy is

$$\frac{C_t^*}{D_t} = \bar{\mathcal{E}}(\boldsymbol{X}_t) \approx \psi(\boldsymbol{X}_t;\chi) = c_t Z_t \tag{136}$$

where $c_t \equiv (1 - \chi\delta(1-\pi\tau_t^K)\beta_t)$ is the time-varying propensity to consume, Z_t stands for agent's current resources. 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

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

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

$$\beta^{\mathcal{P}}_{t+1} = \beta^i_t \nu_{t+1} \Rightarrow \left(\frac{P_{t+1}}{P_t}\right)^{\mathcal{P}} = \beta^i_t \nu_{t+1} \varepsilon^p_{t+1} \Rightarrow \left(\frac{P_{t+1}}{D_{t+1}}\right)^{\mathcal{P}} = \left(\frac{P_{t+1}}{P_t}\right)^{\mathcal{P}} \frac{D_t}{D_{t+1}} \frac{P_t}{D_t}$$

In turn, expected consumption reads

$$\frac{C^{\mathcal{P}}_{t+1}}{D_{t+1}} = (1-\chi\delta(1-\pi\tau^K_{t+1})\beta^{\mathcal{P}}_{t+1}) \left(\left(\frac{P_{t+1}}{D_{t+1}}\right)^{\mathcal{P}} + 1 - \tau^D_{t+1} + \frac{W_{t+1}}{D_{t+1}} - \pi\tau^K_{t+1} \left[\left(\frac{P_{t+1}}{D_{t+1}}\right)^{\mathcal{P}} - \frac{P_t}{D_t}\frac{D_t}{D_{t+1}} \right] \right) \tag{137}$$

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

$$\chi^{j+1} = \chi^j + d(G(\chi^j) - \chi^j) \tag{138}$$

where j iteration number and d the dampening parameter.

To evaluate how good is the approximation, I explore the errors size. The approximating errors are given by

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

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

$$J = \log_{10}\left(\mathbb{E}_t \left| \frac{u_{t+1}}{\frac{C_t}{D_t}} \right| \right) \tag{139}$$

being J a dimension-free quantity expressing that error as a fraction of current consumption, which expresses the results in economic terms. For the baseline model, J = -5.99. It is equivalent to a

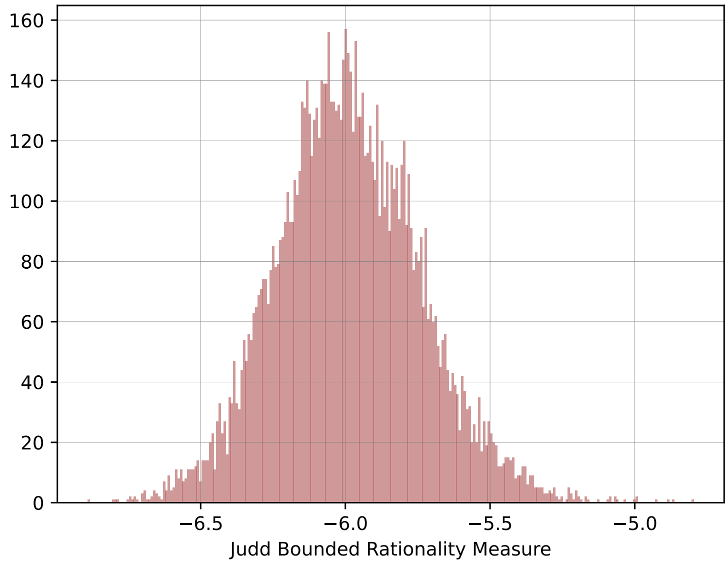

Figure 12: ***Histogram of the Judd Bounded Rationality Measure.*** *The histogram plots the Judd criterion as defined by equation (139) resulting from 10.000 simulations of the model.*

mistake of $1 out of a million. The Mean Square Error is 5.71e-06. Figure 12 plots the histogram of J for 10.000 simulations of the model.

Appendix G: Time-varying elasticity model

To compute the expectations elasticity of the price-dividend ratio I use the following model:

$$\begin{aligned} \ln\frac{P_t}{D_t} &= B + \psi_t \ln\beta_t^h + v_t^p \\ \psi_t &= \psi_{t-1} + w_t^{\psi} \end{aligned} \tag{140}$$

with jointly Normal residuals given by

Therefore, the model has a total of four parameters $\mathcal{A} = \{B, \psi_0, \sigma^2_{v^p}, \sigma^2_{w^\psi}\}$. If they are known, the unknown time-varying elasticity can be estimated using a Kalman filter. The model can be estimate by Maximum Likelihood

$$\log\mathcal{L}(\ln PD_t|\mathcal{A}) = -\frac{1}{2}\sum_{t=1}^{T}\left[log(2\pi) + logV_t + \frac{(\ln PD_t - B - \psi_t \ln\beta_t^h)^2}{V_t}\right] \tag{141}$$

where the prediction error variance is given by

$$V_t \equiv Var_{t-1}[\ln PD_t] = H_t \Sigma_t^{\psi} H_t' + \sigma^2_{v^P}$$

the vector of state coefficients by

$$H_t = [B, \ln\beta_t^h]$$

and the Kalman uncertainty by

$$\Sigma_t^{\psi} \equiv Var_{t-1}[\psi_t] = \Sigma_{t-1}^{\psi} + \sigma^2_{w^{\psi}} - \Sigma_{t-1}^{\psi} H_{t-1}(H_{t-1}\Sigma_{t-1}^{\psi} H_{t-1}')^{-1} H_{t-1}\Sigma_{t-1}^{\psi}$$

with $\Sigma_0^{\psi} = I$ (alternative initial conditions do not change the results). Table 13 reports the estimation results.

Table 13: ***Time-varying elasticity model estimation results.*** *The table shows the MLE of the parameters of the state-space model for expectations elasticity of the price-dividend, as defined by equation 140. Standard errors have been obtained via bootstrapping with 1000 repetitions. The data used for the estimation covers the 1954:I-2012:I period.*

	Coefficient	Std. Errors	z-stat	p-val
B	4.46	0.04	115.80	0.00
ψ_0	-36.82	30.60	-1.20	0.11
$\sigma^2_{v_p}$	0.02	0.01	2.42	0.01
$\sigma^2_{w_{\psi}}$	18.11	3.04	5.95	0.00

Appendix H: Alternative measures of the increase in volatility

In this appendix, I present alternative measures of the increase in the volatility of the Price-Dividend ratio. In the main text, I focus on the unconditional variance. In this appendix, I show that the the conditional variance also went up. Since taxes are a relatively slow-moving variable, I concentrate on the long run component of the conditional variance. First, I present long-run volatility measures based on a GARCH model variant. Second, I include taxes in the long run component. Finally, I show volatility measures for daily price growth and stock returns.

H1.- The AR-GARCH-MIDAS model

The long run price-dividend volatility is the permanent component of the conditional variance. The procedure to obtained it builds upon the GARCH-MIDAS model outlined by Engle et al. (2013). They decompose the conditional variance between a permanent and a transitory component, where the former is a filter over a number of lags of the realized volatility. However, they assume a constant conditional mean that is suitable for stock returns (with a mean close to zero and i.i.d. deviations from the mean) but not for the PD ratio (which is highly persistent, with time varying mean and, as a result, serially correlated deviations from the constant long run mean). To adapt the model for the PD ratio, I introduce an AR(1) model for the conditional mean, in which case the AR(1) residuals become serially uncorrelated and then the GARCH-MIDAS procedure can be applied for the variance.

The model boils down to the following list of equations. Unexpected changes in the price-dividend ratio in quarter q of year t are uncorrelated and normally distributed

$$\frac{P_{q,t}}{D_{q,t}} - \mathbb{E}_{t-1}\Big(\frac{P_{q,t}}{D_{q,t}}\Big) = \epsilon_{q,t}, \quad \epsilon_{q,t} \sim ii\mathcal{N}(0, \sigma^2_{q,t}) \tag{142}$$

with the conditional expectation given by an AR(1) process

$$\mathbb{E}_{t-1}\Big(\frac{P_{q,t}}{D_{q,t}}\Big) = \mu + \rho\frac{P_{q-1,t}}{D_{q-1,t}}$$

The conditional variance model hypothesizes that there is a short run (or transitory) $\mathrm{g}^2_{q,t}$ and a long run (or permanent) variance v^2_t. The permanent component captures an underlying state which makes equivalent surprises in the PD ratio have different effects. For instance, better than expected dividends might have a different impact in stock prices in high or low capital tax environments. As stated by Engle and Rangel (2008), the long-memory component can be interpreted as a trend around which the conditional variance fluctuates. All in all, the errors standard deviation is the product of the short and long run components

$$\sigma_{q,t} = \mathrm{v}_t g_{q,t} \tag{143}$$

It is assumed that the transitory component follows a GARCH(1,1)

$$\mathrm{g}_{q,t}^2 = 1 - \alpha_0 - \alpha_1 + \alpha_0 \frac{\epsilon_{q-1,t}^2}{\mathrm{v}_t^2} + \alpha_1 \mathrm{g}_{q-1,t}^2 \tag{144}$$

In turn, the long run volatility is a MIDAS filter over K past realized volatility

$$\mathrm{v}_t^2 = \phi_0 + \phi_1 \sum_{k=1}^{K} \varphi_k(w) RV_{t-k}^Q \tag{145}$$

with the realized volatility defined as a moving variance over a fixed window of Q quarters

$$RV_t^Q = \frac{1}{Q-1} \sum_{q=1}^{Q} \Big(\frac{P_{q,t}}{D_{q,t}} - \frac{1}{Q} \sum_{q=1}^{Q} \frac{P_{q,t}}{D_{q,t}} \Big)^2 \tag{146}$$

and the weighting scheme given by a beta lag structure, which yields a monotonically decreasing sequence determined by a single parameter

$$\varphi_k(w) = \frac{(1 - k/K)^{w-1}}{\sum_{j=1}^{K} (1 - j/K)^{w-1}} \tag{147}$$

Altogether, the parameter vector θ contains a total of 7 parameters $\theta = \{\mu, \rho, \alpha_0, \alpha_1, \phi_0, \phi_1, w\}$, jointly estimated through Quasi-Maximum Likelihood[65].

For the baseline application, I use an intra-annual standard deviation as the measure of realized volatility (i.e., Q=4) and 10 lags of this realized volatility to compute the long run component v (i.e., K=10). The financial literature usually works with high frequency data (often daily variables) and regard the long run component as the underlying monthly or quarterly trend (e.g., Schwert (1989), Engle et al. (2013)). Differently, here I have adopted a low frequency approach closer to macroeconomics, where the long run frequency tend to go beyond the business cycle. As a result, I have regarded the long run as a 10-year trend. The reason is that taxes are a much slower evolving variable than prices and dividends. In other words, potentially taxes reflect a long-lasting structure or regime, which determines the business conditions for long periods.

Table 15 shows the estimation results. All the coefficients are significant at usual confidence

[65]It is well known that the QML estimator is consistent and asymptotically normal for GARCH(1,1), provided that the innovation distribution has a finite fourth moment, even if the true distribution is far from Gaussian (e.g., see Lumsdaine (1996)). This is the case here indeed: residuals are non-Gaussian (due to fat tails) but exhibit an empirical kurtosis of 5.98 so that quasi-maximum likelihood estimators are asymptotically Normally distributed.

levels. As observed in figure 13, this estimation yields a long run volatility with a steep increase since the 1990s. The variable peaks in the aftermath of the Great Recession but the post-GR volatility is still way above its historical mean. The persistence of higher volatility turns out to be robust across alternative specifications: the number of the realized volatility lags considered for the long run filter; the fixed vs. rolling window specification for realized volatility.

Table 14: ***AR-GARCH-MIDAS model estimation results.*** *The table shows the QML estimation of all the parameters of the AR-GARCH-MIDAS model for the fixed window realized volatility with Q=4 and K=10. The data used for the estimation covers the 1940:I-2018:IV period.*

	Coefficient	Std. Errors	t-stat	p-val
μ	2.91	0.58	5.04	0.00
ρ	0.96	0.01	139.88	0.00
α_0	0.26	0.06	4.11	0.00
α_1	0.67	0.067	10.08	0.00
ϕ_0	10.17	2.60	3.90	0.00
ϕ_1	0.66	0.33	1.99	0.05
w	1.03	0.18	5.63	0.00

The left hand side plot of figure 13 plots the estimation with the fixed window for the realized volatility and different number of lags. As expected, the higher the lags considered, the smoother is the trajectory, without modifying the baseline result. The case of K=5 results in a volatile series, resembling the one for the conditional variance. That illustrates two things: probably 5 year are not enough to capture of long-lasting permanent component; the persistence of volatility is mostly due to the echoes of the Great Recession (i.e., the short run volatility after the GR has not been particularly high).

The right hand side graph of figure 13 compare the fixed to the rolling specification for the realized volatility. The former transforms quarter data into annual long run volatility; the latter keeps the long run volatility at quarter frequency. Both uses 10 years of data to produce the long run measure. As observed, the gray and the red line display the same qualitative trajectory. However, the rolling measure fluctuates way more, reaching a higher peak and reversion.

Finally, to check for the potential model-dependency of the result, I have computed alternative measures. On the one hand, the black line (R2) plots a simple 10 year rolling window standard deviation of the price-dividends deviations from their conditional mean (i.e., the residuals from equation 142). This measure is the one that resembles the most to the baseline one. On the other hand, the dotted black line (R3) shows a 10 year rolling window standard deviation of the Hodrick-Prescott cyclical component of the price-dividend quarterly data. This one follows the rolling window AR-GARCH-MIDAS closely. The correlation matrix among them gives a quantitative view of this comparison (the baseline measure displays correlations above 0.9 with all the alternative measures, except for the medium-term measure).

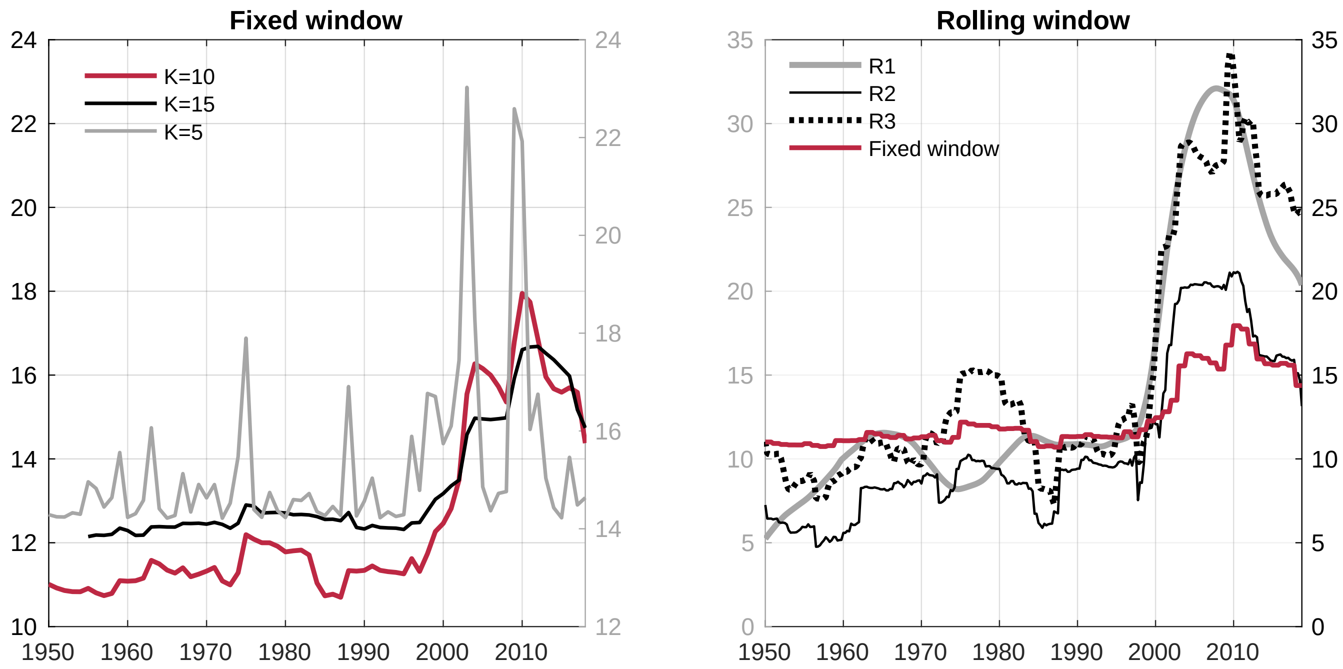

Figure 13: ***Time-varying volatility measures.*** *The left hand side graph plot the baseline model with a Q-fixed window realized volatility (for Q=4) and changes the number of annual lags used for computing the long volatility (equation 145). The red line is the baseline measure (K=10). The right hand side graph plots the baseline fixed window measure (K=10) against a number of rolling window based measures: R1 is the AR-GARCH-MIDAS estimate for the case of a 10-year rolling window for the realized volatility; R2 is a 10-year rolling window standard deviation over the AR(1) residuals of equation 142; R3 is a 10-year rolling window standard deviation over the HP cyclical component of the quarterly price-dividend ratio. The data used for the estimation covers the 1940:I-2018:IV period.*

H2.- Taxes as the long run component

One of the contributions of the Engle et al. (2013)'s GARCH-MIDAS model is allowing for the introduction of macroeconomic variables directly in the long term component. That would avoid a two-step approach (as Schwert (1989)) consisting of measuring volatility and estimate VAR models with volatility proxies and macrovariables. The problem with the 2-step approach is the measure-

*Table 15: **Volatility measures correlation matrix.** The table shows the contemporaneous correlation among all the long run volatility measures. The labels are as in figure 13. All the correlations are significant at 99% confidence level. The data used for the estimation covers the 1940:I-2018:IV period.*

	K=10	K=15	K=5	R1	R2	R3
K=10	1.00					
K=15	0.97	1.00				
K=5	0.50	0.42	1.00			
R1	0.94	0.89	0.50	1.00		
R2	0.94	0.88	0.53	0.96	1.00	
R3	0.97	0.93	0.52	0.94	0.96	1.00

ment error in RV that would bias the coefficient capturing the effect of past volatility on current volatility and on the macroeconomic variables (see Engle et al. (2013)). On the contrary, a single-step procedure would circumvent this problem by adding the macrovariable directly to the variance model. That is one of the contributions of the GARCH-MIDAS approach.

In this paper, though, the impact of taxes on volatility is analyzed through a VAR, that is, via a 2-step approach. The potential cost is the bias effect coming from measurement errors, indeed. However, by assuming that potential cost, I can compare the time series, which opens the door to a much richer analysis than a coefficient significant test (for instance, I can explore the dynamic effects of tax cuts via IRFs). Since that seems to be a substantial gain, the potential measurement error problem is addressed in the VAR context itself (following Forni et al. (2020)). On top of that, the VAR is only used for descriptive purposes and the causal analysis is left for the structural DSGE model.

Nonetheless, for the sake of completeness, in this appendix I report the results of the Engle et al. (2013)'s single-step approach. That amounts to modify the long run component as follows:

$$\mathrm{v}_t^2 = \phi_0 + \phi_1 \sum_{k=1}^{K} \varphi_k(w)\tau_{t-k} \tag{148}$$

where τ is the capital income tax defined in the paper. Table 16 reports the estimation results. Notice that all three coefficients get a decent t-stat value for different number of lags (for K=5, all of them are significant at standard confidence levels).

Table 16: ***AR-GARCH-MIDAS model with taxes as the long run component.*** *The table shows the QML estimation of all the parameters of the model when replacing equation 145 by 148, for the fixed window realized volatility with Q=4. t-statistics in parenthesis. The data used for the estimation covers the 1940:I-2018:IV period.*

	K=5	K=10	K=15
μ	2.19	3.64	4.62
	(3.51)	(6.91)	(9.94)
ρ	0.99	0.98	0.97
	(136.20)	(189.46)	(238.36)
α_0	0.25	0.28	0.28
	(2.98)	(3.12)	(3.46)
α_1	0.63	0.63	0.63
	(7.33)	(7.46)	(7.49)
ϕ_0	21.29	20.87	22.36
	(2.16)	(1.51)	(1.66)
ϕ_1	-3.78	-3.51	-3.93
	(-1.83)	(-1.17)	(-1.31)
w	1.16	3.62	1.02
	(3.01)	(0.94)	(3.81)

H3.- Absolute and Relative Volatility

The baseline volatility measure uses the absolute deviations from the price-dividend conditional mean as the main ingredient (see equation 142). In this sense, it can be regarded as a measure of absolute volatility. That contrasts with **?**'s view, that regard volatility as a relative measure (dealing with price percentage changes or rate of returns). In this section, I comment on the relevance of using price-dividends absolute deviations as well as on the robustness of the rise in volatility when using high-frequency relative measures.

One of the concerns of using absolute measures is that they may "exaggerate" the degree of volatility when the variable level is in high heights. That concern has been repeatedly expressed by Schwert (Schwert (1989), **?**). In this regard, it is convenient to remember that larger deviations from the mean are not a mechanical consequence of a higher mean at all; in other words, they carry some useful information. It suffices to point out the strong positive correlation between absolute price-dividend fluctuations and investment cycles (Cochrane (2017), Adam and Merkel (2019)). The fact that high-mean times tend to go in hand with high-variance times is not a trivial coincidence but something to be explained. The hypothesis of the paper is that the decline in capital taxes is an important driver behind not only the rising trend in stock market valuations (in line with McGrattan and Prescott (2005) and others) but also the larger swings around the trend.

For the particular case of price-dividend ratios, it is worthy it to remember that the variable is a ratio such that pure time-trend effects are accounted for[66]. Besides, the variance of the price-dividend is an important object in the price excess volatility literature, starting with Shiller (1981)[67].

Finally, the rise in long run volatility is also observed when using high frequency price percentage changes or rate of returns. Figure 14 plots a non-parametric measure of it, a 10-year rolling standard deviation over the daily series[68]. In both cases, the long run measure reveals a gradual increase in volatility since the early 1970s.

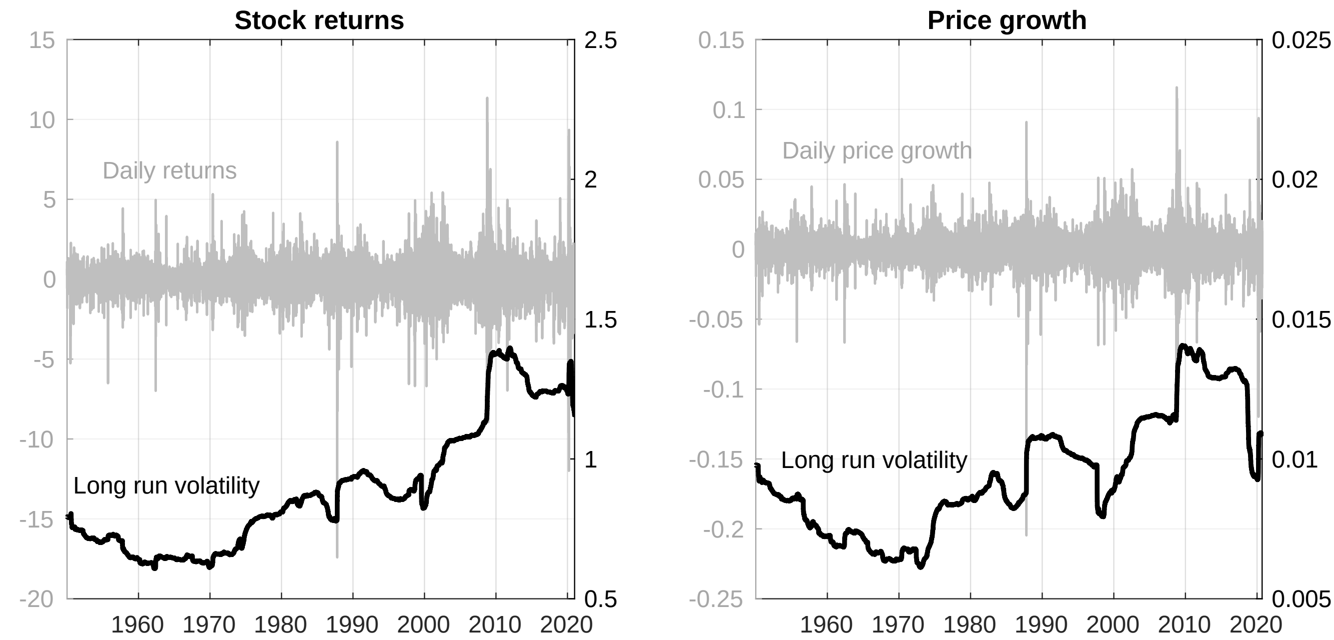

Figure 14: ***Long run volatility of relative variables.*** *The left hand side graph pictures daily stock returns from French's website;the right hand side graph pictures daily price growth from Yahoo's historical data. Both graphs includes the long run volatility, as the 10-year rolling window standard deviation of the series. Data is from 1950 to 2020.*

The rising volatility phenomenon is masked when using relative variables (as returns, price growth and even price-dividend growth) at lower frequency. That signals the existence of some

[66]The results hold when using a detrended price-dividend ratio, no matter the detrending method.

[67]Some of the papers uses the log(PD) and then, focus on percentage changes. However, that is more a requirement of the log-linearization approach to derive the price-dividend variance decomposition than a claim in favour of the relevance of the percentage changes.

[68]Results are robust to detrend the variables.

highly volatile but short-lived events, which when aggregated over a month or quarter they partly offset each other. All in all, the rise in volatility showed consistently by the different measures is just a statistical way of capturing the increase in the frequency and magnitude of well known stock market booms since the 1980s (the late 80s crashes, the Dot-com bubble and the Great Recession and its aftermath).