# Capital Gains Taxation

# and

# Asset Price Cycles


PAU BELDA*

Universitat Autònoma de Barcelona
and Barcelona School of Economics


LATEST VERSION HERE


**Abstract**

Can capital gains taxes prevent asset price cycles? Using an asset pricing model with learning I show that the capital gains tax dampens a key volatility engine -the price-expectations feedback loop- by reducing the sensitivity of prices to expectations fluctuations and by anchoring expectations around their fundamental value. With this theory in hand, I show that the decline in taxes observed the last decades in the United States can explain a substantial part of the increase in the Price-Dividend ratio fluctuations along with the surge in stock market valuations and the equity premium. In fact, the model suggests that the rise in the Price-Dividend ratio volatility would have been entirely avoided in the absence of tax cuts despite the fall in safe real interest rates. Furthermore, I characterize the optimal policy in an economy where asset price volatility leads to business cycles. The optimal tax counteracts too optimistic/pessimistic beliefs, preventing beliefs-driven macroeconomic fluctuations. In some cases, it equals 100%.



*Keywords:* Capital Taxation, Asset Pricing, Expectations, Equity Premium, Macro-Financial Stability.

*JEL codes:* D83, D84, E32, E44, E62, G12, G14.



*PhD candidate. I thank Albert Marcet for his guidance. Besides, I have benefited from comments from Klaus Adam, Ricardo Reis, Dirk Krueger, Jordi Caballé, Luís Rojas, Hugo Rodríguez, Alexander Ludwig, Abhay Abhyankar, Ramon Ruiz and several participants in several seminars. 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.

Some of the most extraordinary asset price booms took place in low capital gains tax regimes. For instance, the rise and fall of Wall Street in the Roaring Twenties was preceded by a cut from a top rate of 73% to 12.5% in 1921; the Japanese stock bubble of the 1980s took off in a tax-free environment; the Dotcom episode was predated by the 1997 capital gains tax cut; and tax breaks preceded the 2000s housing boom in many jurisdictions.[1] Motivated by this anecdotal evidence, this paper studies how capital gains taxes influence asset price cycles.

The question is theoretically explored using a consumption-based asset pricing model with imperfect capital markets characterized by capital taxes, trading constraints and investors with limited market information. With the theory in hand, I approach two issues. First, a quantitative one: How much of the rise in the Price-Dividend (PD) ratio fluctuations observed the last decades in the United States can be explained by the decline in taxes? Second, a theoretical one: What is the optimal tax to prevent beliefs-driven macroeconomic cycles?

**A theory of how the tax level regulates asset price fluctuations.** The basic asset pricing equation points out that an effective tax on future capital gains would influence their present value and then, the current price level. In other words, taxes lower marginal investor's payoffs such that she is willing to pay a lower price for the asset. This is the so-called *tax capitalization hypothesis*.[2] An implication of it is that tax changes would induce price changes and then, tax ups and downs might be behind asset price booms and busts.

Nonetheless, this paper poses that a crucial part of the influence of capital gains taxes on asset price cycles is due not to tax *changes* per se but to the tax *level*. The primary reason is that the tax level affects the stochastic discount factor. In particular, higher taxes would impose a higher discount on future expected payoffs, diminishing their influence on current prices; or to put it differently, higher taxes would reduce the sensitivity of prices to beliefs. Other things equal, a lower sensitivity would scale down price fluctuations.

Additionally, more stable prices would induce more stable beliefs if agents use past prices to forecast future ones in any way. In other words, a second round effect might emerge when in-

[1]For instance, in Spain, the 1998 tax reform reduced a range of tax obligations related to home ownership, including additional capital gains exemptions, see García et al. (2005); similar tax incentives were put in place in Ireland or Denmark, where a tax freeze was introduced in 2001, see Klein et al. (2016); in the US, capital gains taxes were cut further in 2001 and 2003.

[2]Some references are Brennan (1970), McGrattan and Prescott (2005) or Sialm (2009).

vestors are extrapolative: by affecting prices, taxes influence the expectations formation process. Altogether, the tax level emerges as an important driver through their influence on discounting and learning. For example, while a one-time tax cut would increase asset prices once, its effect on price volatility would be persistent, as changes in beliefs would have more influence on prices, then becoming more volatile. Hence, a single tax cut can have long lasting effects on stock market volatility.[3]

These effects of taxes are probably better understood in the context of a theory of asset price cycles. The paper takes an approach with a long tradition in macrofinance: valuation cycles as self-fulfilling prophecies.[4] In a nutshell, investors act on their beliefs and their actions shape reality in line with what they expected, reinforcing the former. In a more formal way, a key engine behind price cycles is a feedback loop between expectations and prices: higher expected returns would drive demand and prices up; higher prices would feedback into higher expected returns. Precisely, taxes would dampen this feedback loop: high taxes would reduce the pass-through from expectations to demand and prices; more stable prices would feedback into more stable beliefs. Thus, it could be said that taxes throw sand in the gears of an important volatility engine. As a consequence, the transmission of shocks through expectations is dampened in higher tax environments, delivering less fragile capital markets.

This view of capital gains taxes as automatic stabilizers of financial markets is at odds with some conventional wisdom that suggests such taxes are in fact destabilizing (e.g. Stiglitz (1983)). The reason would be the so-called *lock-in effect*: taxes would discourage the realization of capital gains during a boom, restricting the aggregate stock supply and raising prices further. Nonetheless, this view focuses on the *realization* clause, ignoring all other channels. Thus, a tax on unrealized capital gains would be unambiguously stabilizing, as anticipated by Haugen and Heins (1969). Even more, I show that even a tax on realized gains would stabilize prices as long as the tax elasticity of realization is not too high (lower than 1 in absolute value), as it is the case according the most

[3]To the best of my knowledge, only Haugen and Heins (1969) and Haugen and Wichern (1973) have highlighted the stabilizing role of capital gains taxes via discounting. Thus, the hypothesis of this paper could be seen as a modern formulation of their work. "Modern" is used to refer to a general equilibrium approach that endogenize expectations. Indeed, expectations modelling turns out to be crucial to understand the stabilizing role of taxes. In standard asset pricing models with Rational Expectations taxes play no role, invalidating the results in the aforementioned papers. When agents learn, taxes turn out to be powerful since they dampen a the price-expectations loop, a key volatility engine. See Section 2 for a formal treatment of these issues.

[4]Back, at least, to Keynes (1936), and further developed by Minsky (1976), Kindleberger (1978) or Shiller (2000).

recent empirical estimations.[5] Hence, the results in this paper suggest an alternative assessment on the stabilization properties of capital gains taxes.

**An application to the US stock market.** I use the theory to connect three changes that took place since the 1980s in the US stock market: the decline in capital taxes, the rise in the aggregate stock market valuation and the larger valuation fluctuations. Thus, the connection between declining taxes and rising valuations suggested by McGrattan and Prescott (2005) among others is extended to valuation fluctuations.

To better understand the drivers of the increase in volatility directly from the data, I first amend the Campbell and Shiller (1988)'s PD ratio variance decomposition by including capital taxes. This extended version shows that about 40% of the PD fluctuations typically attributed to movements in returns seems rather related to tax changes. Besides, the decomposition sets out that behind the rise in the PD ratio variance there are three factors: tax changes, accounting for about 30% of the increase; a divorce between returns and dividend growth, responsible for about 60%; and a higher discount factor in part resulting from lower tax levels that explains about 10%.[6]

Then, a quantitative version of the model is set up with the aim of replicating the documented patterns. 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. The model is fed with the observed history of capital taxes and the registered reduction in income growth volatility, with the remainder parameters estimated using the Simulated Method of Moments. Model-implied statistics move in the right direction, reproducing about half of the increase in the PD level and at least half of the increase in volatility despite the reduction in macroeconomic volatility.

Furthermore, the literature has suggested an alternative explanation for the rise in PD volatility:

[5] Agersnap and Zidar (2021) found an elasticity between -0.5 and -0.3 using state level data from 1980 to 2016. In turn, the Joint Committee on Taxation (JCT) uses an elasticity of -0.68, similar to the -0.72 used by the Treasury. See Gravelle (2020) and the references there.

[6] The fact that returns and dividends comove in a negative rather than in positive way since the 1980s increase the PD volatility. The reason is that while the two of them exhibit comparable volatility throughout the period, before (after) the 1980s their positive (negative) covariance cancelled (adds) an amount of (to) the sum of their variances. Another implication is that, since the 1980s, a high PD ratio seems to forecast higher rather than lower future dividends growth which is in line with asset pricing theory but represents a novelty with respect to the typical puzzling negative relationship found in the literature (see, for instance, Cochrane (2009)).

the decline in safe real interest rates. Although taxes and interest rates are qualitatively equivalent in terms of their influence in the discount factor, a historical calibration of the model reveals their contribution is very different. Thus, the absence of tax cuts would have completely prevented both the rise in stock market valuations and their greater fluctuations despite the fall in real interest rates. On the contrary, the permanence of high interest rates would have avoided only a small part of the increase in volatility.

A remarkable result of the model is its ability to match the equity premium along with a low and stable risk-free rate, realistic consumption and dividend growth processes, a positive discount factor and low risk-aversion. In other words, it can account for the (historical) equity premium puzzle as stated by Cochrane (2017).[7] The reason is twofold. On the one hand, the model generates high volatility coming from beliefs, which makes realistic income processes and low risk aversion compatible with high enough returns. Besides, the inclusion of taxes induce additional volatility plus impart a trend on the PD ratio that helps in getting high returns without exaggerating its volatility, improving the outcomes of otherwise similar learning models. Crucially, these two factors do not affect the risk-free rate.[8]

Finally, I present additional empirical tests of the model's novel mechanism. First, a version of the model without this mechanism -which amounts to a version with Rational Expectations- is estimated. Without this channel, the model completely fails to generate the change in volatility. Besides, I test the mechanism using survey expectations. It turns out the PD ratio became more sensitive to survey beliefs over time, as predicted by the theory. Finally, the model suggested that stock prices would be more responsive to different shocks in low tax environments. The data seems to validate that: the response of the PD ratio to an equivalent price growth shock is twice as big in the low tax than in the high tax regime.

**Optimal capital gains taxation to stabilize capital markets and the macroeconomy.** The last part of the paper digs into the normative side of capital gains taxation. While the extra volatility coming from learning investors was a core element in both the theory and the explanation of the observed data patterns, it is also, however, a manifestation of inefficiency in financial

[7] There is a body of literature concerned with different methods of computing the equity premium. The discussion is about how to compute *expected* rates of return. Along this paper, I take a backward-looking approach, computing the expected rate of return as the historical average return, rather than infer it from discounting future cash flows.

[8] Interest taxes also went down over the period. However, there is no feedback loop affecting bond prices since they are one-period and then, the effect of tax cuts is way less important.

markets. In fact, it could be regarded as a pecuniary externality, since it emerges from the inability of agents to internalize the equilibrium price formation due to imperfect information about other market participants.[9] Thus, rational individuals cope with their imperfect knowledge by using reasonable forecasting models without general equilibrium considerations and learning from new data, which ends up causing aggregate instability in capital markets.

Nonetheless, such externality would be harmless in endowment models as the ones used in the previous analysis. Indeed, a precondition to explore an optimal tax is to establish a connection between asset prices and consumption fluctuations. For that purpose, I setup a tractable model consisting of a two sector growth model with investment adjustment costs that features the Q-investment theory in a context of imperfect information and learning. In this framework, learning markets provoke asset price fluctuations way larger than efficient markets and, since firms determine investment guided by the capital price, these excessive price movements lead to wide cycles of over- and under-accumulation of capital that ultimately affect social welfare.

In such a learning world, a paternalistic social planner is asked to intervene to deliver the best possible competitive equilibrium, having at her disposal a tax on unrealized capital gains and lump-sum taxes. I show these instruments can undo the effects arising from information frictions: the planner uses capital gains taxes to close the gap between learning and efficient prices and rebates the proceeds in a lump-sum manner to avoid income effects. These operations happen to be sufficient to restore First Best allocations or, to put it differently, taxes make learning markets deliver an efficient allocation of resources despite the existence of information imperfections.

The optimal tax turns out to be a function of the subjective expectations deviations from its objective counterpart (call it $\beta^*$), correcting too optimistic/pessimistic hikes. A shortcoming is that the optimal tax is unbounded and inherits the dynamic properties of subjective beliefs, being persistent but potentially quite volatile. Since tax volatility might not be very convenient, an alternative implementation is suggested: instead of solving an explicit optimization problem, asset price volatility is decomposed into a fundamental and a non-fundamental component and the tax is set as to eradicate the latter.[10] It turns out that a constant tax equal to 1 can achieve that. Nonetheless,

[9] When investors homogeneity is not common knowledge, agents cannot neither use market clearing conditions ex-ante nor apply the Law of Iterated Expectations. Therefore, the standard procedure to derive equilibrium prices based on forward iteration on the Euler Equation imposing market clearing is not possible. See Section 2 and Adam and Marcet (2011).

[10] This decomposition is in the spirit of the decomposition between fundamental and non-fundamental trading

such a high tax depresses the capital price a bit, inducing chronic under-investment. This effect can be avoided by resorting to a subsidy to capital rents, which turns out to be a function of just $\beta^*$.[11] Altogether, a combination of a constant tax on capital gains and a relatively stable subsidy on capital rents can restore efficiency in an otherwise too unstable capital market.

All in all, the arguments developed in the paper suggest that a tax on unrealized capital gains could be an effective tool to prevent asset price booms and busts and the associated financial and macroeconomic cycles that often comes with them. In this macroprudential sense, it can thought of as an alternative to the perhaps more popular but less clearly effective Financial Transactions Tax.[12]

The rest of the paper proceeds as follows. Section 1 highlights the contributions to the literature. Section 2 exposes the theory. Section 3 presents a quantitative application of the theory to the US stock market. Section 4 derives optimal policy in a growth model. Section 5 concludes.

## 1.- Related literature

In this section, the main related literature and the innovations with respect to it are detailed. In particular, the paper is related to four bodies of literature: the influence of capital gains taxes on asset price volatility; the effect of capital taxes on asset price levels; stock market excess volatility and its connection with business cycles; and optimal capital taxation and macroprudential policy.

**Capital gains taxes and asset price volatility.** The might be called the dominant view regards the tax as destabilizing due to the lock-in effect: the increase in accumulated unrealized capital gains during a boom (burst) would translate into lower (higher) stock supply in order to avoid (enjoy) the tax burden (subsidy) of realization and then, generate even higher (lower) prices. This view was developed in Somers (1948), Gemmill (1956) or Somers (1960) and was further developed by Stiglitz (1983). There are, however, alternative views that point out the stabilizing effects of capital gains taxes. First, Haugen and Heins (1969) and Haugen and Wichern (1973)

suggested by Dávila (2020) to implement the optimal transaction tax.

[11]Not of subjective beliefs. Since objective beliefs tend to be way more stable than subjective ones, the subsidy would be relatively stable.

[12]The evidence for the ability of Financial Transaction Taxes to curb excess price volatility is at best unclear (e.g., theoretically, Buss et al. (2016), Buss and Dumas (2019), Dávila (2020); empirically, Umlauf (1993), Adam et al. (2015), Cappelletti et al. (2017)).

stress the influence of taxes on the sensitivity of prices to expectations via discounting. Second, Gemmill (1956) argues that the tax works in a countercyclical manner taking (giving) resources in good (bad) times and then stabilizing asset demands and prices. Finally, Dai et al. (2013) found an statistical negative relationship between taxes and returns volatility around the tax reforms of 1978 and 1997 and interpret the result as due to an increase in the risk of the asset due to lower risk-sharing with the government[13]. My paper could be seen as within this view, in particular related to the stabilization property of taxes via their effects on discounting[14].

**Capital taxes and asset prices.** There is a debate between *tax irrelevance* (starting with Miller and Scholes (1978)) and *tax capitalization* (starting with Brennan (1970)) (see Sialm (2009) for a complete review). Although there is no empirical consensus when looking at the returns cross-section[15], the tax capitalization has generated considerable agreement for aggregate market time series. Thus, McGrattan and Prescott (2005) using a quantitative neoclassical growth model and Sialm (2009) via reduced-form evidence showed that the fall in the effective capital tax rate was behind the rising trend in valuations. Recently, this hypothesis has been also confirmed by Brun and González (2017). Tax capitalization comes up very naturally in the paper which contrasts with Sialm (2006)'s results[16]. Besides, I suggest that the decline in capital taxes would be responsible not only for the rise in stock market valuation but also for its larger fluctuations.

[13]This idea of taxes as a risk-sharing device goes back at least to Lerner (1943) and it is formally explored in a body of literature that studies the effect of capital gains taxation on risk-taking and the level of returns that includes Domar and Musgrave (1944), Stiglitz (1975) or, more recently, Sikes and Verrecchia (2012).

[14]Besides, there are some other papers linking capital taxes with volatility. Baker et al. (2019) created a newspaper-based Equity Market Volatility tracker in which tax news drive part of price movements. Ferris (2018) explicitly addressed the relationship between dividend taxes and price volatility using an agency model in which dividend taxes affect the principal-agent costs associated with hiring a manager. Gomme et al. (2011) suggest that stochastic taxes are enough to make an RBC model generate enough returns volatility. However, they focus on a different return, dividends over the capital stock, whose volatility is 22 times lower than equity returns volatility.

[15]For instance, Fama and French (1998) did not find reliable evidence of tax effects using a sample of firms from 1965 to 1992, whereas Sialm (2009) did so when considering a larger period (1913-2006).

[16]He presented a version of the Lucas (1978) model with stochastic taxes. He pointed out that taxes affect asset prices obtaining, however, some undesirable results. On the one hand, the tax level is irrelevant for the price-dividend level under CRRA utility. On the other hand, he gets a positive relation between valuations and taxes under plausible assumptions, which is at odds with the time series evidence aforementioned. These results are mostly driven by the use of an unrealistic tax system that levies on the purchase of new stocks. Sialm (2006) claims that he uses a flat consumption tax. The agent's budget constraint (for the case of a single asset) reads as $c_t + (1-\tau_t)p_t s_t = (1-\tau_t)(p_t + d_t)s_{t-1} + (1-\omega)T_t$. As observed, the tax rate $\tau_t$ is taxing the acquisition of stocks ($(1-\tau_t)p_t s_t$), the dividend income ($(1-\tau_t)d_t s_{t-1}$) and the wealth of the agent ($(1-\tau_t)p_t s_{t-1}$). Nonetheless, buying stocks is not a taxable event in the US. There exists no federal wealth tax either. This is not the case in other countries. For instance, in the UK there exists a 0.5% duty on share purchases. Moreover, the tax on stock sellings (i.e., capital gains tax), which is quantitatively way more important than tax on purchases, is missing. Sialm recognized that the tax system he analyzed was not realistic. Thus, my work can also be seen as a more realistic introduction of stochastic capital taxes in a Lucas-like setup.

**Excess volatility.** Since the initial observation that prices fluctuated way more than the constantly discounted stream of future dividends (LeRoy and Porter (1981) and Shiller (1981)), the excess volatility literature have tried to understand the drivers of these exuberant fluctuations. Empirically, the literature has resorted to a variance decomposition of the PD ratio following Campbell and Shiller (1988). I extend this decomposition by including capital taxes and show that almost half of the PD variance typically attributed to returns seems rather related to tax changes. Theoretically, the paper is related to the Internal Rationality literature, following Adam and Marcet (2011). In particular, the quantitative model extends the analysis in Adam et al. (2017) by including stochastic taxes which improves its empirical performance in terms of the equity premium and the change in PD mean and volatility. Besides, there is a body of learning papers linking stock market and business cycles via labour demand (Adam and Merkel (2019)), collateral constraints (Winkler (2020)) and wealth effects (Ifrim (2021)). This paper contributes to that effort by introducing learning in a Q-investment theory setup.

**Optimal policy.** The paper is related to a recent literature on optimal policy in models with pecuniary externalities. Two prominent topics in the literature are credit cycles and capital flows in models with endogenous collateral constraints. Capital controls and taxes on borrowing have been proposed to correct these disfunctionalities (e.g., Lorenzoni (2008), Jeanne and Korinek (2010), Dávila and Korinek (2018), Jeanne and Korinek (2019)). This paper applies a similar macroprudential logic to beliefs-driven cycles. In fact, following Benigno et al. (2019)'s insight, optimal taxes are used as to restore full efficiency rather than constrained one. Although the purpose of the optimal tax is not financing public spending, the analysis of capital gains taxation in a two sector model represents an innovation with respect to the standard optimal capital taxation analysis that focuses on one sector models excluding then any consideration of capital gains taxes even when they use rich tax systems as in Chari et al. (2020). Finally, the optimal policy analysis speaks to the literature on asset price targetting and monetary policy, starting with Bernanke and Gertler (2000). In particular, this research offers evidence for a tool that can be effective at regulating asset prices preserving the use of interest rates to target consumption prices.

## 2.- Theory: taxes as stabilizers of valuation cycles

This section is a theoretical exploration of the role of capital gains taxes in asset price cycles in a minimalist asset pricing model with incomplete markets, imperfect information and learning. It stresses out the main theoretical proposition of the paper, that is, the variance of the PD ratio is decreasing on the tax 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 analyzed. 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 general equilibrium asset pricing model with a constant tax on realized capital gains and dividends is set up. Its basic layer is the Lucas (1978)'s tree model with i.i.d. endowments growth. On top of it, information frictions are allowed by modelling a general subjective probability measure. Finally, capital gains taxes are modelled as to incorporated its two opposite effects: capitalization and lock-in.

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

*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, obeying 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 $\varepsilon^d_t \sim \log \mathcal{N}(1, e^{s^2_d} - 1)$ an i.i.d. innovation. Capital gains are endogenously determined.

*Markets.* Financial markets are competitive but incomplete. 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 $\tau^D$ on dividends. Capital gains are taxed in a linear way at a rate $\tau^K$ on a realization basis. Each period a fraction $\pi \in [0, 1]$ of agents is forced to realize their one-period capital gains and pay the correspondent taxes while the remaining $1-\pi$ agents cannot realize their gains.[17] The fraction of agents realizing their gains is assumed to be decreasing on the tax level, that is, $\pi = \Pi(\tau^K)$ with $\Pi_{\tau^K} < 0$.[18] Symmetrically, capital losses give rise to transfers. All the revenues (outflows), call them $T_t$, are transferred to (taxed from) the individuals in a lump-sum way.

*Investors' information set.* Investors know the tax rates but ignore their marginal contribution to lump-sum transfers. Besides, they might face information frictions related to the dividend process or the knowledge of other investors' type. These possible frictions are 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 complete information. Thus, the underlying probability space is given by $(\Omega, \mathcal{B}, \mathcal{P}^i)$ with $\mathcal{B}$ denoting the corresponding $\sigma$-algebra of Borel subsets of $\Omega$ and $\mathcal{P}^i$ agent's i subjective probability measure over $(\Omega, \mathcal{B})$. For generality, we include prices in the the state space $\Omega$, with $\omega = \{D_t, P_t\}_{t=0}^{\infty}$ as a typical element.[19]

*Investors' program.* Each investor faces a consumption-savings problem: she chooses sequences of consumption and stock holdings $\{C_t^i, S_t^i\}_{t=0}^{\infty}$ by solving an optimization program using their subjective probability measure $\mathcal{P}^i$:[20]

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

subject to

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

$$0 \leq S_t^i \leq \bar{S}, \quad \text{given} \quad S_{-1}^i = 1 \tag{4}$$

The utility function is a time-separable continuous, increasing in consumption $U'(C_t^i) > 0$

[17]This is equivalent to assume that investors anticipate realizing a fixed proportion of the expected capital gains as in Sialm (2009)

[18]This is a shortcut that allows to incorporate the two effects that literature on capital gains taxation has pointed out (i.e., capitalization and lock-in) without having to deal with optimal trading strategies that would increase the complexity of the model remarkably.

[19]Note the exclusion of prices from the state space imposes a knowledge of the mapping from fundamentals to prices. See Adam and Marcet (2011) for a detailed discussion about it.

[20]This setup with optimizing investors that have a view on prices that possibly diverge from the objective one has been labelled as Internal Rationality. See Adam and Marcet (2011).

but concave $U''(C_t^i) \leq 0$ function. In this section, I assume risk-neutrality.[21] 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 Ponzi schemes and individual trading strategies directed to avoid taxes.

**Model Equilibrium.** The investor's program consists of maximizing a bounded continuous function over a compact non-empty feasible set.[22] 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 Karush-Kuhn-Tucker (KKT) theorem. Given the fiscal system, investor i's optimality conditions boil down to the following Euler Equation

$$
\begin{aligned}
P_t &= \delta \mathbb{E}_t^{\mathcal{P}^i}\Big[(1-\tau^D)D_{t+1} + (1-\pi)P_{t+1} + \pi[P_{t+1} - \tau^K(P_{t+1} - P_t)]\Big] \\
&= \delta \mathbb{E}_t^{\mathcal{P}^i}\Big[(1-\tau^D)D_{t+1} + P_{t+1} - \pi\tau^K(P_{t+1} - P_t)\Big]
\end{aligned}
\tag{5}
$$

along with a non-bubble condition, and the sequence of budget constraints and market clearing conditions. 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 agents cannot apply the Law of Iterated Expectations because the probability measure of future periods marginal agent is unknown.

### 2.2.- Capital gains taxes and the Price-Dividend ratio level

This section studies the effect of capital gains taxes on the asset price level. The literature has focused on two questions: whether taxes affect asset prices and if they do, whether the effect is positive or negative. Regarding the former, although several trading strategies that can circumvent taxes have been pointed out (e.g., Constantinides (1983), Stiglitz (1983)), it is well-known that capital market frictions as short-selling constraints make these strategies unfeasible. Hence, in the context of imperfect capital markets as in the model, capital gains taxes affect asset prices; the question is how. On the one hand, taxes would deprive investors from part of their reward making them less willing to pay for the asset. This is the capitalization effect, that imply a negative relationship between the tax and the asset price. Besides, taxes would make sellers ask for higher

[21] This assumptiont is just for the sake of simplifying the analytical derivation of the tax hypothesis of volatility. In the quantitative model of Section 3 , I introduce risk-averse investors.

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

prices because they have to pay taxes when selling. This is the lock-in effect, that induce a positive relationship between taxes and asset prices.[23]

The equilibrium PD ratio is given by

$$\frac{P_t}{D_t} = \frac{\delta(1-\tau^D)a}{1-\delta\beta_t^p(1-\pi\tau^K)-\pi\delta\tau^K} \tag{6}$$

with $\beta_t^p \equiv \mathbb{E}_t^{\mathcal{P}^i}\left[\frac{P_{t+1}}{P_t}\right]$. In line with the literature, $\tau^K$ becomes irrelevant if all investors decide to lock-in its capital gains (i.e., if $\pi = 0$). However, in general, $\tau^K$ would affect asset prices.[24] Its net effect is ambiguous as the model incorporates the two effects: through the budget constraint, higher taxes lower expected payoffs, demand and prices; through the effect of $\tau^K$ on the realization rate $\pi$, taxes push to lock-in gains and increase prices. Their balance depends on the tax elasticity of capital gains realization. *Proposition 1* shows that the capitalization effect prevails as long as the realization rate is not too sensitive to taxes.

***Proposition 1: Capital gains tax effect on the PD ratio level.*** *The relationship between the capital gains tax and the PD ratio level is negative (positive) if and only if the tax elasticity of realization is greater (smaller) than minus 1. In other words,*

$$\frac{P_t}{D_t}\left(\tau^K,\cdot\right) \lessgtr \frac{P_t}{D_t}\left(\tilde{\tau}^K,\cdot\right) \quad \text{for} \quad \tau^K > \tilde{\tau}^K$$

*if* $-1 \lessgtr \varepsilon_\tau^\pi$, *with* $\varepsilon_\tau^\pi \equiv \frac{\partial\Pi(\tau^K)}{\partial\tau^K}\frac{\tau^K}{\Pi(\tau^K)}$.

*Proof.* The capital gains tax affect the PD ratio through two channels. First, consider the Euler Equation, ignoring its effect on the realization ratio. For that, take the partial derivative of the equilibrium PD ratio with respect to $\tau^K$, keeping $\pi$ fixed at some value $\bar{\pi}$. Then,

$$\left.\frac{\partial P_t/D_t}{\partial\tau^K}\right|_{\Pi(\tau^K)=\bar{\pi}} = -\frac{\delta^2(1-\tau^D)a\pi(\beta_t-1)}{(1-\delta\beta_t^p(1-\pi\tau^K)-\pi\delta\tau^K)^2} < 0 \tag{7}$$

where the negative sign holds for positive expected capital gains (i.e., $\beta_t > 1$). Second, the

[23] See Dai et al. (2008) for a literature review about the two effects.

[24] Tax capitalization holds for dividend taxes, via a demand channel. Thus, the model capture the insights of McGrattan and Prescott (2005) in an endowment economy. The effects of $\tau^K$ and $\tau^D$ on prices contrasts with what Sialm (2006) obtained in a similar framework. He found that only tax changes would have an effect on asset prices. His result arises out of a very particular fiscal system, with a tax on stock purchases. Instead, a bit more realistic types as dividends and capital gains taxes can deliver the negative relationship between taxes and prices very naturally.

effect of $\tau^K$ through the realization rate is given by

$$\frac{\partial P_t/D_t}{\partial \pi}\frac{\partial \pi}{\partial \tau^K} = -\frac{\delta^2(1-\tau^D)a\pi(\beta_t - 1)}{(1-\delta\beta_t^p(1-\pi\tau^K)-\pi\delta\tau^K)^2}\frac{\partial \pi}{\partial \tau^K} > 0 \tag{8}$$

for $\beta_t > 1$ since it was assumed $\frac{\partial \pi}{\partial \tau^K} < 0$. Thus, higher taxes are associated to higher prices via a lower realization rate. The positive effect dominates if

$$\left.\frac{\partial P_t/D_t}{\partial \tau^K}\right|_{\Pi(\tau^K)=\bar{\pi}} + \frac{\partial P_t/D_t}{\partial \pi}\frac{\partial \pi}{\partial \tau^K} < 0 \tag{9}$$

which requires

$$\pi + \tau^K \frac{\partial \pi}{\partial \tau^K} > 0 \tag{10}$$

Rearranging this last condition and using the definition of the tax elasticity of realization $\varepsilon_\tau^\pi > -1$ is obtained. Likewise, the negative effect dominates if $\varepsilon_\tau^\pi < -1$ as stated in the proposition. ■

## 2.3.- Capital gains taxes and the Price-Dividend ratio volatility

This section scrutinizes the relationship between capital gains taxes and PD ratio fluctuations. I show that if the capitalization effect dominates, capital gains taxes are stabilizing since they reduce the sensitivity of prices to beliefs. Thus, higher taxes impose a higher discount on future expected payoffs, reducing their influence on current prices and then, reducing the pass-through from beliefs fluctuations to the PD ratio. On the other hand, taxes are destabilizing if the lock-in effect dominates since taxes would discourage the realization of capital gains during booms, restricting supply and raising prices further. Thus, the analysis encompasses the two views in the literature, the destabilizing (e.g., Somers (1948), Stiglitz (1983)) and the stabilizing one (Haugen and Heins (1969), Haugen and Wichern (1973)). *Proposition 2* formalizes the idea.

***Proposition 2a: Stabilization properties of capital gains taxes.*** *If there is a negative (positive) relationship between taxes and the PD ratio, the unconditional variance of the PD ratio* v *is decreasing (increasing) on the capital gains tax level. In other words, -1* $\lessgtr \varepsilon_\tau^\pi$ *implies that*

$$\mathrm{v}\left(\tau^K, \cdot\right) \lessgtr \mathrm{v}\left(\tilde{\tau}^K, \cdot\right) \quad \text{for} \quad \tau^K > \tilde{\tau}^K$$

*Proof.* Take a 1st order Taylor approximation on the equilibrium PD ratio (6) around $\beta_t^p = 1$.

Then apply the variance on both sides to obtain

$$\text{v} \equiv \mathbb{V}ar\left[\frac{P_t}{D_t}\right] \approx \underbrace{\omega(\tau^K)^2}_{\text{Tax wedge}} \times \mathbb{V}ar(\beta_t^p) \tag{11}$$

where $\omega(\tau^K) \equiv \frac{\partial P_t/D_t}{\partial \beta_t^p}$ evaluated at $\beta_t^p = 1$. It turns out

$$-1 \lesseqgtr \varepsilon_\tau^\pi \quad \Rightarrow \quad \frac{\partial \omega}{\partial \tau^K} = \frac{\partial^2 P_t/D_t}{\partial \beta_t^p \partial \tau^K}\Big|_{\beta_t^p=1} \lesseqgtr 0$$

and then higher taxes reduce (increase) the pass-through from beliefs to PD fluctuations since the quadratic function is monotonic for positive taxes, that is,

$$\frac{\partial \omega}{\partial \tau^K} < 0 \quad \Rightarrow \quad \frac{\partial \omega^2}{\partial \tau^K} < 0$$

Notice that the proposition holds for $\beta_t^p > 1$ and reverses its sign for expected capital losses. ■

*Proposition 2* kept $\beta_t^p$ as given. Thus, the proposition would hold for any expectations model. Now I explore two different expectations models that deliver additional insights. First, under Rational Expectations $\tau^K$ becomes neutral (i.e., the tax level does not affect the PD ratio volatility). Next, the case of learning about prices within an Internal Rational framework. In this second case, $\tau^K$ becomes particularly powerful because it intervenes over the feedback loop between beliefs and prices emerging from learning. The following Propositions elaborate on these points.

***Proposition 2b: Tax neutrality under complete information.*** *Assume i) agents know the dividend process as defined by expression (1) and ii) investors homogeneity is common knowledge in the sense of Aumann (1976). This is the case of Rational Expectations. Then, the capital gains tax level $\tau^K$ is neutral, that is,*

$$\text{v}\left(\tau^K, \cdot\right) = \text{v}\left(\tilde{\tau}^K, \cdot\right) \quad \text{for} \quad \tau^K > \tilde{\tau}^K$$

*Proof.* Using the Euler Equation, market clearing and a non-bubble condition, the representative investor can derive the following equilibrium prices[25]

$$\frac{P_t^{RE}}{D_t} = \frac{\delta(1-\tau^D)a}{1-\delta\pi(1-\pi\tau^K)a - \delta\pi\tau^K} \tag{12}$$

[25]Notice that this pricing formula is exactly the same as the one obtained from Gordon's model with constant

Since the PD ratio becomes a constant, its variance is zero and taxes cannot play any role. Is this neutrality trivial? ■

Now assume investors know the dividends process but homogeneity is not common knowledge. As a result, agents cannot deduce equilibrium prices from what they know since they cannot impose market clearing (i.e., an *aggregate* equilibrium condition) ex-ante. Thus, price formation (a general equilibrium process) is not internalized by the agents; instead, prices become part of the state space $\Omega$. Since agents are ignorant about the equilibrium pricing function, they need a model to forecast future prices. The proposed subjective model generalizes the complete information one, allowing price growth to differ from dividends growth:[26]

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

$$\ln p_t = \ln p_{t-1} + \ln \eta_t \tag{15}$$

with i.i.d. normally distributed disturbances with zero mean and constant variance. The permanent component of price growth $p_t$ is unobserved and has to be estimated from price signals. For that purpose, investors use a Kalman filter. The posterior (conditional on the observed price history) is given by $\ln p_t \sim \mathcal{N}(\ln \beta_t, \sigma_p^2)$ where $\sigma_p^2$ is the steady state Kalman filter uncertainty and the posterior mean evolves recursively following

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

where $g$ is the steady state Kalman gain. As it is standard in the literature, equation (16) contains lagged price growth and no variance correction terms.[27] Hence, subjective price expectations are given by $\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{P_t}\right] = \beta_t$. In this setup, the following Proposition holds.

***Proposition 2c: Reinforced tax non-neutrality with incomplete information and learn-***

expected returns, dividend growth and taxes. Thus, letting $\delta = 1/(1+r)$ and $a = 1+g$, equation (12) becomes

$$P_t^{\text{Gordon}} = \frac{(1-\tau^D)(1+g)}{r - g(1-\pi\tau^K)} D_t$$

[26]The implicit price model under RE mimics the dividends dynamics:

$$\ln P_t = \ln a + \ln P_{t-1} + \ln \varepsilon_t^d \tag{13}$$

[27]The reason is that it is assumed agents observe in period t information about the lagged transitory component $\ln \varepsilon_{t-1}^P$. 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.

***ing.*** *Assume that fundamentals processes are known, homogeneity is not common knowledge and agents learn about capital gains using equation (16). In this case, the effect of taxes on the PD ratio volatility is reinforced, that is,* $-1 < (>)\varepsilon^{\pi}_{\tau}$ *implies that*

$$\mathrm{v}\Big(\tau^K,\cdot\Big) << (>>)\mathrm{v}\Big(\tilde{\tau}^K,\cdot\Big) \quad \text{for} \quad \tau^K > \tilde{\tau}^K$$

*Proof.* The reason of the extra effect of taxes on volatility is that they also influence beliefs formation, that is, $\beta_t = \beta(\tau^K,\cdot)$. Without loss of generality, assume $-1 < \varepsilon^{\pi}_{\tau}$. Under learning, equilibrium prices reads as

$$\frac{P^L_t}{D_t} = \frac{\delta(1-\tau^D)a}{1-\delta(1-\pi\tau^K)\beta_t - \delta\pi\tau^K} \tag{17}$$

$\{\beta_t\}$ is fully determined by the following 2nd order difference equation, that arises from combining equilibrium prices (17) and the expectations updating equation (16):

$$\beta_t = (\beta_{t-1})^{1-g}\left(a\varepsilon^d_t + \frac{\delta(1-\pi\tau^K)\Delta\beta_{t-1}a\varepsilon^d_t}{1-\delta(1-\pi\tau^K)\beta_{t-1} - \delta\pi\tau^K}\right)^g \tag{18}$$

The basic properties of this equation can be illustrated in a two-dimensional phase diagram on the $(\beta_t, \beta_{t+1})$ plane (keeping dividend shocks at their mean values), as the one on the left of figure 1. Four phases come up: boom (A), bust (C) and two reversion areas (B and D). Thus, expectations display momentum and mean reversion (see Adam et al. (2016)). Importantly, $\tau^K$ plays a key role in $\beta_t$ dynamics. As shown on the right graph of figure 1, the higher the tax the smaller the boom area. In other words, a high tax regime would anchor expectations, making them to orbit closer to its fundamental value. This is a logic consequence of learning: since higher taxes makes price more stable, price expectations will become less volatile as long as agents learn from actual prices. All in all, when expectations are influenced by market outcomes, taxes not only affect the sensitivity of prices to beliefs but also beliefs dynamics. Hence, the tax non-neutrality gets reinforced:

$$\mathbb{V}ar\left[\frac{P^L_t}{D_t}\right] \approx \omega(\tau^K)^2 \times \mathbb{V}ar(\beta_t(\tau^K,\cdot)) \tag{19}$$

■

Finally, an unexplored question in this framework is whether the dividends tax exhibits similar stabilization properties. To answer this question, we repeat the previous analysis for $\tau^D$. *Proposition 3* argues that a dividends tax can indeed stabilize the PD ratio but in a weaker way than capital gains taxes. Do I need it here? Better in an appendix?

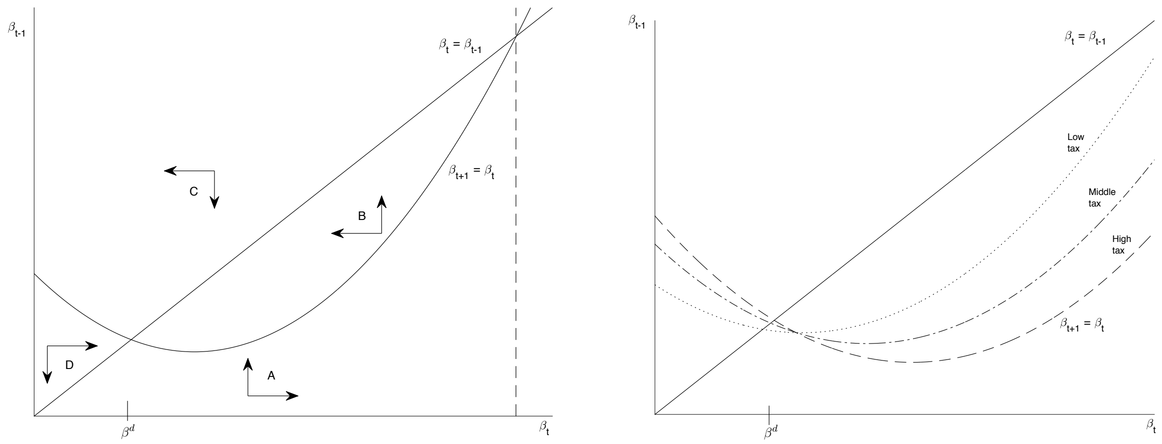


*Figure 1:* ***Subjective Expectations Phase Diagram.*** *The panel on the left illustrates the dynamics of subjective capital gains expectations given by the 2nd Order Difference Equation(18) for the case of* $-1 < \varepsilon^{\pi}_{\tau}$. *It considers* $\Delta\beta_{t+1} = \beta(\beta_t, \beta_{t-1})$ *with the other variables being fixed at mean values. The solid lines pictures the nullclines that makes* $\Delta\beta_t$ *and* $\Delta\beta_{t+1}$ *equal to zero, respectively. The panel on the right shows the* $\Delta\beta_{t+1} = 0$ *curve at different* $\tau^K$ *levels.*

***Proposition 3a: Stabilization properties of dividends taxes.*** *The unconditional variance of the PD ratio is decreasing on the dividend tax, that is,*

$$\mathrm{v}\Big(\tau^D, \cdot\Big) \lesseqgtr \mathrm{v}\Big(\tilde{\tau}^D, \cdot\Big) \quad \text{for} \quad \tau^D > \tilde{\tau}^D$$

*Proof.* Proceeding as in *Proposition 2a*, it turns out

$$\frac{\partial \omega}{\partial \tau^D} = \frac{\partial^2 P_t/D_t}{\partial \beta^p_t \partial \tau^D}\Big|_{\beta^p_t = 1} = -\frac{\delta^2 a(1 - \pi\tau^K)}{(1 - \delta\beta^p_t(1 - \pi\tau^K) - \pi\delta\tau^K)^2} < 0$$

Hence, $\tau^D$ reduces the pass-through from beliefs volatility to PD ratio changes:

$$\mathrm{v} \equiv \mathbb{V}ar\left[\frac{P_t}{D_t}\right] \approx \underbrace{\omega(\tau^D)^2}_{\text{Tax wedge}} \times \mathbb{V}ar(\beta^p_t) \tag{20}$$

■

***Proposition 3b: The stabilization power of dividends taxes is weaker than that of capital gains taxes.*** *The effect of dividends taxes on PD ratio volatility is smaller than that of capital gains taxes when the net effect of capital gains on the PD ratio is negative. In* *other words,* $-1 \lesseqgtr \varepsilon_\tau^\pi$ *implies that*

$$\left|\frac{\partial \text{v}}{\partial \tau^D}\right| \lesseqgtr \left|\frac{\partial \text{v}}{\partial \tau^K}\right|$$

*Proof.* I show that $\omega$ and $\mathbb{V}ar(\beta_t^p)$ react more to $\tau^K$ than to $\tau^D$ while keeping the right sign. I proceed in two steps. First, combining the derivatives with respect to $\omega$, it turns out

$$\frac{\partial \omega}{\partial \tau^K} = \mathcal{A}\frac{\partial \omega}{\partial \tau^D} - \mathcal{C}$$

with $\mathcal{A} = (1-\tau^D)2\delta(\beta_t^p-1)(\pi+\tau^K\frac{\partial\pi}{\partial\tau^K})$ and $\mathcal{C} = -\delta^2 a(1-\tau^D)(\pi+\tau^K\frac{\partial\pi}{\partial\tau^K})(1-\delta\beta_t^p(1-\pi\tau^K)-\delta\pi\tau^K)$. Note that $\mathcal{A} > 0$ and $\mathcal{C} < 0$ if $\pi + \tau^K\frac{\partial\pi}{\partial\tau^K} > 0$, which implies $\varepsilon_\tau^\pi > -1$. In this case,

$$\frac{\partial \omega}{\partial \tau^K} < \frac{\partial \omega}{\partial \tau^D} < 0$$

On the other hand, $\tau^D$ does not affect beliefs formation either. Consider the subjective price model specified before, which implies that beliefs are fully characterized by equation (18). It is clear that the path of capital gains expectations does not depend on $\tau^D$. Even if agents learn about dividends, dividends beliefs would not depend on $\tau^D$ as the later does not affect dividends that are assumed to be exogenous. ■

Altogether, this section has shown that capital gains taxes depress stock prices when investors do not adjust their capital gains realizations too much when taxes change, as seems to be the case in reality (see Agersnap and Zidar (2021)). When that negative relationship holds, capital gains taxes stabilize stock prices.

## 3.- Quantitative Analysis

This section uses the theory to understand important changes in the US stock market since the 1980s as a consequence of the decline in capital taxes. In Section 3.1, I document two major changes: a rise in the stock market valuation and in its volatility. A new version of the Campbell and Shiller (1988) PD variance decomposition with taxes is presented. Section 3.2 presents an extended version of the model. Section 3.3 describes the estimation results and its robustness in a number of dimensions. Section 3.4 presents evidence consistent with the model main mechanism, that is, taxes affect both the sensitivity of prices to beliefs and the beliefs formation process.

### 3.1.- Facts

This section documents two major changes in the stock market since the 1980s that cohabited with a decline in capital taxes. First, the PD ratio went up due to the stronger growth of capital gains over dividends. Second, using an extended version of the Campbell and Shiller (1988) variance decomposition, I show that the PD volatility went up as a result of both the decoupling of returns from dividend growth and tax cuts.

**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).[28] Thus, the dividends tax rate is

$$\tau_t^D = \tau_t^d(1-\eta_t) \tag{21}$$

the capital gains tax

$$\tau_t^K = (\phi\tau_t^{skg} + (1-\phi)\tau_t^{lkg})(1-\eta_t) \tag{22}$$

and finally, the interest tax

$$\tau_t^B = \tau_t^b(1-\eta_t) \tag{23}$$

In the previous expressions, $\tau_t^d$, $\tau_t^{skg}$, $\tau_t^{lkg}$ and $\tau_t^b$ are the effective average marginal rates on dividends, short, long capital gains and interest income respectively; $\phi$ is the average weight of short capital gains on total capital gains; $\eta_t$ 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

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

to a massive change in asset holdings from taxable to non-taxable accounts (see McGrattan and Prescott (2005)).[29,30]

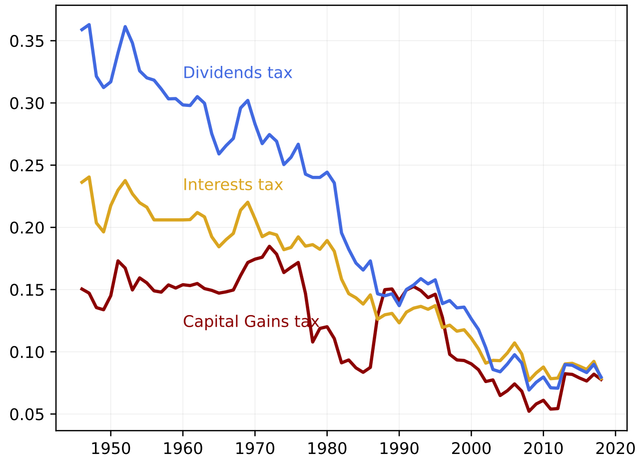


*Figure 2:* ***Capital taxes rates along the postwar period****. The graph plots the capital taxes on dividends (blue), interest income (yellow) and capital gains (red) as defined by equation 21, 22 and 23. Annual series 1946-2018. See Appendix A for data sources and Appendix C for details on the computations.*

**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[31].

**Fact 3: Rise in valuations volatility.** The rise in levels have gone hand in hand with larger fluctuations of the PD ratio.[32] Indeed, the PD ratio standard deviation turns out to be two and half times higher after 1982 than before.[33] Thus, exuberance (higher valuations) has gone hand in

[29]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.

[30]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))

[31]This fact has been extensively documented among others by ...

[32]This fact has received much less attention. To the best of our knowledge, it has just been pointed out by Adam (2020).

[33]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).

hand with instability (larger fluctuations).

To understand the variation in the PD ratio, the literature has resorted to a so-called "dynamic accounting equation", first derived by Campbell and Shiller (1988) to analyze the PD variations. The CS Equation (after Campbell-Shiller) reads as follows:

$$p_t - d_t \approx \text{constant} + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}}_{\bar{d}_t} - \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}}_{\bar{r}_t} \tag{24}$$

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.[34]. However, this dynamic equation overlooks the role of taxes in shaping returns. To overcome this absence, a version of the CS Equation with taxes is derived which allows to decompose the contribution of pre-tax returns between after-tax returns and taxes. Here I sketch the key steps in the derivation; the complete procedure can be found in Appendix XX. First, start with this identity

$$1 = \hat{r}_{t+1}^{-1} \hat{r}_{t+1} \tag{25}$$

with

$$\hat{r}_{t+1} = (1 - \pi\tau_{t+1}^{K}) \Big( \frac{P_{t+1} - P_t}{P_t} \Big) + (1 - \tau_{t+1}^{D}) \frac{D_{t+1}}{P_t} \tag{26}$$

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

$$\frac{P_t}{D_t} = \tilde{R}_{t+1}^{-1} \frac{D_{t+1}}{D_t} \Big( (1 - \pi\tau_{t+1}^{K}) \Big( \frac{P_{t+1}}{D_{t+1}} \Big) + 1 - \tau_{t+1}^{D} \Big) \tag{27}$$

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

$$p_t - d_t = -\tilde{r}_{t+1}^{-1} + \Delta d_{t+1} + \ln \Big( e^{ln(1-\pi\tau_{t+1}^{K})} e^{p_{t+1} - d_{t+1}} + e^{ln(1-\pi\tau_{t+1}^{D})} \Big) \tag{28}$$

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

[34] The decomposition raised by Campbell and Shiller (1988) and Cochrane (1992) assumes the stationarity of the log PD ratio. It is not clear that this is the case. The lack of stationarity is a potential problem since the decomposition is based on an expansion around the long-term mean, which may not exist. Nevertheless, it turns out that the approximation works quite well, as set out by Dybvig and Zhang (2018) and showed in this paper.

is obtained:

$$\begin{aligned} p_t - d_t \approx & \text{constant} + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j}) - \sum_{j=1}^{\infty} \rho^{j-1}(\tilde{r}_{t+j}) \\ & + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \rho \ln(1 - \pi\tau^K_{t+j})}_{\equiv \bar{\tau}^K_t} + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \tilde{\rho} \ln(1 - \tau^D_{t+j})}_{\equiv \bar{\tau}^D_t} \end{aligned} \tag{29}$$

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:

$$\begin{aligned} \mathbb{V}ar(p_t - d_t) \approx \quad & \mathbb{C}ov\Big(p_t - d_t, \bar{d}_t\Big) - \mathbb{C}ov\Big(p_t - d_t, \bar{r}_t\Big) \\ & + \mathbb{C}ov\Big(p_t - d_t, \bar{\tau}_t{}^K\Big) + \mathbb{C}ov\Big(p_t - d_t, \bar{\tau}_t{}^D\Big) \end{aligned} \tag{30}$$

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 $\rho$.

This decomposition allows to address at least two questions of interest for the topic at hand. First: Do tax changes play an important role or, very much like future dividends, are largely irrelevant to explain PD fluctuations? Second: How much of the rise in PD volatility can be directly linked to tax changes? Table 1 compiles the estimation outcomes of the CSC Equation with and without taxes for different samples.[35] The first message is that taxes account for an important chunk of the PD variance. For the whole postwar period, the well-established result that returns explain almost all the variance is only true when taxes are ignored. 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. In other words, an important part of the PD ratio fluctuations typically attributed

[35] 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). I impose short-run restrictions, with the variable ordering being dividends, taxes, returns and the PD ratio. It is common practice to exclude returns from the VAR, and recover them as a residual instead (i.e., the part non-explained by dividends and taxes). The reason is that including returns would give rise to very high multicollinearity, if the approximation works well. However, that procedure imposes an exact approximation, biasing the true returns share. I check that the approximation error is time-varying, so including returns is not so harming in terms of collinearity. Anyway, results are very similar when excluding returns from the VAR.

to movements in discount rates seems related to changes in capital taxes instead.

Second, this decomposition is helpful to figure out the sources of the PD volatility rise. 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). Thus, higher valuations anticipated future high -not low- dividend growth since the 1980s, which makes the PD variance move up.[36] The second factor is the decline in capital taxes which would induce PD ratio adjustments. Thus, tax volatility would be responsible for about 20% of the increase in PD volatility. Finally, a part of these effects is driven by a greater discount factor.[37] These statistics are also collected in table 2, documenting Fact 3.

*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 (30). 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 $\rho$ | 0.9784 | | 0.9726 | | 0.9815 | |

**Relation to the theory.** Can the theory exposed in Section 2 help to explain these facts? First, by the capitalization hypothesis, the decline in taxes would generate a rise in prices for a

[36]An implication of this change is that returns and dividend growth got decoupled since the 1980s, varying in opposite directions; high valuations still forecast low future returns but not low dividend growth. Then, it must be that returns became more dependent on capital gains.

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

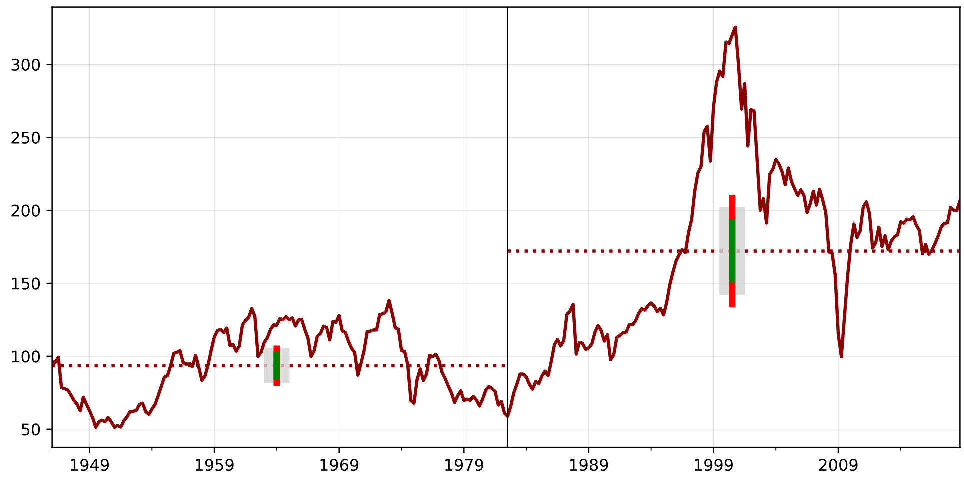


*Figure 3:* ***Change in the mean and standard deviation of the Price/Dividend ratio.*** *The graph plots the evolution of the PD ratio in the 1946:I-2018:IV period. The dotted lines plot the mean of each subperiod 1946:I-1982:II and 1982:III - 2018:IV. The gray box depicts the standard deviation. The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

given level of dividends. In other words, the hypothesis implies that lower taxes would increase the PD ratio by increasing capital gains. Besides, an implication of the hypothesis is that tax volatility would translate into PD volatility, in line with what was documented. Furthermore, the tax volatility hypothesis conjectures that lower taxes would also boost PD fluctuations by turning prices more sensitive to beliefs. By this channel, the model has a chance to link the change from negative to positive covariance between PD and future dividends as a result of lower taxes.

### 3.2.- Extended model and a new solution algorithm.

This section extends the simple model set up in Section 2 to better equip it to replicate the stylized facts reported in the previous section. Besides, it introduces a new algorithm to solve based on the Parameterized Expectations Algorithm.

The model described in Section 2.1 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 $\gamma$ regulating its risk aversion level. Second, an additional source of exogenous income is introduced, to avoid a too high correlation between dividends and consumption. In particular, it is assumed agents get a wage endowment $W_t$ each period, following

*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 |

this process:

$$\ln\Big(1+\frac{W_t}{D_tS_t}\Big)=(1-p)\ln(1+\rho)+p\ln\Big(1+\frac{W_{t-1}}{D_{t-1}S_{t-1}}\Big)+\ln\varepsilon_t^w \tag{31}$$

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

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

Third, the aggregate stock supply is assumed to evolve stochatiscally according to the following process:

$$S_t=s_0+s_1S_{t-1}+\varepsilon_t^s \tag{32}$$

with $s_0<0$, $s_1<1$ such that it fluctuates around a long-run mean given by $\bar{S}=\frac{s_0}{1-s_1}$ and $\varepsilon_t^s\sim\mathcal{N}(0,\sigma_s^2)$. Stock changes can be interpreted as the net balance between stock issuance and repurchases. 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 set up in Section 2 (determined by equation (14) and (15), with the updating equation (16)).

**Competitive Equilibrium**. Given initial endowments $S^i_{-1} = S_{-1}$, and the probability measure $\{\mathcal{P}_i\}_{i=1}^I$ (involving the dividends, wage and stock supply processes (expressions (1), (31), (32), the price model (14) and the updating equation (16)), a Competitive Equilibrium consists of sequences of allocations $\{\{C^i_t, S^i_t, B^i_t\}_{t=0}^\infty\}_{i=1}^I$ and prices $\{P_t\}_{t=0}^\infty$ such that:

1. Investors behave optimally, satisfying:

   (a) KKT First Order Conditions. They boil down to the following Euler Equations[38]

$$(C^i_t)^{-\gamma} = \delta \mathbb{E}^{\mathcal{P}_i}_t \left( [P_{t+1} + D_{t+1} - \tau^D_{t+1} D_{t+1} - \pi \tau^K_{t+1}(P_{t+1} - P_t))] P_t^{-1} (C^i_{t+1})^{-\gamma} \right) \tag{33}$$

$$(C^i_t)^{-\gamma} = \delta(1 + (1 - \tau^B_t) r^b_t) \mathbb{E}^{\mathcal{P}_i}_t \left( (C^i_{t+1})^{-\gamma} \right) \tag{34}$$

   and the budget constraint:

$$C^i_t + P_t S^i_t + B^i_t \leq W_t + (P_t + D_t) S^i_{t-1} + (1 + (1 - \tau^B_t) r^b_{t-1}) B^i_{t-1} + T_t - (\tau^D_t D_t + \pi \tau^K_t (P_t - P_{t-1})) S^i_{t-1} \tag{35}$$

   (b) A transversality condition:

$$\lim_{j \to \infty} \left( \frac{\delta}{1 - \delta \pi \tau^K_t} \right)^j \mathbb{E}^{\mathcal{P}_i}_t \left[ \frac{C^i_{t+j}}{C^i_t} (1 - \pi \tau^K_{t+j}) P_{t+j} S_{t+j} \right] = 0 \tag{36}$$

2. Markets clear:

$$\text{Equities: } \int_0^1 S^i_t di = S_t \tag{37}$$

$$\text{Bonds: } \int_0^1 B^i_t di = 0 \tag{38}$$

$$\text{Goods: } \int_0^1 C^i_t di = S_t D_t + W_t \tag{39}$$

**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, $\boldsymbol{X}_t = (D_t, W_t, \tau_t, T_t, S_{t-1}, S_t, P_t, \beta^p_t)$[39]. Given

[38] 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.

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

the homogeneity property of the CRRA function, the state vector can be reduced to $\boldsymbol{X}_t = (\frac{W_t}{D_t S_t}, \tau_t, T_t, S_{t-1}, S_t, \frac{P_t}{D_t}, \beta_t^{\mathcal{P}})$. 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.** The proof of existence of a recursive equilibrium stated in Adam et al. (2017) for the same model without taxes continuous to hold in the model with taxes. Therefore, the recursive solution boils down to a time-invariant stock demand function $S_t = S(\boldsymbol{X}_t)$. The main difficulty to derive such invariant function is that optimality conditions includes an unknown conditional expectation. Thus, the optimal consumption-dividends ratio must satisfy

$$\frac{C_t}{D_t} = \left\{\delta \mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{P_{t+1}}{D_{t+1}} + 1 - \tau_{t+1}^D - \pi\tau_{t+1}^K\left(\frac{P_{t+1}}{D_{t+1}} - \frac{P_t}{D_t}\frac{D_t}{D_{t+1}}\right)\right)\frac{D_t}{P_t}\left(\frac{C_{t+1}}{D_{t+1}}\right)^{-\gamma}\left(\frac{D_{t+1}}{D_t}\right)^{1-\gamma}\right]\right\}^{-1/\gamma} \tag{40}$$

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

$$\frac{C_t}{D_t} = \left(\delta \mathcal{E}(\boldsymbol{X}_t)\right)^{-\frac{1}{\gamma}} \equiv \bar{\mathcal{E}}(\boldsymbol{X}_t) \tag{41}$$

To solve the model, $\bar{\mathcal{E}}(\boldsymbol{X}_t)$ must be computed somehow. The Parameterized Expectations Algorithm (PEA) is one of the alternatives[40]. PEA consists of replacing the conditional expectation $\mathcal{E}(\boldsymbol{X}_t)$ by some parametric function $\psi$. $\psi$ is not unique; popular possibilities are polynomials, splines or neural networks. In this model, there is no practical difference between approximating the conditional expectation $\mathcal{E}(\boldsymbol{X}_t^i)$ and approximating the policy function $\bar{\mathcal{E}}(\boldsymbol{X}_t^i)$. 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 $\psi$:

$$\frac{C_t^*}{D_t} = \bar{\mathcal{E}}(\boldsymbol{X}_t) \approx \psi(\boldsymbol{X}_t; \chi) = c_t Z_t \tag{42}$$

where $c_t \equiv (1 - \chi\delta(1 - \pi\tau_t^K)\beta_t^i)$ 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[41]) and $\chi$ is a parameter of $\psi$ to

[40] 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).

[41] In the baseline case, I set $T_t = 0$. Tax rebates are explored in the robustness analysis.

be estimated. Then, the stock policy can be obtained using the budget constraint

$$S_t^* = S(\boldsymbol{X_t}) \approx (1 - c_t) Z_t \frac{D_t}{P_t} \tag{43}$$

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 discounting mechanism outlined in Section 2 that explicitly uses investors' demand as mediator. Figure 4 illustrates these mechanics.

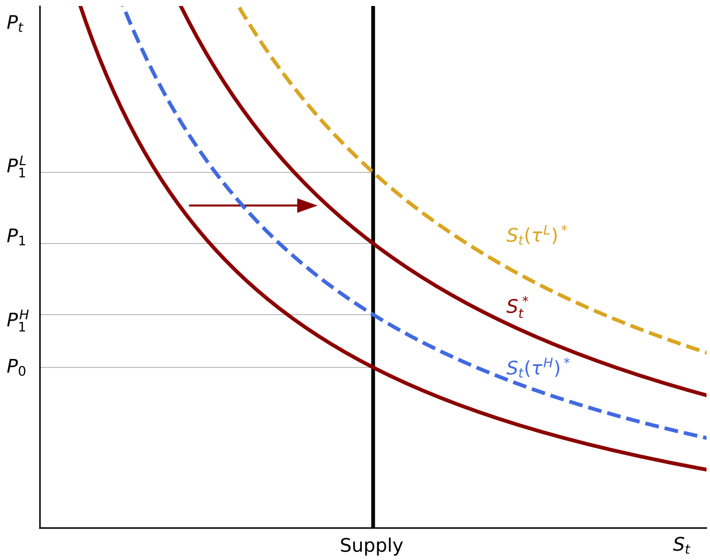


*Figure 4:* ***Response of stock demand to an increase in optimism at different tax levels.*** *The graph plots the stock policy function (equation (43)) keeping everything constant except prices. Then, as $\beta_t^P$ increases, the curve moves rightwards. This displacement is shown at three different tax levels: low (blue), moderate (the baseline, in red) and high (yellow).*

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

$$\frac{P_t}{D_t} = \frac{(1 - c_t)\left(\frac{W_t}{D_t S_t} S_t + \left(1 - \tau_t^D + \pi \tau_t^K \frac{P_{t-1}}{D_{t-1}} \frac{D_{t-1}}{D_t}\right) S_{t-1}\right)}{S_t - (1 - c_t)(1 - \pi \tau_t^K) S_{t-1}} \tag{44}$$

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.[42] 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 XX 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; the parameterization of the model. The former is solved by the introduction of a projection facility. The model is parameterized following a mix strategy, with some parameters being picked directly from the US data and the rest being estimated via the Simulated Method of Moments.

The model simulation requires to deal with two issues. On the one hand, the equilibrium PD ratio faces a discontinuity. As is standard in this 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 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 15 parameters; a subset of them is picked from US data and the rest is estimated. Specifically, the vector $\widetilde{\theta} = \{a, \sigma_D, \sigma_W, \sigma_{WD}, p, \rho, \pi, \sigma_s\}$ is picked directly from US data. I calibrate $a, \sigma_D, \sigma_W, \sigma_{WD}$ distinguishing between the two studied subperiods. 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 $\pi$, which is not directly observable. To find empirical counterparts, I decompose it into three ratios:

$$\pi \equiv \frac{\text{Equity RKG}}{\text{Equity KG}} = \frac{\text{Equity RKG}}{\text{RKG}} \frac{\text{RKG}}{\text{KG}} \frac{\text{KG}}{\text{Equity KG}} \tag{45}$$

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, re-

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

spectively.[43] 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 $\pi$ is equal to 8%, that is, only 8% of equity capital gains were realized and then, taxable. Thus, 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]

The remaining parameters, collected in the vector $\theta = \{\delta, g, \gamma, s_1, \bar{S}, PD^L, PD^U\}$, are estimated via the Simulated Method of Moments. Aiming at testing the power of taxes to explain the various observed changes, all of them are kept fixed throughout the two subperiods. 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 a total of M=18 moments. The vector $\theta$ is chosen as to minimize the distance between model $\widetilde{\mathcal{S}}(\theta)$ and data $\hat{\mathcal{S}}$ statistics, that is,

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[\hat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta)\right]' \hat{\Sigma}_{\mathcal{S}}^{-1} \left[\hat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta)\right] \tag{46}$$

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. Appendix Z offers more details about statistics generation.

Finally, the model is fed with the empirical time series for capital taxes. Thus, the simulated series can potentially exhibit the same trajectories and trends as the observed ones.[45]

[43] The time coverage is unequal, though. IRS data only covers the 1997-2012 period along with the year 1985, whereas the US Financial Accounts cover the whole postwar period. See Appendix A for more details.

[44] The idea that the effective tax on capital gains is just a fraction of the statutory rate is in line with the literature; see Sialm (2009) for references. Besides, 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] If, instead, average tax rates are introduced as a parameters and the two subsamples are simulated as two long lasting regimes, 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 3:* ***Benchmark parameterization.*** *This table reports the values of the model parameters used for the quantitative analysis. Panel a) lists the calibrated parameters, using various data sources specified in Appendix A. Panel b) lists the estimated parameters obtained from the SMM estimation of the model. Column "Without $r_t^b$" shows the parameters for the baseline case excluding the risk-free rate; column "With $r_t^b$" shows $\hat{\theta}$ when the risk-free rate is included.*

| a) Calibrated parameters | | 1946-1982 | 1982-2018 |
|---|---|---|---|
| Mean dividend growth | $a$ | 1.0049 | 1.0075 |
| Dividends growth standard deviation | $\sigma_D$ | 0.0252 | 0.0197 |
| Wage shocks standard deviation | $\sigma_W$ | 0.0261 | 0.0196 |
| Covariance (wage, dividend) | $\sigma_{WD}$ | -0.0006 | -0.0004 |
| Persistence consumption-dividend ratio | $p$ | 0.99 | |
| Average consumption-dividend ratio | $1+\rho$ | 20.03 | |
| Fraction of realized capital gains | $\pi$ | 0.08 | |
| **b) SMM estimated parameters $\hat{\theta}$** | | **Without $r_t^b$** | **With $r_t^b$** |
| Discount factor | $\delta$ | 1.00 | 1.00 |
| Kalman gain | $g$ | 0.335 | 0.336 |
| Risk aversion | $\gamma$ | 1.30 | 0.56 |
| Stock long-run mean | $\bar{S}$ | 4.41 | 3.99 |
| Stock supply persistence | $s_1$ | 0.98 | 0.95 |
| Projection facility start | $PD^L$ | 230.07 | 267.65 |
| Projection facility upper bound | $PD^U$ | 424.01 | 489.04 |

## 3.4.- Baseline results

In this section, the baseline results are reported and analyzed. Table 4 contains the statistics from the US data and the estimated models. The upper part details the targeted statistics using observed and simulated data from the model and its Rational Expectations counterpart; besides, additional non-targeted statistics are reported. The estimated parameter vector $\hat{\theta}$ is reported in the column "Without $r_t^b$" of table 3's panel b).

Broadly speaking, the learning model produces variables that move in line with the documented changes, passing the individual t-test in most of the cases. I articulate the discussion in 3 dimensions: the PD level, the PD volatility and the equity premium. First, the model produces on average[46] about half of the increase in the PD level due to its ability to generate capital gains

[46] An alternative is to look at the whole distribution of outcomes produced by the simulated model. In this case, the PD ratio 5th percentile and 95th percentile for the 1946-1982 period are 21.25 and 33.80, respectively and 31.71 and 43.33 for the 1982-2018 period.

larger than the dividend growth. Second, the model matches well both the level and the change in PD volatility. The reason is that it replicates decently the reduction in the covariance between the PD ratio and both future dividend growth and future taxes. However, an important part of the increase in the PD variance results from a change in the covariance between the PD ratio and future returns that, although within the confidence bands, is somehow excessive. A related observation is that the volatility of returns is too low in the first period. This simply points out a limitation of the main mechanism: while tax cuts certainly increased volatility, quantitatively it does not explain all the rise in volatility; in this event, the model produces the needed volatility by resorting to an increase in the expectations volatility, which leads to a certain mismatch of the PD volatility sources. Finally, although the model overstates the risk-free rate, it produces a decent equity premium (about 90% of the observed one) as well as replicates its decline.

To have a more traditional benchmark, the outcomes implied by a Rational Expectations version of the model are reported in the last two columns of table 4. The RE version is also capable to produce an increase in the mean PD ratio, although overstating its level. Nonetheless, it performs badly in the other two dimensions. On the one hand, RE fails to generate neither enough PD volatility nor a remarkable increase in it. This result illustrates the theoretical reasoning of Section 2: with RE the tax level becomes neutral and only tax changes spur PD volatility, which proves utterly insufficient. Besides, RE equity premium is only about one third of the observed one and goes up instead of down. The main reason of that is the lack of volatility; see next section.

Results are very similar when the risk-free rate is included in the estimation targets, as reported in table 5. 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, as shown in panel's b) right column of table 3, such that volatility is increased. 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 the next section.

Finally, to isolate the effects of tax changes from those coming from the increase in mean dividend growth, table 6 reports a counterfactual exercise. First, the learning model is simulated with the baseline parameterization but keeping dividends and capital gains taxes constant. In this case, the increase in the mean PD ratio is negligible, the PD volatility goes down instead of up and the equity premium goes up instead of down. In other words, given the dividend process, capital taxes seem decisive to explain the rise in mean valuations and their fluctuations. The next columns explores the different role of the different taxes. Loosely speaking, while both dividends and capital gains taxes are equally responsible for the increase in stock market valuations the latter are way more important to explain the rise in volatility. In fact, when the capital gains tax is kept at its 1946 level, volatility is low and does not rise whereas when the dividends tax is fixed the volatility is a bit lower but essentially the same increase is observed. That illustrates quantitatively the particular relevance of capital gains taxes pointed out in the theory. Lastly, I simulate the model keeping the mean dividend growth $a$ at its mean 1946-1982 level always (instead of increasing it in the second subsample). In this case, the PD volatility doubles but its level is too low. All these counterfactuals suggest that taxes alone would explain about 50% of the rise in the model's volatility with the rest coming from the interaction between lower taxes and higher dividend growth.

### 3.4.- 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[47]

$$\left(\prod_{t=1}^{N}\frac{P_t+D_t}{P_{t-1}}\right)^{\frac{1}{N}}=\underbrace{\left(\prod_{t=1}^{N}\frac{D_t}{D_{t-1}}\right)^{\frac{1}{N}}}_{R_1}\underbrace{\left(\frac{PD_N+1}{PD_0}\right)^{\frac{1}{N}}}_{R_2}\underbrace{\left(\prod_{t=1}^{N-1}\frac{PD_t+1}{PD_t}\right)^{\frac{1}{N}}}_{R_3} \tag{47}$$

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 7 reports the decomposition using empirical and simulated data. Since the dividend

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

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 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)[48]. 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[49].

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

$$r_t^b = \left( \frac{1}{\delta \mathbb{E}_t \left[ \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]} - 1 \right) \frac{1}{1 - \tau_t^b} \tag{48}$$

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

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 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

[48] 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).

[49] 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.

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 $\tau^b$[50]. 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.

### 3.5.- Robustness tests

Tax anticipation. Tax rebates. Tax and dividends learning. HA. Table 8.
UNFINISHED.

### 3.6.- Testing the mechanism

This section presents empirical evidence of the model's novel mechanism. According to the model, tax cuts increased PD volatility because they incremented both the sensitivity of prices to beliefs and beliefs fluctuations, other things equal. While previous sections have tested the model in terms of the relationship between taxes and PD volatility, the suggested mechanism has not been tested yet. To do so, an IRFs analysis is proposed, comparing the response of the PD ratio, capital gains expectations and the expectations elasticity of the PD ratio to a capital gains negative tax shock using both simulated and observed data. Remarkably, the observed data uses survey expectations, offering then a model-free empirical benchmark.

To get analytical intuition, let's retrieve the model with capital gains learning of Section 2. The sensitivity of prices to beliefs can be measured in terms of a time-varying elasticity that reads as

$$\epsilon_t \equiv \frac{\partial P_t/D_t}{\partial \beta_t^p} \frac{\beta_t^p}{P_t^H/D_t} = \frac{\delta \beta_t^p (1-\tau^K)}{1-\delta(1-\tau^K)\beta_t^p - \delta\tau^K} \tag{49}$$

[50]In this case, there is no feedback loop affecting the bond price due to its one-period maturity. Hence, $\tau^b$ level is neutral in the sense of Section 2 and only tax changes have an impact on bond prices. These impact is very small so that the risk-free rate is very stable.

It turns out that this elasticity depends on the capital gains tax level. In particular,

$$\frac{\partial \epsilon_t}{\partial \tau^K} < 0 \text{ if } \delta < 1 \tag{50}$$

Besides, from the difference equation (18) governing expectations dynamics, it can be shown that

$$\frac{\partial \beta_t}{\partial \tau} < 0 \text{ if } \delta^2(1-\tau)(\beta_{t-1}-1) < 1 \tag{51}$$

The condition is satisfied for standard values of the parameters. These two partial derivatives give us the sign of the reaction of the objects of interest to a tax change. To get the dynamic response, I resort to simulations. In order to apply the same treatment to simulated and observed data, I run a VAR with capital gains taxes, the PD ratio, expectations and the time-varying elasticity using both simulated and observed data[51]. Simulated expectations are easily recoverable from the model; observed expectations are taken from surveys[52] Moreover, the time-varying elasticity can be obtained using a state-space model as

$$\begin{aligned} \ln\frac{P_t}{D_t} &= \mu + \tilde{\epsilon}_t \ln\beta_t + v_t^p \\ \tilde{\epsilon}_t &= \tilde{\epsilon}_{t-1} + w_t^{\tilde{\epsilon}} \end{aligned} \tag{52}$$

where $\tilde{\epsilon}_t$ is the desired elasticity, $\beta_t$ 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 $\{\tilde{\epsilon}_t\}$[53]. Figure 5 plots the elasticity.

[51]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.

[52]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).

[53]In the model, the elasticity could be pinned down analytically. Instead, I applied the same techniques as with empirical data, that is, I computed the time-varying elasticity using the aforementioned state-space model and run the elasticity regression. In this way, the statistics from the observed and simulated data are fully comparable.

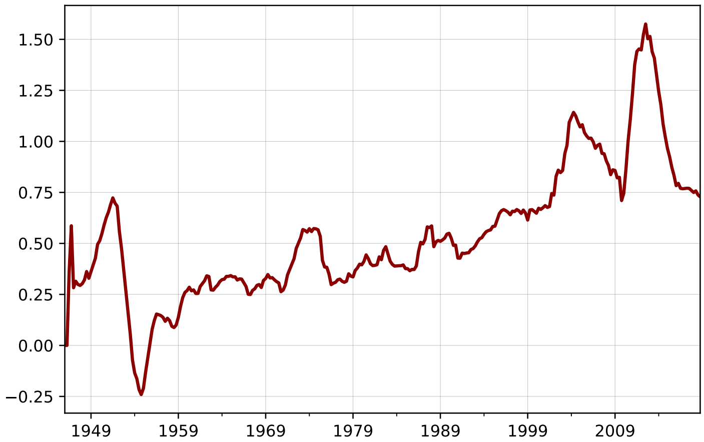


*Figure 5:* ***Estimated expectations elasticity of the PD ratio.*** *The line plots the time series of* $\{\tilde{\epsilon}_t\}$ *as defined in equation 52, using the parameters from the MLE (see Appendix E). The time series cover the 1946:I-2018:IV period.*

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

$$\boldsymbol{Y}_t = A + B\boldsymbol{Y}_{t-1} + \boldsymbol{u}_t \tag{53}$$

with $\boldsymbol{u}_t \sim \mathcal{N}(0, \Omega)$ and $\boldsymbol{Y}_t = [\tau_t^K, \beta_t, P_t/D_t, \tilde{\epsilon}_t]$. I impose $\boldsymbol{u}_t = S\bar{\boldsymbol{u}}_t$ with $S$ being the Cholesky factor of the matrix $\Omega$ and $\bar{\boldsymbol{u}}_t \sim \mathcal{N}(0, I)$ . Figure 6 shows the response of expectations, the PD ratio and the elasticity 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. Remarkably, the model-implied responses are within the empirical confidence bands despite these statistics have not been targeted. Altogether, the observed data seems in line with the model predictions[54,55]. Appendix XX explores the relationship between taxes and survey expectations using all available surveys[56].

[54]This result suggests an explanation for an issue pointed out by Adam et al. (2017). They showed that the correlation between prices and beliefs turned stronger after the year 2000. They hypothesized it may have to do with lower interest rates. Here, I explore that relationship more generally (with time-varying coefficients), showing that the rising trend started before and it seems a product of tax cuts, at least in part (in my regression, $R^2 = 0.52$, so there is still room for other factors).

[55]In a recent work, Giglio et al. (2021) have showed that lower taxes makes portfolio decisions more sensitive to beliefs. Thus, the results in this paper complements their evidence: both stock holdings and price sensitivity to expectations seems to depend on taxes.

[56]I use the Duke CFO Global Business Outlook, Shiller survey for both institutional and individual investors, the Livingston survey from the Philadelphia Fed and an alternative extension of the UBS/Gallup survey . The positive reaction of expectations to a negative tax shock broadly hold. However, it turns out there is a positive correlation between tax changes and expectations since the Great Recession (small tax increases happened together with a recovery in expectations) that deliver the opposite result for surveys with short time span.

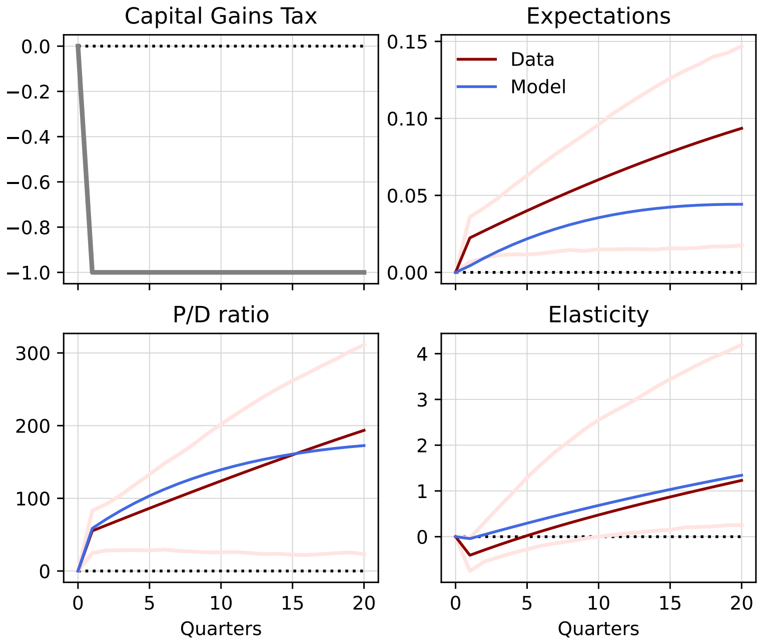


*Figure 6:* ***Responses to a persistent reduction in*** $\tau^K$***.*** *The figure plots the response of different variables to a 1pp persistent reduction in the capital gains tax. Red lines are the responses using the observed data for the US. Blue lines are the responses implied by the quantitative model with the baseline parameterization of table 3. Bands at 68% confidence level have been computed via data simulations using 1000 repetitions.*

### 3.7. Taxes and the stock market fragility

In this section, another prediction of the model that relies only on observables. According to the model, the response of the PD ratio to shocks, for instance to a price growth shock, 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 would lean against this amplification.

Thus, a prediction of the model would be that the response of the PD ratio to a shock, for instance a price growth shock, is significantly higher in the period with lower taxes. To test it, a minimalist VAR model with price growth and the PD ratio is estimated for the two subsample. 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 7 illustrates.

### 3.8.- A historical test of competing hypothesis

In this section, I study the properties of the model when using the observed time series not only of capital taxes but also of price and dividend growth and the wage-dividends ratio. I use this

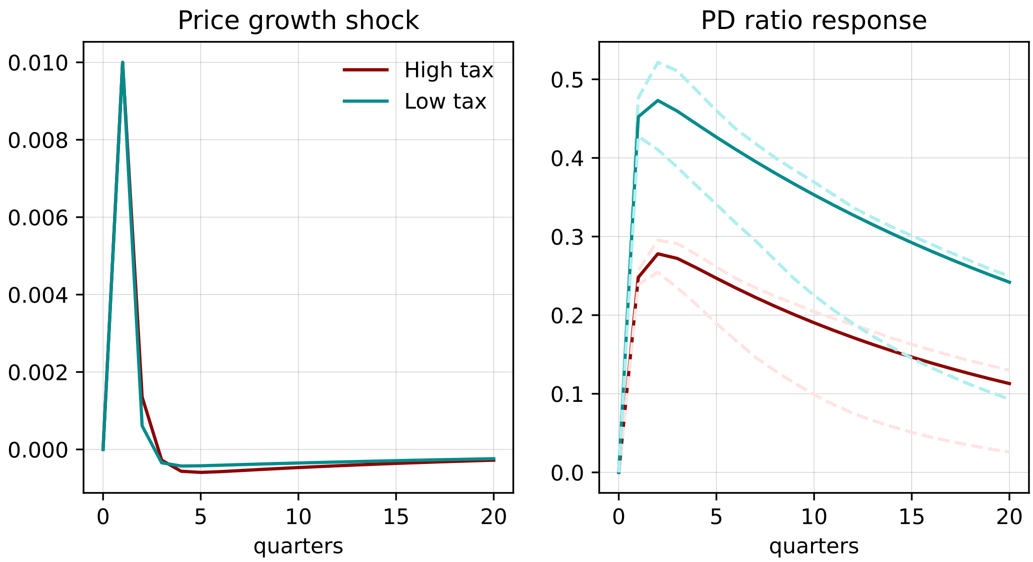


*Figure 7:* ***Empirical response of the PD ratio to a price shock in different tax regimes.*** *The graph plots the response of the PD ratio to an equivalent price shock in two different subsamples: red lines signals the 1946-1982 sample characterized by higher taxes; blue lines are for the 1982-2018, characterized by lower taxes. 68% confidence bands are computing by simulations using 1000 repetitions.*

historical calibration of the model to evaluate some alternative candidates to explain the increase in volatility, particularly the fall in safe real interest rates. Through the lens of the model, interest rates played a minor role in the whole story, though.

The connection between interest rates and asset prices booms and busts have been suggested in some papers (e.g., Taylor (2007)) 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 restrict the discussion to its effects on the discount factor, which is the case made by Adam and Merkel (2019).[57] By the same analysis of section, it is simple to show that $\delta$ acts in an equivalent way to $\tau^K$: other things equal, a lower safe rate would increase the $\delta$ and then the sensitivity of prices to beliefs. As before, this effect is reinforced by capital gains learning. Hence, when considered together, higher taxes can mitigate the effects of lower rates or on the other hand, lower interest rates can amplify the effect of tax cuts.

Now, I describe the procedure. First, I compute a series of model-implied expectations $\{\beta_t\}$ by using the model's updating equation (16) with the historical price growth data. Then, I recover a series of $\{\delta_t\}$ using the model's Euler Equation for bonds (34) with the empirical risk-free time

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

series, consumption growth and the estimated risk aversion. In other words, $\{\delta_t\}$ is calibrated to replicate the historical trajectory of the risk-free rate. Finally, I calibrate a series of stock supply shocks as to get a perfect match between the model-implied and the empirical PD ratio.

With that calibration, I run a counterfactual analysis to assess the marginal contribution of different factors as shown in figure 8 and quantified in table 9. First, capital gains taxes are kept constant at its 1946 level. Thus, in the event of high constant capital gains taxes, the increase in the level and fluctuations of the PD ratio would have been almost entirely avoided, despite the fall in interest rates, the existence of important equity supply shocks or the rise in capital income. This constitutes additional evidence in favour of the model's mechanism. Contrarily, the same experiment for dividend taxes deliver more modest results: higher dividend taxes would have prevented part of the Dotcom bubble and the subsequent high levels in line with McGrattan and Prescott (2005) with only minor effects on volatility.

Besides, the role of interest rates is explored by keeping them at its 1982 level, which is among the highest in the whole period. Although qualitatively equivalent to capital gains taxes, quantitatively its role seems rather secondary, delivering a reduction in volatility of less than 10% compared to a reduction of 50% delivered by a high $\tau^K$.[58] Additionally, supply shocks appear as an important force to explain the rise in levels although not much for volatility. Finally, the Wage-Dividends ratio exhibited a declining trend, in line with the surge of the dividends national income share. The model predicts that a lower Wage-Dividends ratio would have driven down the PD ratio which would have been offset by the boosting effect of lower taxes and interest and negative supply shocks. Unsurprisingly, the lack of a declining trend in the Wage-Dividends ratio would leave alone all the expansionary forces, delivering a wilder stock market with even more exuberant levels and volatility.

## 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

[58] This result must be interpreted within the model's environment. Interest rates can play a particular important role in driving credit cycles that are an important ingredient in asset price booms and busts.

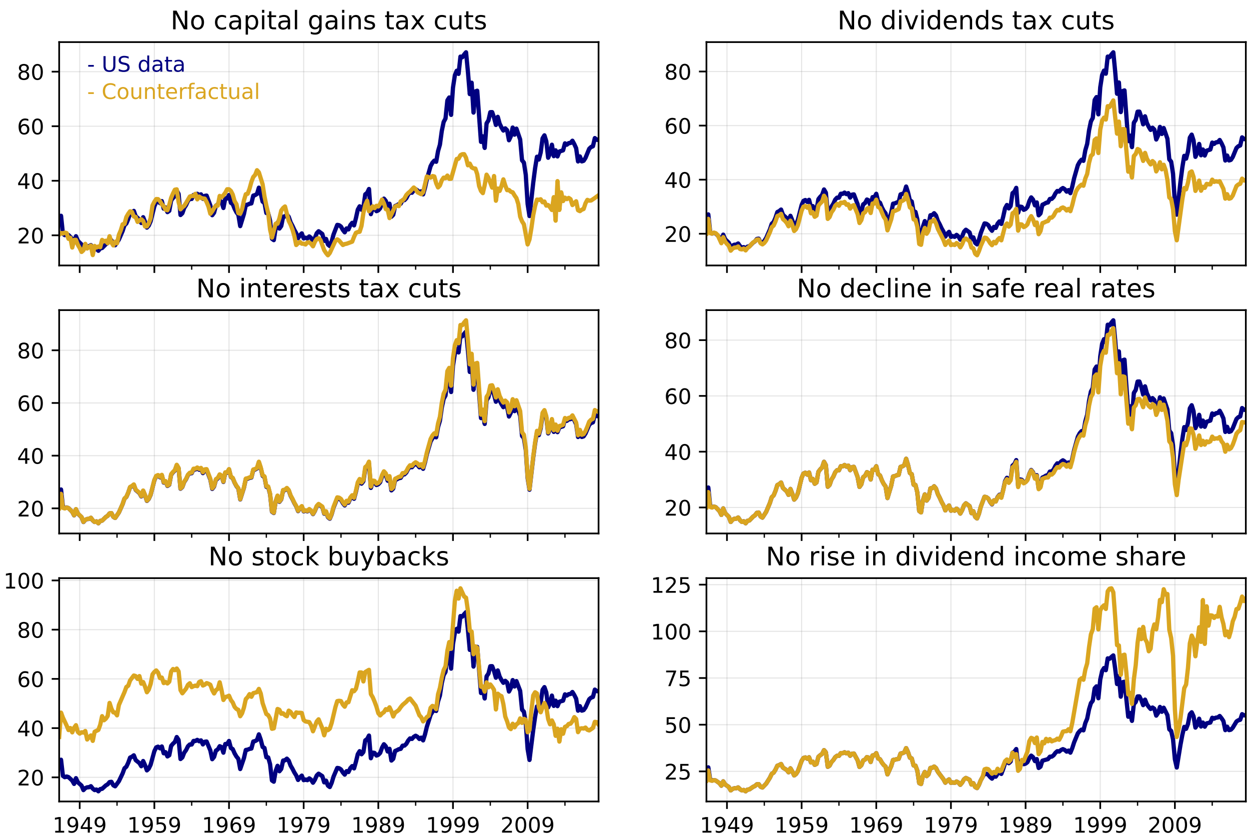


*Figure 8:* ***Historical counterfactuals.*** *The graph plots the US Price-Dividend ratio and its counterfactual path under alternative scenarios using the calibrated historical model.*

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[59]. Thus, the lack of knowledge of the true determinants of prices pushes agents to make decisions using forecasts derived from their subjective models that do not consider the effect of the own forecasts on market prices and everybody else's predictions. In short, smart learning individuals trying to make the best prediction about future prices but missing general equilibrium effects end up causing excessive volatility.

Note, though, this externality is harmless in any of the versions of the model used so far due to the exogenous nature of production. Thus, a precondition to explore an optimal tax is to establish a connection between capital price and consumption fluctuations. In the context of learning models, at least three channels have been suggested: labour demand (Adam and Merkel (2019)); firms' collateral constraints (Winkler (2020)); wealth effects and aggregate demand (Ifrim (2021)). Although the logic I explore would hold similarly in these setups, I suggest another -perhaps more traditional- connection via the neoclassical (or Q) investment theory (built upon Tobin (1969)'s ideas). I do that on two grounds; i) the model's tractability[60]; ii) the good empirical performance of the Q-theory[61]. Thus, in the context of the Q-investment theory, excess price volatility due to learning would boost business cycles and capital gains taxes will have a chance to improve economic welfare.

This exercise is in the spirit of the literature on optimal policy in models with pecuniary externalities more than traditional optimal capital taxation that looks for the best way of financing exogenous sequences of government spending. In the externalities field, research has mostly focused on agents' failure to internalize the effect of their action on other agents borrowing constraints leading to inefficient fluctuations. Capital controls and taxes on borrowing have been proposed to correct these disfunctionalities (e.g., Lorenzoni (2008), Jeanne and Korinek (2010), Dávila and Korinek (2018), Jeanne and Korinek (2019)). While typically the optimal policy has been derived

[59]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.

[60]The previous 3 references are quantitative models. In this section, though, my primary goal is to have an analytical derivation of the optimal tax.

[61]Especially in the last decades and with intangible capital, there has been document a strong relationship between the Tobin's Q and investment. See, for instance, Gutiérrez and Philippon (2016), Peters and Taylor (2017), Andrei et al. (2019) or Melcangi and Sterk (2020).

with respect to a constrained efficient benchmark, recently Benigno et al. (2019) have pointed out that the same instruments can be use better to undo the effects of the borrowing constraint, restoring full efficiency. I follow this approach, deriving the optimal capital gains tax as to undo the effects of learning in prices and allocations.

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 with learning markets. Section 4.4. studies an optimal taxation problem, exploring the use of a tax on unrealized capital gains to restore the First Best allocation in a market economy with incomplete information. Finally, Section 4.5. proposes an alternative implementation of the optimal policy that relies on a constant capital gains tax along with a pretty stable tax on capital rents.

### 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 $G''(\cdot) > 0$[62].

[62] This specification implicitly assumes diminishing returns to scale in adjustment costs. In this way, Hayashi

*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:

$$\max_{\{C_t, I_t, K_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \delta^t U(C_t) \tag{54}$$

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

$$C_t + I_t + G(I_t) \leq F(Z_t, K_{t-1}) \tag{55}$$

ii) Capital-goods resource constraint:

$$K_t \leq I_t + (1-d)K_{t-1} \tag{56}$$

iii) Non-negative consumption:

$$C_t \geq 0 \tag{57}$$

**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 (55)-(56) are satisfied.

2. First order conditions:

$$u_t^c = \lambda_t \tag{58}$$

$$1 + G_t^I = q_t/\lambda_t \tag{59}$$

$$\frac{q_t}{\lambda_t} = \delta \mathbb{E}_t \left[ \frac{\lambda_{t+1}}{\lambda_t} \left( F_{t+1}^k + (1-d) \frac{q_{t+1}}{\lambda_{t+1}} \right) \right] \tag{60}$$

where $G_t^I = \frac{\partial G(I_t)}{\partial I_t}$, $u_t^c = \frac{\partial U(C_t)}{\partial C_t}$, $F_{t+1}^k = \frac{\partial F(Z_{t+1}, K_t)}{\partial K_t}$; $\lambda_t$ 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, $q_t/\lambda_t$ reflects the marginal value of capital (in terms

(1982)'s theorem does not hold. The violation of the theorem can be avoided by assuming $G$ also depends negatively 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.

of goods).

3. A transversality condition.

$$\lim_{j\to\infty} \delta^j \mathbb{E}_t \left[ \frac{u^c_{t+j}}{u^c_t} \frac{q_{t+j}}{\lambda_{t+j}} K_{t+j} \right] = 0 \tag{61}$$

Altogether, the model equilibrium is crucially characterized by the capital market. There is a capital supply function that arises from combining equation (59) . By the properties of $G$, it poses a positive relationship between the stock $K$ and the shadow price of capital $\bar{q} \equiv q/\lambda$. Besides, there Euler Equation can be read as a capital demand function. By the properties of $F$, it poses 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 $\bar{q}$[63].

### 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 $Q_t$[64]. 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.

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

[64] 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

$$\frac{P_t}{D_t} = \frac{(1-d)}{x}\frac{Q_t}{F^k_t} + \frac{1-x}{x} \tag{62}$$

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.

Then, we have to 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, 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 $\Pi_t$. Their static problem can be stated as

$$\max_{\{I_t\}_{t=0}^{\infty}} \Pi_t = Q_t I_t - I_t - G(I_t) \tag{63}$$

**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

$$\max_{\{C_t,K_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \delta^t U(C_t) \tag{64}$$

s.t. i) Budget constraint:

$$C_t + Q_t K_t \leq F(Z_t, K_{t-1}) + (1-d)Q_t K_{t-1} + \Pi_t \tag{65}$$

ii) Non-negative consumption:

$$C_t \geq 0 \tag{66}$$

**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

$$Q_t = 1 + G_t^I \tag{67}$$

2. Households behave optimally, satisfying:

   (a) The sequence of budget constraints (65).

   (b) The sequence of Euler Equation

$$Q_t = \delta \mathbb{E}_t \left[ \frac{u_{t+1}^c}{u_t^c} \left( F_{t+1}^k + (1-d)Q_{t+1} \right) \right] \tag{68}$$

(c) A transversality condition

$$\lim_{j\to\infty} \delta^j \mathbb{E}_t \left[ \frac{u^c_{t+j}}{u^c_t} Q_{t+j} K_{t+j} \right] = 0 \tag{69}$$

3. Markets clear:

$$\text{Goods: } C_t + I_t + G(I_t) = F(Z_t, K_{t-1}) \tag{70}$$

$$\text{Capital: } K_t = I_t + (1-d)K_{t-1} \tag{71}$$

**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

$$1 + G^I_t = \delta \mathbb{E}_t \left[ \frac{\lambda_{t+1}}{\lambda_t} \left( F^k_{t+1} + (1-d)(1 + G^I_{t+1}) \right) \right] \tag{72}$$

which implies

$$Q_t = \bar{q}_t \tag{73}$$

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.- Learning 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 $\sigma$-algebra of Borel subsets of $\Omega$ and $\mathcal{P}^i$ agent's i subjective probability measure over $(\Omega, \mathcal{B})$. For generality, we include prices in the the state space $\Omega$, 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

$$\max_{\{C_t,K_t\}_{t=0}^{\infty}} \mathbb{E}_0^{\mathcal{P}} \sum_{t=0}^{\infty} \delta^t U(C_t) \tag{74}$$

Hence, the Euler Equation reads as

$$Q_t = \delta \mathbb{E}_t^{\mathcal{P}} \left[ \frac{u_{t+1}^c}{u_t^c} \Big( F_{t+1}^k + (1-d) Q_{t+1} \Big) \right] \tag{75}$$

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

$$\frac{Q_{t+1}}{Q_t} \frac{u_{t+1}^c}{u_t^c} = \theta_t + \varepsilon_t^P \tag{76}$$

$$\theta_t = \theta_{t-1} + \nu_t \tag{77}$$

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

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

where $\sigma_\theta^2$ is the steady state Kalman estimate uncertainty and the posterior mean evolves recursively following

$$\beta_{t+1} = \beta_t + g \Big( \frac{Q_t}{Q_{t-1}} \frac{u_t^c}{u_{t-1}^c} - \beta_t \Big) \tag{78}$$

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 (71), the Euler Equation (75) and the expectations updating equation (78). 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.

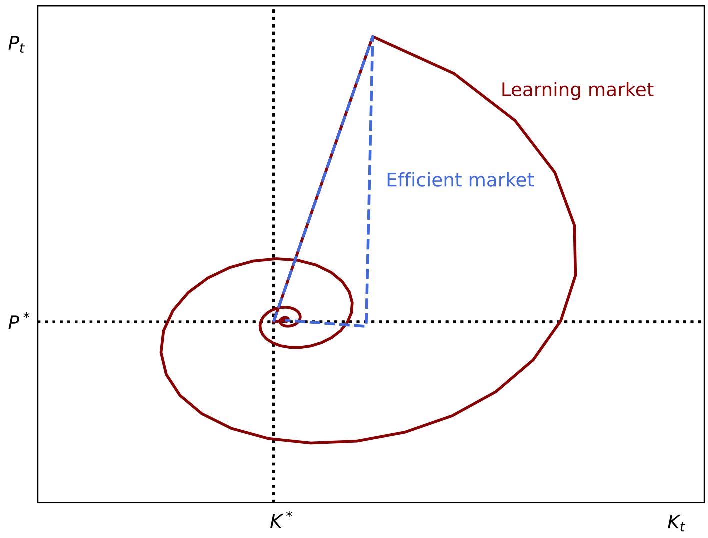


*Figure 9:* ***Response of capital and capital price to a transitory productivity shock under Rational Expectations and Learning.*** *The graph uses a plane with the capital stock on the x-axis and the capital price on the y-axis. Starting in the steady state, the economy is perturbed by a one-off productivity 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.*

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 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 learning 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[65].

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

$$C_t + Q_t K_t \leq F(Z_t, K_{t-1}) + (1-d)Q_t K_{t-1} + \Pi_t - \tau_t^K (Q_t - Q_{t-1})(1-d)K_{t-1} + T_t \tag{79}$$

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

$$\tau_t^K (Q_t - Q_{t-1})(1-d)K_{t-1} = T_t \tag{80}$$

Then, the Euler Equations becomes

$$Q_t = \delta \mathbb{E}_t^{\mathcal{P}} \left[ \frac{u_{t+1}^c}{u_t^c} \left( F_{t+1}^k + (1-\tau_{t+1}^K)(1-d)Q_{t+1} + \tau_{t+1}^K (1-d) Q_t \right) \right] \tag{81}$$

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 (79); the government budget constraint (80); the capital producers' profits equation[66]; goods and capital market clearing conditions (70, 71); the investment function (67); the households' Euler Equation (81); and the beliefs updating equation (78).

[65] 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.

[66]

$$\Pi_t = Q_t I_t - I_t - G(I_t)$$

**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:

$$Q_t^{RE} = \frac{\delta \mathbb{E}_t \left[ \frac{u_{t+1}^c}{u_t^c} F_{t+1}^k \right]}{1 - \delta(1-d)\beta_t^*} \tag{82}$$

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 reads as

$$Q_t^L = \frac{\delta \mathbb{E}_t \left[ \frac{u_{t+1}^c}{u_t^c} F_{t+1}^k \right]}{1 - \tau_{t+1}^K \delta(1-d) - (1-\tau_{t+1}^K)\delta(1-d)\beta_t} \tag{83}$$

Then, the learning market inefficiency, 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 $X$:

$$\tau^* \iff X_t(\tau^*) = 0 \tag{84}$$

Then, the optimal tax is given by

$$\tau_{t+1}^* = 1 - \frac{\beta_t^* - 1}{\beta_t - 1} \tag{85}$$

Thus, the tax is a non-linear function of the deviation of capital gains subjective expectations from its objective value. Two limit cases can be derived. First, when subjective beliefs tend to the objective one 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[67]:

$$\tau_{t+1}^* = \begin{cases} > 0 \text{ if } \begin{cases} \beta_t > \beta_t^* \mid \beta_t > 1 & (A) \\ \beta_t < \beta_t^* \mid \beta_t < 1 & (B) \end{cases} \\ < 0 \text{ if } \begin{cases} \beta_t > \beta_t^* \mid \beta_t < 1 & (C) \\ \beta_t < \beta_t^* \mid \beta_t > 1 & (D) \end{cases} \end{cases} \tag{86}$$

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, $\beta_t^*$ is a

[67] 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 denominator.

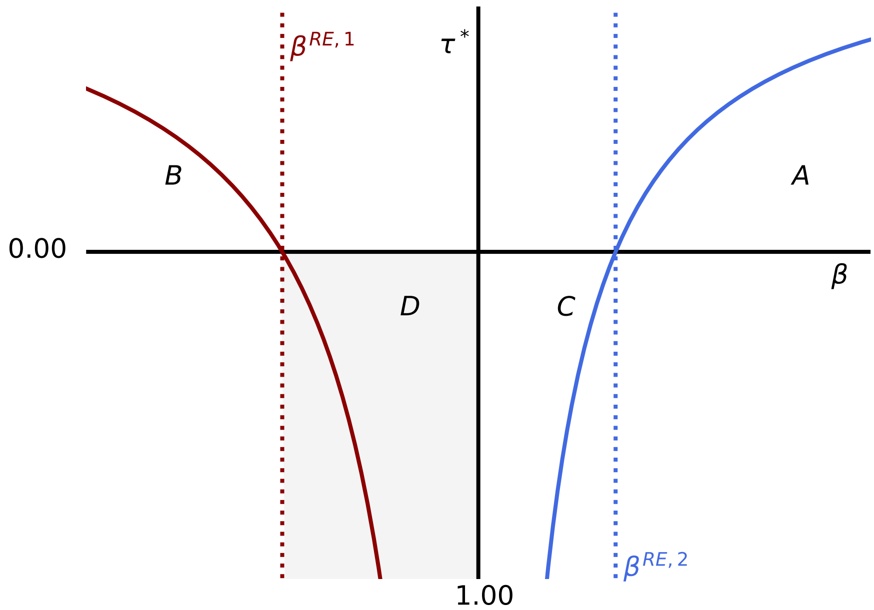


*Figure 10:* ***Optimal capital gains tax.*** *The figure shows the optimal capital gains tax $\tau^*$ as a function of subjective expectations $\beta_t$ for two different values of objective expectations $\beta^{RE}$ In the text, $\beta^{RE} = \beta_t^*$, change the graph., one above and one below 1. Letters signal the 4 four cases highlighted in expression (86). 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).*

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 $\beta_t$ 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. Thus, the optimal tax might take extreme values at times and be remarkable volatile. From a policy standpoint, that is probably an important shortcoming; next section deals with it.

### 4.5.- An alternative implementation

In this section, an alternative implementation of the optimal policy is presented. It relies on an alternative procedure: instead of solving a well defined optimization problem, it decomposes the volatility into fundamental and non-fundamental and finds the tax that eradicate the latter. It turns out it is a constant tax equal to 1. However, such a high tax depresses a bit the capital price inducing chronic under-investment. Such a bad outcome can be avoided by a subsidy on capital rents which depend on $\beta_t^*$, being pretty stable. Thus, a combination of two pretty stable can implement First Best allocations.

This implementation of optimal tax via volatility needs a decomposition of total volatility between fundamental and non-fundamental. Following the procedure in Section 2, the variance of the capital price can be approximated by

$$\mathbb{V}ar(Q_t) \approx \underbrace{z^2\mathbb{V}ar(Z_t) + k^2\mathbb{V}ar(K_{t-1})}_{\text{Fundamental}} \quad + \quad \underbrace{b^2\mathbb{V}ar(\beta_t)}_{\text{Non-Fundamental Volatility} \equiv \mathcal{V}} \tag{87}$$

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

$$\tau^* \iff \mathcal{V}(\tau^*) = 0 \tag{88}$$

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

$$b = \frac{\partial Q_t}{\partial \beta_t}(Z^*, K^*, 1) = \frac{\delta^2 F(Z^*, K^*)(1-d)(1-\tau^K_{t+1})}{(1-\delta(1-d)\tau^K_{t+1} - \delta(1-d)(1-\tau^K_{t+1}))^2} \tag{89}$$

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

$$Q^L_t = \frac{\delta \mathbb{E}_t\left[\frac{u^c_{t+1}}{u^c_t} F^k_{t+1}\right]}{1-\delta(1-d)} \tag{90}$$

which is exactly the price under Rational Expectations when $\beta^*_t = 1$. In other words, this approximate 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. Thus, an implementation based on the volatility decomposition could be seen as a limit case of the optimal policy problem.

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 $\tau^r$, which would resemble a corporate tax in this environment. With this new instrument, equation (91) becomes

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

Hence, by setting

$$(\tau_{t+1}^r)^* = 1 - \frac{1-\delta(1-d)}{1-\delta(1-d)\beta_t^*} \tag{92}$$

Thus, if RE implies an almost constant capital gains expectations, $(\tau_{t+1}^r)^*$ would be almost constant and then, the First Best can be implemented by a constant $\tau^K$ and a not-too volatile $\tau^r$ along with lump-sum taxes[68]. Altogether, the alternative implementation offers a way of stabilizing capital markets avoiding excessive volatility in taxes.

## 5.- Conclusions

In this paper, I study how capital gains taxes influence asset price cycles. Using a learning model, I showed that the tax level can stabilize the PD ratio by reducing its sensitivity to beliefs fluctuations which feeds back into more stable beliefs. The reason is that taxes affect both the discount factor and the beliefs formation process.

The theory has been used to tie together three changes observed in the US the last decades: the decline in capital taxes, the rise in the stock market valuations and its larger fluctuations. Indeed, the absence of tax cuts would have avoid the rise in volatility entirely despite the fall in safe real

[68]Note that $(\tau_{t+1}^r)^*$ would be less volatile than $\tau_{t+1}^*$ as long as $\beta_t^*$ is more stable than $\beta_t$.

interest rates.

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 key to explain stock market volatility, the excess volatility they cause can be seen as a pecuniary externality that might cause undesirable real fluctuations. In such case, a tax on unrealized capital gains that corrects too optimistic/pessimistic beliefs proves able to restore the First Best. On the wrong side, the optimal tax is unbounded and potentially very volatile which represent clear implementation problems. To overcome them, I suggest an alternative implementation that combines a high constant tax on capital gains and a smaller and pretty stable subsidy on capital rents.

Altogether, the arguments developed in the paper suggest that a tax on unrealized capital gains can be an effective tool to prevent asset price booms and the financial and macroeconomic fluctuations associated to them. In this sense, it emerges as an interesting macroprudential instrument, especially in front of the more popular but less clearly effective Financial Transactions Tax.

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 rather 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. Besides, an international analysis could help and see their possible interaction with other important changes as financial deregulation or capital flows liberalization.

Theoretically, the model has abstracted of a crucial element in asset price cycles, credit. An interesting question is whether a capital gains tax as the one suggested would be sufficient to prevent credit booms or whether a tax on borrowing is needed.[69] Besides, I have taken payout policies as given, ignoring their possible reaction to tax changes and its impacts on both asset prices and physical investment. Furthermore, the optimal capital taxation literature has not considered the use of capital gains taxes so far due to their focus on one-sector models.[70] Thus, the optimal use of capital gains taxes to fund government spending has not been explored. Finally, it is well known

[69] As the one suggested in some of the recent macroprudential literature, for instance Jeanne and Korinek (2019).

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

that capital gains have important redistributive implications.[71] The analysis in this paper would suggest that a capital gains tax could help in fighting that redistribution not only by taxing and redistributing capital gains but by preventing the part of them that comes from inefficient price movements.

[71] See Fagereng et al. (2002) for a recent analysis.

## References


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

Adam, K., Beutel, J., Marcet, A., and Merkel, S. (2015). Can a financial transaction tax prevent stock price booms? *Journal of monetary economics*, 76:S90–S109.

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.

Andrei, D., Mann, W., and Moyen, N. (2019). Why did the q theory of investment start working? *Journal of Financial Economics*, 133(2):251–272.

Aumann, R. J. (1976). Agreeing to disagree. *The annals of statistics*, 4(6):1236–1239.

Baker, S. R., Bloom, N., Davis, S. J., and Kost, K. J. (2019). Policy news and stock market volatility. Technical report, National Bureau of Economic Research.

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.

Bernanke, B. S. and Gertler, M. (2000). Monetary policy and asset price volatility.

Brennan, M. J. (1970). Taxes, market valuation and corporate financial policy. *National tax journal*, 23(4):417–427.

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

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 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.

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. (1992). Explaining the variance of price–dividend ratios. *The Review of Financial Studies*, 5(2):243–280.

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

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

Constantinides, G. M. (1983). Capital market equilibrium with personal tax. *Econometrica: Journal of the Econometric Society*, pages 611–636.

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.

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

Domar, E. D. and Musgrave, R. A. (1944). Proportional income taxation and risk-taking. *The Quarterly Journal of Economics*, 58(3):388–422.

Dybvig, P. H. and Zhang, H. (2018). That is not my dog: Why doesn't the log dividend-price ratio seem to predict future log returns or log dividend growths? *Available at SSRN 3123595.*

Fagereng, A., Gomez, M., Gouin-Bonenfant, E., Holm, M., Moll, B., and Natvik, G. (2002). Asset-price redistribution. *Manuscript.*

Fama, E. F. and French, K. R. (1998). Taxes, financing decisions, and firm value. *The journal of Finance*, 53(3):819–843.

Ferris, E. E. S. (2018). Dividend taxes and stock volatility. *International Tax and Public Finance*, 25(2):377–403.

García, L. et al. (2005). La vivienda y la reforma fiscal de 1998: un ejercicio de simulación. *Hacienda Pública Española/Revista de Economía Pública*, 175:123–147.

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.

Gomme, P., Ravikumar, B., and Rupert, P. (2011). The return to capital and the business cycle. *Review of Economic Dynamics*, 14(2):262–278.

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

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

Gutiérrez, G. and Philippon, T. (2016). Investment-less growth: An empirical investigation. Technical report, National Bureau of Economic Research.

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.

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

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

Klein, A., Hviid, S. J., Hvolbøl, T. S., Kramp, P. L., and Pedersen, E. H. (2016). House price bubbles and the advantages of stabilising housing taxation. *Danmarks Nationalbank Monetary Review*, 55(2):39–55.

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

LeRoy, S. F. and Porter, R. D. (1981). The present-value relation: Tests based on implied variance bounds. *Econometrica: Journal of the Econometric Society*, pages 555–574.

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.

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

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.

Melcangi, D. and Sterk, V. (2020). Stock market participation, inequality, and monetary policy. *FRB of New York Staff Report*, (932).

Miller, M. H. and Scholes, M. S. (1978). Dividends and taxes. *Journal of financial economics*, 6(4):333–364.

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

Peters, R. H. and Taylor, L. A. (2017). Intangible capital and the investment-q relation. *Journal of Financial Economics*, 123(2):251–272.

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).

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.

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.

*Table 4:* ***Baseline results.*** *This table reports the data and model moments, together with standard errors and t-statistics. In particular, the first two columns report the observed moments for the two periods. The third and fourth columns shows the learning model implied statistics which are the average across 1000 simulations. The last two columns reports the performance of the model under Rational Expectations. Both models use the parameterization described in Table 3. Numbers in parenthesis are Newey-West standard errors. Objects in square brackets are t-statistics, with the null of equality between observed and simulated moments. The t-ratio is the ratio of the gap between the model and the data moment over the Newey-West standard error of the moment in the data. Hence, values lower than 1.96 indicates that data and model moments are statistically equal. The top panel reports the targeted moments; the bottom panel the non-targeted ones. Data sources are in Appendix A.*

| | US data $\hat{\mathcal{S}}_i$ | | Learning Model $\widetilde{\mathcal{S}}_i(\hat{\theta})$ | | RE Model $\widetilde{\mathcal{S}}_i(\hat{\theta})$ | |
|---|---|---|---|---|---|---|
| | 1946-1982 | 1982-2018 | 1946-1982 | 1982-2018 | 1946-1982 | 1982-2018 |
| | | | Targeted moments | | | |
| $\mathbb{E}(PD_t)$ | 25.48 | 47.09 | 27.10 | 37.21 | 52.58 | 74.93 |
| | (1.55) | (4.04) | [-1.04] | [2.42] | [-17.47] | [-6.89] |
| $\mathbb{E}(D_t/D_{t-1})$ | 0.49 | 0.75 | 0.49 | 0.75 | 0.49 | 0.75 |
| | (0.35) | (0.34) | [-0.02] | [0.01] | [-0.02] | [0.01] |
| $\mathbb{E}(P_t/P_{t-1})$ | 0.48 | 1.84 | 0.77 | 1.22 | 0.64 | 0.86 |
| | (0.06) | (0.06) | [-0.86] | [0.86] | [-0.49] | [1.36] |
| $\mathbb{E}(r_t^s)$ | 4.73 | 4.33 | 4.81 | 4.46 | 2.67 | 2.28 |
| | (0.76) | (0.76) | [-0.08] | [-0.16] | [1.82] | [2.69] |
| $\mathbb{V}ar(p_t - d_t)$ | 7.15 | 13.98 | 5.85 | 15.23 | 2.63 | 3.25 |
| | (1.35) | (3.56) | [0.96] | [-0.35] | [3.33] | [3.01] |
| $\mathbb{C}ov(p_t - d_t, \bar{d}_t)$ | -1.98 | 2.37 | -1.97 | -0.01 | -0.24 | 0.18 |
| | (0.60) | (0.56) | [-0.02] | [4.24] | [-2.91] | [3.90] |
| $\mathbb{C}ov(p_t - d_t, \bar{r}_t)$ | -9.24 | -9.68 | -8.37 | -14.23 | -2.84 | -1.46 |
| | (1.51) | (3.04) | [-0.58] | [1.50] | [-4.23] | [-2.71] |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^K)$ | -0.29 | 1.73 | -0.42 | 0.68 | 0.00 | 1.41 |
| | (0.30) | (0.31) | [0.44] | [3.40] | [-1.01] | [1.02] |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^D)$ | 0.36 | 1.00 | 0.51 | 0.12 | 0.28 | 0.50 |
| | (0.33) | (0.33) | [-0.47] | [2.69] | [0.22] | [1.53] |
| | | | Non-targeted moments | | | |
| $\mathbb{E}(r_t^b)$ | 0.42 | 0.38 | 0.98 | 0.81 | 0.83 | 1.14 |
| | (0.02) | (0.03) | [28.55] | [12.84] | [21.31] | [22.88] |
| $\sigma(r_t^s)$ | 7.87 | 7.41 | 3.64 | 7.39 | 2.24 | 2.03 |
| | (0.72) | (0.81) | [-5.87] | [-0.02] | [-7.77] | [-6.63] |
| $corr(PD_t, PD_{t-1})$ | 0.96 | 0.98 | 0.97 | 0.96 | 0.97 | 0.98 |
| | (0.13) | (0.07) | [0.11] | [-0.32] | [-0.38] | [0.06] |

*Table 5: **Results targeting the risk-free interest rate.** This table reports the data and model moments, together with standard errors and t-statistics. In particular, the first two columns report the observed moments for the two periods. The third and fourth columns shows the learning model implied statistics which are the average across 1000 simulations. The last two columns reports the performance of the model under Rational Expectations. Both models use the parameterization described in the column "With $r_t^b$" in panel b) of Table 3. Numbers in parenthesis are Newey-West standard errors. Objects in square brackets are t-statistics, with the null of equality between observed and simulated moments. The t-ratio is the ratio of the gap between the model and the data moment over the Newey-West standard error of the moment in the data. Hence, values lower than 1.96 indicates that data and model moments are statistically equal. The top panel reports the targeted moments; the bottom panel the non-targeted ones. Data sources are in Appendix A.*

| | US data $\hat{\mathcal{S}}_i$ | | Learning Model $\widetilde{\mathcal{S}}_i(\hat{\theta})$ | | RE Model $\widetilde{\mathcal{S}}_i(\hat{\theta})$ | |
|---|---|---|---|---|---|---|
| | 1946-1982 | 1982-2018 | 1946-1982 | 1982-2018 | 1946-1982 | 1982-2018 |
| | | | Targeted moments | | | |
| $\mathbb{E}(PD_t)$ | 25.48 | 47.09 | 31.38 | 42.70 | 60.92 | 85.42 |
| | (1.55) | (4.04) | [-3.80] | [1.09] | [-22.86] | [-9.49] |
| $\mathbb{E}(D_t/D_{t-1})$ | 0.49 | 0.75 | 0.49 | 0.75 | 0.49 | 0.75 |
| | (0.35) | (0.34) | [-0.02] | [0.01] | [-0.02] | [0.01] |
| $\mathbb{E}(P_t/P_{t-1})$ | 0.48 | 1.84 | 0.72 | 1.26 | 0.69 | 0.85 |
| | (0.06) | (0.06) | [-0.38] | [0.92] | [-0.35] | [1.57] |
| $\mathbb{E}(r_t^s)$ | 4.73 | 4.33 | 4.19 | 4.13 | 2.44 | 2.09 |
| | (0.76) | (0.76) | [0.71] | [0.27] | [3.01] | [2.96] |
| $\mathbb{E}(r_t^b)$ | 0.42 | 0.38 | 0.42 | 0.39 | 0.37 | 0.53 |
| | (0.02) | (0.03) | [0.00] | [-0.10] | [2.48] | [-4.53] |
| $\mathbb{V}ar(p_t - d_t)$ | 7.15 | 13.98 | 5.48 | 16.16 | 2.49 | 3.16 |
| | (1.35) | (3.56) | [1.23] | [-0.61] | [3.44] | [3.04] |
| $\mathbb{C}ov(p_t - d_t, \bar{d}_t)$ | -1.98 | 2.37 | -1.60 | -0.00 | 0.03 | 0.19 |
| | (0.60) | (0.56) | [-0.63] | [4.23] | [-3.36] | [3.88] |
| $\mathbb{C}ov(p_t - d_t, \bar{r}_t)$ | -9.24 | -9.68 | -7.64 | -15.18 | -1.18 | -1.31 |
| | (1.51) | (3.04) | [-1.05] | [1.81] | [-5.33] | [-2.76] |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^K)$ | -0.29 | 1.73 | -0.51 | 0.68 | 0.96 | 1.47 |
| | (0.30) | (0.31) | [0.73] | [3.40] | [-4.25] | [0.85] |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^D)$ | 0.36 | 1.00 | 0.47 | 0.10 | 0.26 | 0.49 |
| | (0.33) | (0.33) | [-0.36] | [2.76] | [0.28] | [1.57] |
| | | | Non-targeted moments | | | |
| $\sigma(r_t^s)$ | 7.87 | 7.41 | 3.51 | 7.78 | 2.07 | 2.01 |
| | (0.72) | (0.81) | [-6.01] | [0.46] | [-8.02] | [-6.66] |
| $corr(PD_t, PD_{t-1})$ | 0.96 | 0.98 | 0.97 | 0.95 | 0.97 | 0.98 |
| | (0.13) | (0.07) | [0.10] | [-0.38] | [0.06] | [0.05] |

*Table 6:* ***Capital taxes marginal effect.*** *This table reports the baseline estimation results from the learning model and different counterfactuals. The third and fourth column shows the model implication in the case that both dividends and capital gains taxes remained at its 1946 level. The next (last) two keep the capital gains (dividends) tax constant.*

| | Learning Model | | Constant $\tau^K$, $\tau^D$ | | Constant $\tau^K$ | | Constant $\tau^D$ | | Constant $a$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1946-82 | 82-2018 | 1946-82 | 82-2018 | 1946-82 | 82-2018 | 1946-82 | 82-2018 | 1946-82 | 82-2018 |
| $\mathbb{E}(PD_t)$ | 27.10 | 37.21 | 24.42 | 26.01 | 26.07 | 31.52 | 25.14 | 30.67 | 27.08 | 34.46 |
| $\mathbb{E}(D_t/D_{t-1})$ | 0.49 | 0.75 | 0.49 | 0.75 | 0.49 | 0.75 | 0.49 | 0.75 | 0.49 | 0.49 |
| $\mathbb{E}(P_t/P_{t-1})$ | 0.77 | 1.22 | 0.48 | 0.79 | 0.59 | 0.83 | 0.62 | 1.14 | 0.74 | 0.67 |
| $\mathbb{E}(r_t^s)$ | 4.81 | 4.46 | 4.85 | 4.82 | 4.67 | 4.11 | 4.96 | 5.07 | 4.71 | 3.92 |
| $\mathbb{V}ar(p_t - d_t)$ | 5.85 | 15.23 | 3.21 | 1.77 | 3.04 | 1.14 | 5.25 | 13.78 | 4.72 | 8.54 |
| $\mathbb{C}ov(p_t - d_t, \bar{d}_t)$ | -1.97 | -0.01 | 0.16 | 0.11 | -0.06 | 0.04 | -0.27 | -0.03 | -0.50 | -0.03 |
| $\mathbb{C}ov(p_t - d_t, \bar{r}_t)$ | -8.37 | -14.23 | -3.05 | -1.64 | -2.98 | -1.03 | -5.28 | -13.02 | -4.63 | -7.66 |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^K)$ | -0.42 | 0.68 | 0.00 | 0.00 | 0.00 | 0.00 | 0.20 | 0.57 | 0.23 | 0.69 |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^D)$ | 0.51 | 0.12 | 0.00 | 0.00 | 0.24 | 0.11 | 0.00 | 0.00 | 0.51 | 0.19 |
| $\mathbb{E}(r_t^b)$ | 0.98 | 0.81 | 0.90 | 1.23 | 1.01 | 1.17 | 0.83 | 0.88 | 0.95 | 0.72 |

*Table 7:* ***Decomposition of the stock return geometric mean.*** *The table shows the stock returns mean decomposition according to expression (47). The first column uses U.S. data; the second, simulated data using the learning model; the third, simulated data using the RE model. 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 |
| $R_1$ | 1.0046 | 1.0074 | 1.0049 | 1.0075 | 1.0049 | 1.0075 |
| $R_2$ | 0.9973 | 1.0088 | 1.0028 | 1.0013 | 1.0026 | 1.0013 |
| $R_3$ | 1.0410 | 1.0237 | 1.0363 | 1.0317 | 1.0178 | 1.0123 |

*Table 8:* ***Robustness analysis.*** *This table reports the data and model moments, together with standard errors and t-statistics. In particular, the first two columns report the observed moments for the two periods. The third and fourth columns shows the learning model implied statistics for the case of tax rebates. The next column reports results for the case of anticipated tax changes. The last ones for the case of tax learning. The models use the baseline calibration reported in Table 3; the rest of the parameters have been reestimated for each model and are reported in Appendix XX. Numbers in parenthesis are Newey-West standard errors. Objects in square brackets are t-statistics, with the null of equality between observed and simulated moments. The t-ratio is the ratio of the gap between the model and the data moment over the Newey-West standard error of the moment in the data. Hence, values lower than 1.96 indicates that data and model moments are statistically equal. The top panel reports the targeted moments; the bottom panel the non-targeted ones. Data sources are in Appendix A.*

| | US data $\hat{\mathcal{S}}_i$ | | Tax Rebates $\widetilde{\mathcal{S}}_i(\hat{\theta})$ | | Tax anticipation $\widetilde{\mathcal{S}}_i(\hat{\theta})$ | | Tax learning $\widetilde{\mathcal{S}}_i(\hat{\theta})$ | |
|---|---|---|---|---|---|---|---|---|
| | 1946-1982 | 1982-2018 | 1946-1982 | 1982-2018 | 1946-1982 | 1982-2018 | 1946-1982 | 1982-2018 |
| Targeted moments | | | | | | | | |
| $\mathbb{E}(PD_t)$ | 25.48 | 47.09 | 35.16 | 41.48 | | | | |
| | (1.55) | (4.04) | [-6.24] | [1.39] | | | | |
| $\mathbb{E}(D_t/D_{t-1})$ | 0.49 | 0.75 | 0.49 | 0.75 | 0.49 | 0.75 | | |
| | (0.35) | (0.34) | [-0.02] | [0.01] | [-0.02] | [0.01] | | |
| $\mathbb{E}(P_t/P_{t-1})$ | 0.48 | 1.84 | 0.66 | 1.35 | | | | |
| | (0.06) | (0.06) | [-0.28] | [0.78] | | | | |
| $\mathbb{E}(r_t^s)$ | 4.73 | 4.33 | 3.65 | 4.14 | | | | |
| | (0.76) | (0.76) | [1.41] | [0.25] | | | | |
| $\mathbb{V}ar(p_t - d_t)$ | 7.15 | 13.98 | 3.18 | 10.77 | | | | |
| | (1.35) | (3.56) | [2.94] | [0.90] | | | | |
| $\mathbb{C}ov(p_t - d_t, \bar{d}_t)$ | -1.98 | 2.37 | -2.15 | 0.04 | | | | |
| | (0.60) | (0.56) | [0.28] | [4.14] | | | | |
| $\mathbb{C}ov(p_t - d_t, \bar{r}_t)$ | -9.24 | -9.68 | -6.32 | -9.99 | | | | |
| | (1.51) | (3.04) | [-1.93] | [0.10] | | | | |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^K)$ | -0.29 | 1.73 | -0.59 | 0.44 | | | | |
| | (0.30) | (0.31) | [0.99] | [4.18] | | | | |
| $\mathbb{C}ov(p_t - d_t, \bar{\tau}_t^D)$ | 0.36 | 1.00 | 0.22 | 0.06 | | | | |
| | (0.33) | (0.33) | [0.41] | [2.87] | | | | |
| Non-targeted moments | | | | | | | | |
| $\mathbb{E}(r_t^b)$ | 0.42 | 0.38 | 0.61 | 0.48 | | | | |
| | (0.02) | (0.03) | [9.83] | [3.07] | | | | |
| $\sigma(r_t^s)$ | 7.87 | 7.41 | 3.58 | 8.25 | | | | |
| | (0.72) | (0.81) | [-5.92] | [1.04] | | | | |
| $corr(PD_t, PD_{t-1})$ | 0.96 | 0.98 | 0.95 | 0.89 | | | | |
| | (0.13) | (0.07) | [-0.07] | [-1.25] | | | | |

*Table 9:* ***Counterfactual historical Price-Dividend ratio in different scenarios.*** *The table reports the mean and standard deviation of the PD ratio using observed US data and simulated data from the calibrated historical version of the model.*

| | US data | | No $\tau^K$ cuts | | No $\tau^D$ cuts | | No $\tau^B$ cuts | |
|---|---|---|---|---|---|---|---|---|
| | $\mathbb{E}(PD)$ | $\sigma(PD)$ | $\mathbb{E}(PD)$ | $\sigma(PD)$ | $\mathbb{E}(PD)$ | $\sigma(PD)$ | $\mathbb{E}(PD)$ | $\sigma(PD)$ |
| 1946-1982 | 25.48 | 6.52 | 26.02 | 7.90 | 23.32 | 6.03 | 25.55 | 6.60 |
| 1982-2018 | 47.02 | 16.52 | 32.89 | 8.17 | 36.16 | 13.17 | 48.14 | 17.32 |
| Ratio | 1.85 | 2.53 | 1.26 | 1.03 | 1.55 | 2.18 | 1.88 | 2.63 |
| | | | No $r^b$ decline | | No buybacks | | No W/D decline | |
| 1946-1982 | | | 25.42 | 6.54 | 49.86 | 8.01 | 25.42 | 6.54 |
| 1982-2018 | | | 43.90 | 15.18 | 52.23 | 13.63 | 73.68 | 33.60 |
| Ratio | | | 1.73 | 2.32 | 1.05 | 1.70 | 2.90 | 5.14 |