Capital Taxation and Stock Market Volatility

Pau Belda[†]

November 2020

Preliminary version.

Abstract

This paper argues that the long run volatility of the price-dividend ratio is decisively driven by capital taxes. First, a new empirical fact is documented: capital taxes alone are able to explain up to 85% of the permanent increase in the long-run volatility observed since the mid 1990s in the United States. Second, the role of stochastic capital taxes is analyzed in a consumption-based asset pricing model. In a standard Rational Expectations setup with stochastic taxes, tax volatility capitalizes into price volatility since taxes affect investor's net cash-flows. However, this fundamental channel only gives rise to a transitory volatility response. Alternatively, a model in which rational investors face uncertainty about how fundamentals map into prices and learn about it brings up a new non-fundamental channel: taxes determine the elasticity of stock prices to investors beliefs. In this case, if a tax rate change is persistent, this second channel produces a persistent rise in volatility as the one observed. Third, different versions of the model are structurally estimated. The learning model quantitatively replicates a list of statistics made of traditional asset pricing facts as well as the new tax-volatility link. Remarkably, a high enough equity premium (along with a low and stable risk-free rate) is obtained with low risk aversion levels. Contrarily, the RE counterpart performs poorly in all the relevant dimensions.

Keywords: Capital Taxation, Stock Market Volatility, Internal Rationality, Equity Premium.

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

[†]PhD Candidate at Universitat Autònoma de Barcelona (UAB) and Barcelona Graduate School of Economics (BGSE). Contact: pau.beldaitortosa@gmail.com. This is a preliminary version; do not distribute it. I thank Albert Marcet for his guidance. Besides, I have benefited from comments from Klaus Adam, Dirk Krueger, Luís Rojas, Hugo Rodríguez, Alexander Ludwig, Abhay Abhyankar, Ramon Ruiz and several participants in seminars at Barcelona GSE Jamboree 2019 and 2020, Mannheim PhD Seminar 2020, IDEA students seminar 2020 and the Bellaterra MacroClub 2020. This research has received funding from the European Research Council under the European Union's Horizon 2020 linked to the project "Asset Prices and Macro Policy when Agents Learn and are Heterogeneous" (PI: Albert Marcet). The usual disclaimers apply.

Stock market volatility is back. The fluctuations of the price-dividend ratio have become larger and more frequent since the 1980s, linked to turbulent events as the late 1980s crashes, the Dot-com bubble or the Great Recession. Although shorter in duration, these episodes resembled the turbulent times surrounding the Great Depression and, in any case, represented a clear ending of the relatively calmed period of the central decades of the 20th century[1]. This paper argues that the return of volatility is intimately related to the sharp decline in capital income tax rates.

There is now substantial evidence of the fact that capital taxes influence the time series of stock prices. Particularly, the fall in effective capital tax rates has been regarded as a crucial determinant of the rise in aggregate market valuations that took place in the United States since the 1980s (e.g., McGrattan and Prescott (2005)). On top of reinforcing this tax capitalization hypothesis, I claim that the decline in capital taxes was also responsible for the increase in the aggregate stock market volatility occurring roughly over the same period. Thus, the paper digs into the causes of stock prices excess volatility, bringing in a macroeconomic fundamental which has been largely overlooked in most of the previous research. The contributions of the paper to the literature are explored in depth in Section 1.

In the paper, a new empirical fact is documented: there is a strong negative correlation between capital taxes and the long run volatility of the Price-Dividend ratio in the United States along the 1954-2012 period. Thus, taxes alone explain up to 85% of the long run volatility variation. The good prediction of the volatility time trajectory is due to a level effect: the lower the initial tax rate, the larger the impact of a tax cut on volatility. This non-linearity hints the relevance of the tax rate level and not just of tax changes in determining volatility. The empirical analysis is found in Section 2.

The strong negative correlation is taken as a signal of a causal connection running from capital taxes to low frequency stock market fluctuations. Then, causality is analyzed in a general equilibrium model in which taxes drive volatility through two channels. On the one hand, tax volatility directly capitalizes into price volatility because taxes determine the net cash flow accruing to investors. In this way, lower taxes would increase after-tax returns which -because of a substitution effect- would boost stock demand and -given a short run inelastic stock supply- prices. In other

[1]The increase in volatility since the 1980s is also observed for daily series of stock returns and percentage price changes. See Appendix A

words, prices reflect changes in fundamentals. This direct effect is a logical corollary of the tax capitalization hypothesis and was already explored in a general equilibrium framework by Sialm (2006). Nevertheless, this channel shows unable to produce a permanent increase in volatility because tax changes eventually cease to happen and then prices would stabilize.

On the other hand, taxes play a key role in beliefs-driven booms and busts. This kind of fluctuations have a long tradition in finance (e.g., Keynes (1936), Kindleberger (1978) or Shiller (2000)) and recently have been proven quantitatively powerful (e.g., Adam et al. (2017)). Their core mechanism is a a self-referential loop: optimistic price expectations increase demand and actual prices, and the latter feed back into investors' expectations. In this context, there emerges a new channel through which taxes rule price-dividend volatility: the tax rate level acts as a wedge between investors' expectations and prices. As a result, the higher the taxes, the lower the elasticity of stock prices to capital gains expectations. In other words, if optimism boosts up expected capital gains, taxes absorb part of this rise, making the after-tax expected return lower and then prices less responsive to optimism. It follows that the higher the capital taxes, the less likely the booms and then, the faster beliefs tend to revert towards fundamentals. Altogether, the tax rate level determine the exposure of prices to non-fundamental fluctuations. If the tax rate level changes permanently, this second channel is able to produce a durable rise in volatility as the one documented in Section 2. This story is formalized via a consumption based general equilibrium asset pricing model with rational investors who have imperfect knowledge about how prices are formed but learn about it from what they observed as in Adam et al. (2016), augmented with stochastic capital taxes on dividends and realized capital gains. It is exposed in Section 3.

The way of assessing the rejectability of the causal hypothesis is to confront the model to a list of asset pricing statistics from US data for the 1954-2012 period. The list consists of a set of traditional measures of the stock market behaviour, including features of the dividends process, the equity premium, the volatility of the market and the predictability of stock returns. This list is augmented by the new statistics brought up in this paper as the mean and dispersion of the long run volatility and the correlation between taxes and the long run volatility. Finally, it also includes the correlation between the tax level and market valuation (in line with McGrattan and Prescott (2005) or Sialm (2009)). In other words, the model is asked to replicate the new facts documented in Section 2 together with many of the facts traditionally highlighted in the literature. It is estimated through the Simulated Method of Moments, having 7 free parameters to match 13 statistics.

It turns out that all the model generated statistics are able to pass individual significance tests, many of them with t-statistics below 1. As opposed to that, its Rational Expectations counterpart misses not only the volatility (as systematically pointed out in the literature) but also the association between taxes and the long run volatility. The quantitative analysis is displayed in Section 4.

Remarkably, the model is able to generate a high enough historical equity premium with a low and stable risk-free rate, low risk aversion, positive subjective discount factor, high enough market volatility and realistic dividends and consumption growth processes. In this sense, it raises a solution to the (historical) equity premium puzzle (as stated in Cochrane (2017)). It does so because the decline in taxes gives rise to an upward trend in the price-dividend ratio that naturally increases the sample stock returns mean but barely affects the risk-free rate. In other words, whereas usually consumption-based models generate trendless price-dividends and have to resort to exorbitant risk-aversion levels in one way or another, the upward PD ratio trend emerging from the decline in taxes that matches the observed (trended) PD ratio does the job with realistically low risk aversion.

The model also passes an additional number of tests. First, the results are robust to alternative information set specifications (for instance, assuming that agents fully anticipate tax changes in the near future rather than be surprised by them) and to alternative learning schemes. Second, the model performs well in a historical test consisting on replacing the Montecarlo simulation by the observed realization of dividends and consumption shocks. Finally, Section 5 explores a distinction between dividends and capital gains tax rates, proving that the latter are more powerful to regulate the overall level of volatility as long as agents learn about prices and irregardless if they also learn about dividends.

Altogether, the paper shows the crucial role of capital taxes regarding not only asset valuations but also aggregate volatility in the stock market. In other words, it raises a tax theory of the aggregate long run financial volatility.

1.- Contributions to the literature

The connection between stock market volatility and capital taxes brought up by this paper lies at the intersection of two major debates on finance and macroeconomics: the causes of stock market volatility on the one hand, and the relation between capital taxes and stock prices on the other hand. With few exceptions, the former lacks any consideration of taxes and the latter of volatility. Therefore, the contribution of this paper is to bring these topics together by providing both empirical evidence and explicit causal mechanisms in a general equilibrium framework.

There are two main bodies of literature analyzing stock market volatility. On the one hand, the excess volatility literature; on the other hand, a plethora of time series models for forecasting returns 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, it has abundantly resorted to variance decompositions of the price-dividend ratio (see Campbell and Shiller (1988) or Cochrane (1992) for early versions), showing that, as a matter of an identity, the volatility of the price-dividend ratio was almost completely explained by future expected returns -or discount rates- variation rather than expected market fundamentals (dividends or even earnings growth) (see Cochrane (2009)). Since the covariance between price-dividends and future expected returns is high, it must be the case that a regression of the latter on the former must show considerable predictive power[2]. Thus, the empirical excess volatility literature has turned into an allegedly equivalent discussion about returns predictability (as claimed by Cochrane (2017)). As opposed to that, in this paper I stick to the original emphasis of the price-dividend variation, focusing on its low frequency movements and linking it to one overlooked fundamental: capital taxes.

Several general equilibrium models have tried to incorporate different mechanisms to generate a large enough volatility[3,4]. Following a long tradition in finance (e.g., Kindleberger (1978), Minsky (1986)), this paper claims that volatility is crucially caused by investors' subjective per-

[2] On top of dividend yields, a number of variables (Price-Earnings ratios, bond yields, the aggregate consumption-wealth ratio, etc.) have revealed significant predictive power. However, predictability is still highly controversial given the existence of contradictory evidence in favour of the predictability of excess returns in the short rather than in the long run Ang and Bekaert (2007) or in favour of the forecastability of dividend growth rather than returns Dybvig and Zhang (2018), when alternative corrections are made for both serial correlation and heteroscedasticity.

[3] Cochrane (2017) reviewed most of them -always within the consumption-based class- pointing out how they did the job. The common strategy among them consisted in increasing the variability of the stochastic discount factor by bringing in some additional variable on behalf of generalizing the standard consumption-based model. The number of stories is magnificent, ranging from consumption habits or long run risks to agents heterogeneity or directly agents' irrationality.

[4] Out of the consumption-based class of asset pricing models, it is worth it to highlight the work of Iraola and Santos (2017), who explained up to 2/3 of price-dividend long run volatility via price markups volatility using a growth model.

ceptions. This idea is formalized through the Internal Rationality approach put forward by Adam and Marcet (2011), which slightly deviates from Rational Expectations by generalizing the pricing function. Thus, the fact that agents have imperfect knowledge about how fundamentals maps into prices and are extrapolative regarding their capital gains beliefs gives rise to self-fulfilling booms and busts in line with the empirical evidence (Adam et al. (2016) and Adam et al. (2017)). My paper shows that bringing in capital taxes into the Adam et al. (2016) model improves its performance in the dimensions they considered (particularly, in terms of the equity premium), and additionally makes it able to account for the additional facts brought up in this paper, without making any additional core assumption.

The second body of literature tackling volatility consists of a vast number of time series models organized around GARCH (starting with Engle (1982)) and Stochastic Volatility models (first introduced by Taylor (1986))[5]. Within this literature, there is a subset of works attempting to find the macroeconomic roots of stock market volatility, starting with Officer (1973) and Schwert (1989)[6]. This macroeconomic approach to financial volatility has highlighted the counter-cyclical nature of aggregate volatility and the predictive power of several macroeconomic variables (such as inflation, production or money aggregates)[7]. Nevertheless, a general macroeconomic theory of the returns volatility is not in place yet (Engle and Rangel (2008)). My paper contributes to the macrofoundations of aggregate volatility by presenting both reduced-form time series evidence as well as a proper causal analysis in a structural model, posting capital taxes as a primary candidate to explain the very long run volatility[8].

Despite its underestimation in the volatility literature, taxes have explicitly been claimed a determinant of asset prices. Two competing views have co-existed here: the tax irrelevance (starting with Miller and Scholes (1978)) and the tax capitalization hypothesis (starting with Brennan (1970); see Sialm (2009) for a complete review). Although there is no empirical consensus when looking at the returns cross-section[9], the tax capitalization has generated considerable agreement

[5]Their main goal has been to forecast stock returns volatility while accounting for key return stylized facts (such as volatility clustering, fat tails or gain/loss asymmetry; see Cont (2001) for a stylized list of stock returns facts).

[6]Some more recent examples are Errunza and Hogan (1998), Beltratti and Morana (2006), Diebold and Yilmaz (2008), Campbell and Diebold (2009), Corradi et al. (2012), Asgharian et al. (2013) or Conrad and Loch (2015).

[7]The evidence in favour of macroeconomic causes of aggregate volatility is large, but not unanimous. For instance, Christiansen et al. (2012) showed that pure macroeconomic variables had poor predictive power when compared to financial ones (valuation ratios, credit risk, interest rates, etc.).

[8]In this paper I focus on the volatility of the market price-dividend ratio, but the analysis is extensible to stock returns.

[9]For instance, Fama and French (1998) did not find reliable evidence of tax effects using a sample of firms from

at aggregate valuations time series domain. Thus, Sialm (2009) via reduced-form evidence and McGrattan and Prescott (2005) using a quantitative neoclassical growth model, showed that the fall in the effective dividend tax rate was behind the rising trend in valuations. Recently, this hypothesis has been also confirmed by Brun and González (2017). On top of reinforcing the tax capitalization hypothesis, this paper extends the relevance of capital taxes to prices second moment.

Besides, Sialm (2006) presented a version of the Lucas (1978) model with stochastic taxes. He pointed out that taxes affect asset prices via substitution, income and growth effects obtaining, however, some undesirable results. On the one hand, only tax changes are relevant for the price-dividend ratio level. In other words, constant taxes play no role in determining the asset price level, so that, a 100% or a 0% rate would deliver the same asset price as long as it remains constant[10]. On the other hand, he got a positive relation between valuations and taxes under plausible assumptions[11], 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[12,13]. The model presented in my paper also analyzes stochastic taxation in a Lucas economy, but incorporating a very different set of assumptions. First, I use a more realistic tax system. On the one hand, I consider personal capital income taxes, levying on both dividends and realized capital gains, as is the case in the US tax code. By doing so, the price-dividend ratio becomes an unambiguously negative function of the tax rate level. On the other hand, I use an AR(1) process for taxes instead of a two-state Markov process, which is able to replicate the observed US time series. Second, the Rational Expectations assumption is replaced by a more general information setup, giving rise to a new channel through which the tax rate level influences price volatility, which turns out to be crucial for the good quantitative performance of the model.

Finally, there exists some works that raised a connection between taxes and volatility, in a way

1965 to 1992, whereas Sialm (2009) did so when considering a larger period (1913-2006).

[10]This is the case in a constant relative risk aversion environment. He only obtained a non-neutral tax rate level with respect to stock prices under constant absolute risk aversion.

[11]The tax-price positive relationship comes up whenever dividend and tax growth processes are dependent or when they are orthogonal and the constant risk aversion coefficient is low and a part of the tax revenues are rebated to investors.

[12]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 τ_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 ($(1-\tau_t)d_t s_{t-1}$).

[13]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 generalization of the Lucas model by bringing in capital taxes.

or another. In this sense, Baker et al. (2019) created a newspaper-based Equity Market Volatility tracker in which news about taxes (not taxes themselves) 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) reported that an RBC model with stochastic taxes was able to generate about the 80% of the capital return volatility. The thing is that the rate of return they built was a ratio of dividends over physical capital which exhibited a volatility 22 times lower than the volatility of SP500 returns. In this sense, their model did not address the stock market volatility issue. Finally, Dai et al. (2013) showed that US 1978 and 1997 capital gains tax cuts gave rise to an increase in return volatility due to the reduction in risk-sharing between government and stockholders. As opposed to that, this paper is the first to explore the relationship between capital taxes (concerning both dividend and capital gains) and stock market volatility systematically, providing both reduced-form time series evidence and explicit mechanisms in a general equilibrium model.

2.- A new fact

In this section, a new fact is documented: the strong negative correlation between the long run price-dividend volatility and the capital tax rate for the United States over the 1954-2012 period. On the one hand, both variables are operationalized: the long run volatility is the permanent component of the conditional variance and the capital tax rate is a convex combination of the effective average marginal rates on dividends, short and long capital gains. The two objects display a strong negative correlation equal to -0.84. On the other hand, this association is further explored by showing the ability of tax rates alone to predict up to 85% of the volatility time trajectory.

2.1. Empirical Measurement

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

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

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

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

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

This is equivalent to write the price-dividend innovation $\epsilon_{q,t}$ as a product of the conditional variance and a white noise shock as

$$\epsilon_{q,t} = \sigma_{q,t}\varepsilon_{q,t}, \quad \varepsilon_{q,t}|I_{q-1,t} \sim \mathcal{N}(0,1) \tag{2}$$

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

$$\sigma_{q,t} = \mathrm{v}_t g_{q,t} \tag{3}$$

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

$$\mathrm{g}^2_{q,t} = \alpha_0 + \alpha_1\frac{\epsilon^2_{q,t}}{\mathrm{v}^2_t} + \alpha_2\mathrm{g}^2_{q-1,t} \tag{4}$$

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

$$\mathrm{v}_t^2 = \phi_0 + \phi_1 \sum_{k=1}^{K} \varphi_k(w) RV_{t-k}^2 \tag{5}$$

with the realized volatility defined as a moving standard deviation over a fixed window[14] of Q quarters

$$RV_t = \sqrt{\frac{1}{Q-1} \sum_{q=1}^{Q} \Big(\frac{P_{q,t}}{D_{q,t}} - \frac{1}{Q} \sum_{q=1}^{Q} \frac{P_{q,t}}{D_{q,t}} \Big)^2} \tag{6}$$

and the weighting scheme given by a beta lag structure, which yields a monotonically decreasing sequence determined by a single parameter[15]

$$\varphi_k(w) = \frac{(1-k/K)^{w-1}}{\sum_{j=1}^{K}(1-j/K)^{w-1}} \tag{7}$$

Altogether, the parameter vector θ contains a total of 7 parameters $\theta = \{\mu, \rho, \alpha_0, \alpha_1, \phi_0, \phi_1, w\}$, jointly estimated for $(Q, K) = (4, 10)$ through Quasi-Maximum Likelihood[16]. The long run coefficients $\{\phi_0, \phi_1\}$ turns out to be significant at usual levels of confidence. Appendix B reports the baseline estimation results as well as a number of alternative specifications of the long run volatility that essentially display the same time trajectory.

As showed in the right bottom panel of figure 1, the long run volatility was far from constant throughout the 1954-2012 period. In general, it experienced ups and downs, alternating calmed with turbulent times, in consonance with the well-known fact of volatility clustering (as first noted by Mandelbrot (1963)). Nevertheless, the novel fact that this measure brings up is the steady rise in long run volatility since the 1990s. Thus, there was a transition from a relatively low and stable regime up to the 1980s to a highly volatile environment afterwards. As a result, the volatility ended up at levels almost 30% higher than at the beginning of the sample period. This shift appears as very persistent, with no signals of mean-reverting behavior. Interestingly, the rise in volatility mirrored the increasing trend in the PD ratio level, but by no means the former is a me-

[14] Results are robust to consider a rolling window. However, I use a fixed window to get a volatility time series at the same frequency of the tax time series (i.e., annual). See Appendix B.

[15] Engle et al. (2013) also propose alternative schemes: a beta polynomial with two parameters and an exponential weighting. Results are robust to different weighting schemes. See appendix B.

[16] It is well known that the QML estimator is consistent and asymptotically normal for GARCH(1,1), provided that the innovation distribution has a finite fourth moment, even if the true distribution is far from Gaussian (e.g., see Lumsdaine (1996)). This is the case here indeed: residuals are non-Gaussian (due to fat tails) but exhibit an empirical kurtosis of 6.16 so that quasi-maximum likelihood estimators are asymptotically Normally distributed.

chanic consequence of the latter (higher PD levels are perfectly compatible with lower fluctuations).

The rise in financial volatility[17] strongly contrasts with the regime switch to lower volatility of real macroeconomic indicators occurring in the last 15 years of the 20th century, which has been well documented in the literature (see, e.g., Blanchard and Simon (2001) or Stock and Watson (2002)) and popularly known as the Great Moderation (Bernanke (2004)). In other words, the B-side of the Great Moderation in real activity seems to be a sort of "Great Agitation" in capital markets[18].

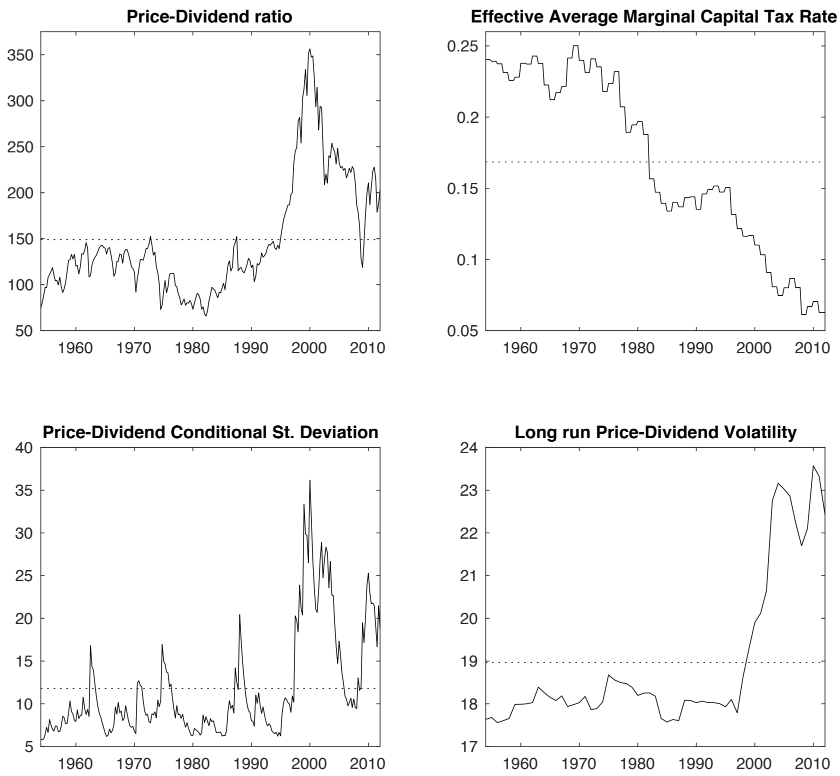

Figure 1: ***Time series of the main variables.*** *The top left panel pictures the Price-Dividend ratio for the S&P 500, with deseasonalized dividends. The top right panel shows the effective average marginal capital tax rate, defined in equation (8). The bottom panels depict the PD ratio conditional standard deviation and the long run volatility, computed according equation (3) and (5). The dotted lines pictures their long run sample averages. Computations used data from 1927:I to 2012:I for the United States. Graphs plot the 1954:I to 2012:I subsample. See Appendix A for data sources.*

As a measure for the tax rates levying equity income, an effective average marginal capital income tax rate is proposed. It is a convex combination of dividend, short and long capital gains

[17] The increase in volatility is also observed for both stock returns and percentage price changes. See appendix B.
[18] The possible connection between this two facts is left to future research.

tax rates given by

$$\tau_t = \Big(\tau_t^d \nu_1 + \tau_t^{skg} \nu_2 + \tau_t^{lkg}(1 - \nu_1 - \nu_2)\Big)(1 - \eta_t) \tag{8}$$

where τ_t^d, τ_t^{skg} and τ_t^{lkg} are the average effective marginal tax rates on dividends, short and long capital gains respectively. These are the rates implemented to an additional unit of capital income, which are the relevant ones to analyze the effects of taxes on economic decisions (see Barro and Sahasakul (1983) or McGrattan and Prescott (2005)). The TAXSIM program of the National Bureau of Economic Research (NBER) supplies these rates on an annual basis from 1960 to 2018 at federal level and from 1979 to 2008 at state level. Then, I adjusted for state and local taxes before 1979 and after 2008 (following Sialm (2009)) as well as for the distinction between qualified and non-qualified dividends from 2003 on to get a complete series for the 1960-2018 period. Before 1960, τ_t^d, τ_t^{skg} and τ_t^{lkg} rates are taken from Sialm (2009).

Altogether, the resulting rate accounts for most of the features of the tax code (maximum and minimum tax, partial inclusion of social security, earned income credit, phaseouts of the standard deduction and lowest bracket rate, etc.), as opposed to the nominal nature of statutory rates. Nevertheless, there is an important feature the NBER rates do not consider: qualified dividends and capital gains accruing to pension funds, individual retirement accounts, and nonprofit organizations which are subject to a tax-deferred scheme[19]. As typical in the literature, I also adjust the rates for the non-taxable share of equity income η_t. It was computed following Rosenthal and Austin (2016)[20]. Finally, ν_1 and ν_2 are the average share of dividends and short capital gains on reported equity income, based on Internal Revenue Service data[21]. See Appendix A for more details.

The top right panel of figure 1 plots the evolution of the resulting capital tax rate τ_t. Broadly

[19]The tax-deferred scheme implies that these funds are only taxed upon withdrawal. In this case, they face an ordinary income tax. As a result, they do not affect stocks equilibrium valuations. To see that, take a Traditional IRA, for instance. Contributions to it are deductible (under some conditions), which means that the investor is paying lower taxes (\$1 invested in the equity fund costs \$$(1-\tau)$). So, she saves taxes today with her contributions. She gets some income once she withdraw her money, which ends being \$$(1-\tau)$ per each dollar withdrawn. Hence, these future taxes are already offset by today taxes and have nothing to do with asset pricing. In fact, the tax deduction can imply an income tax cut since when the investor is working she usually is in a higher income bracket than when she gets retired, but this is another topic.

[20]Alternative computations have been done by the Federal Reserve, McGrattan and Prescott (2005) or Sialm (2009). Ultimately, they all try to compute the fraction of stocks owned by American households. The evolution coming out of the different methods is the same, although they differ in the levels.

[21]The average is taken for isolating tax rates variations from changes in the relative importance of dividends and capital gains. Sialm (2009) also builds a capital tax rate as a combination of dividend, short and long capital gains rates but using time-varying yields (i.e., dividends at time t over the price at t-1) as a weights instead. The resulting rates are very similar as the ones used in this paper.

speaking, since the removal of the dividends tax exemption in the Internal Revenue Code of 1954[22], the capital tax rate was relatively high and roughly constant until the late 1970s, when it started a long gradual dropping until the 2000s; then, it stabilized at a much lower rate. Three factors drove this decline. First, statutory tax rates experienced a long decline boosted by 1978 Congress capital gains tax reduction, Reagan's 1982 and 1986 tax acts (intensively reducing top marginal rates), Clinton's 1997 Taxpayer Relief Act (cutting capital gains tax rates) and Bush's 2001 and 2003 reforms (affecting mostly dividend tax rates). Second, there were changes in corporate payout policies, moving from dividends to capital gains via buybacks, that reduced effective taxes by decreasing the weight of dividends in total equity income. Finally, important regulatory changes involving pensions savings vehicles took place along the period (see McGrattan and Prescott (2005) for details on these reforms). The result of these regulatory reforms was an enormous switch in corporate equities ownership, moving from households (taxable units) to pensions funds and IRA accounts (non-taxable). According to my estimates, the share of equity income paying taxes drop from 86% in 1954 to just 29% in 2012[23].

2.2. Predicting volatility with taxes

The hypothesis of the paper is that the capital tax rate level is crucial to understand the movement of the price-dividend ratio volatility. Intuitively, the relative calmed period cohabited with relatively high and stable tax rates and the rise in long run volatility with the decline of tax rates. The left graph on figure 2 shows a first piece of evidence backing such hypothesis: high (low) one-period ahead rates are strongly associated with low (high) long-run volatility levels, exhibiting a correlation of -0.84 (using log-levels).

However, a closer examination of the time series graphs reveals that while there were two rounds of tax cuts with roughly the same intensity, only the tax break starting in 1995 was contemporaneous to an increase in volatility. This asymmetry is the result of a level effect coming from the non-linearity in the volatility-tax relationship as well as a change in the coefficients that govern it.

On the one hand, the correlation presented has used the variables in logarithms. In other words,

[22]Dividends were exempted of taxes from 1913 to 1953, excepting the 1936-1939 period when dividends deduction was temporarily removed in the Revenue Act of 1936.

[23]This sharp decline is in line with the estimations of McGrattan and Prescott (2005), Sialm (2009) or Rosenthal and Austin (2016).

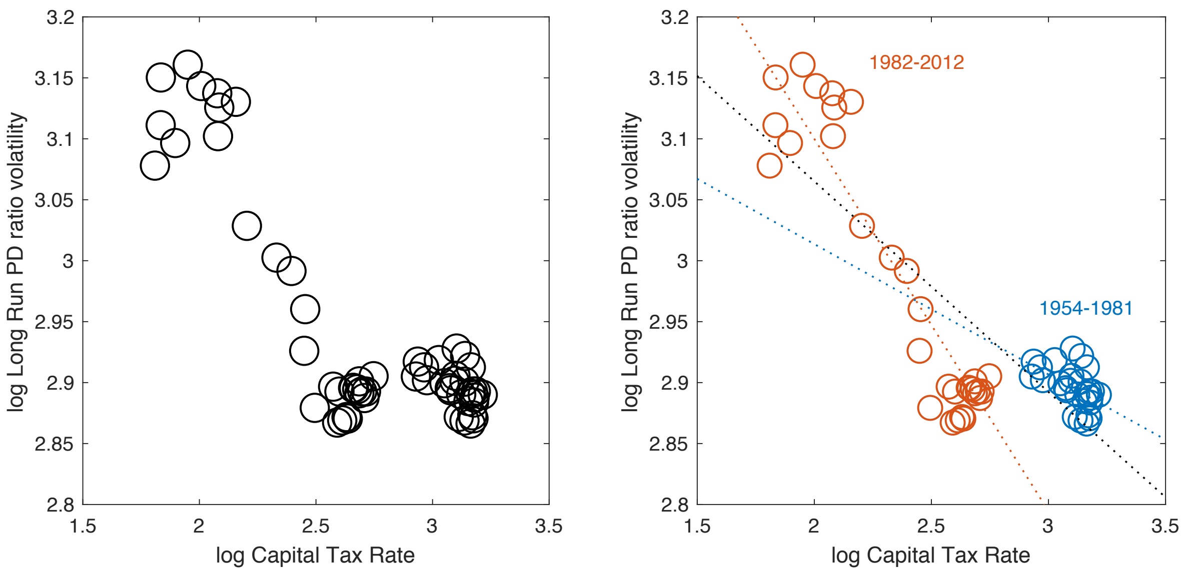

Figure 2: ***Scatter plot of the capital tax rate against the Price-Dividend long run volatility.*** *The two graphs plot the log effective average marginal tax rate against the log of the Price-Dividend long run volatility. The sample is 1954-2012, annual data. The left graph simply plots the data. The graph on the right plots the data distinguishing by two time periods, 1954-1981 and 1982-2012. Best fit lines are also plotted, showing the structural change in the tax-volatility relationship. The correlation amounts to -0.84 (p-value = 0.00) for the whole 1954-2012 sample, -0.52 (p-value = 0.01) for 1954-1981 and -0.94 (p-value = 0) for the 1982-2012 period.*

the relationship between capital taxes and volatility is best captured[24] by a power function as

$$\mathrm{v}_t = e^{\lambda_0}\tau_{t+1}^{\lambda_1} \tag{9}$$

This non-linear relationship implies a level effect: a tax cut has a much larger effect when taxes are already low. Hence, the reason why there was not any increase in long run volatility during the first tax cut is precisely the small effect of a tax cut at the 1980s tax level, which could have been easily offset by other forces (for instance, the reduction of macroeconomic volatility that began precisely in the mid 1980s). An alternative way to set out the role of the tax level is by looking at Impulse-Response Functions. Thus, if the tax level is (ir)relevant for volatility, a tax level shock must trigger a (transitory) permanent response in volatility. To test for that possibility, a minimal-

[24]According to all the usual information criteria.

ist VAR system comprising capital tax rates and long run volatility in log-levels is estimated[25,26]. Indeed, there is a permanent effect of a tax shock on the long run PD volatility. This fact posses a challenge to models in which the tax rate level is neutral regarding the price-dividend fluctuations, as is the case under Rational Expectations models. This question will be explored further in the quantitative analysis of the model.

Finally, a natural concern in time series is the presence of structural breaks. In fact, coefficient stability before and after 1982 in the log-linearized version of (9) is strongly rejected [27]. Thus, even though taxes and volatility move in the opposite direction throughout the sample, the intensity of this comovement changes, becoming much stronger since 1982 - namely, since the President Reagan's first tax cut. This fact is clearly observed in the right side scatter plot in figure 2: the best fit line turned much steeper since 1982. In this regard, figure 3 plots the observed long run volatility and the predicted one using the log-linearized version of (9), with and without structural break. The basic shape is well captured by the non-linearity alone, which explains up to 68.47% of the total variance of the volatility. Nevertheless, the structural break in coefficients clearly improves the model fitting, explaining up to 85.80% of the total variance.

To sum up, this section has robustly documented a new fact: the significant negative association between capital tax rates and the long run volatility of the Price-Dividend ratio throughout the 1954-2012 period in the United States. Particularly, it has highlighted that the tax level itself -and not only tax changes- plays a crucial role in (statistically) explaining volatility. The next section builds a structural model with a explicit causal mechanism linking taxes to volatility.

[25]There is some debate about whether is better to use first differences and then cumulate the effects to go back to levels or to compute directly a level VAR. It turns out that differencing is not an appropriate practice whenever the variables are cointegrated (in which case, differencing generates mis-specification) (see Sims (2011)). It is not clear that this is the case here; the ADF test yields (no) cointegration using (0) 1 lag. Anyway, this discussion is not very relevant here because using either levels or first difference with cumulation I get the same result: a permanent effect of a tax shock on PD volatility.

[26]The result is robust to the inclusion of additional variables in the VAR system, such as the GDP.

[27]According to the Chow test, with a p-value=0.00 (the test statistic is 43.48 while the critical value is 3.17). The estimated parameters are: $\{\hat{\lambda}_0, \hat{\lambda}_1\}_{\text{full sample}} = \{-0.32, -0.46\}$; $\{\hat{\lambda}_0, \hat{\lambda}_1\}_{\text{1954-1981}} = \{-0.28, -0.48\}$;$\{\hat{\lambda}_0, \hat{\lambda}_1\}_{\text{1982-2012}} = \{-1.14, -0.81\}$.

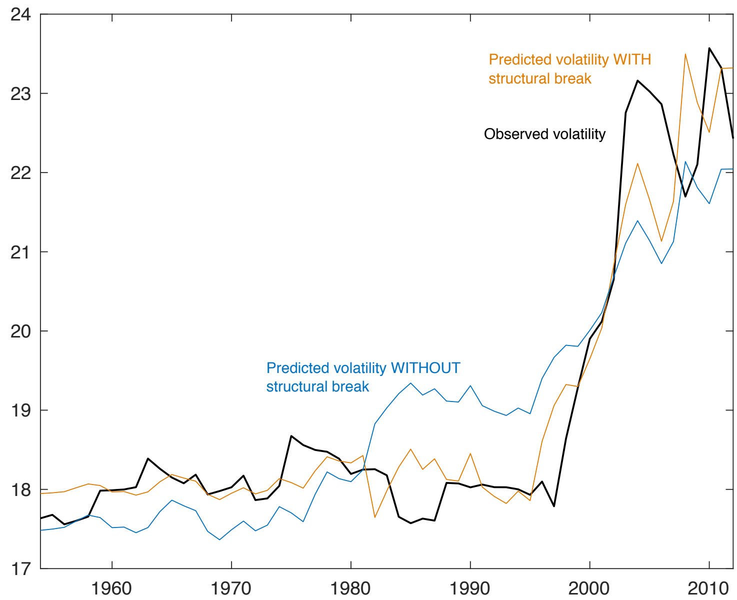

Figure 3: ***Observed and predicted long run price-dividend volatility.*** *The scatter plot on the left pictures the log of the Effective Average Marginal Capital Tax Rate against the log volatility level. The graph depicts the observed volatility v with the predicted volatility using the log-linearized version of equation (9) with and without structural change in λ in 1982. The data sample is 1954-2012.*

3.- The Model

In this section, a general equilibrium asset pricing model with capital taxes is set up and the equilibrium prices and their dynamics are analyzed. Section 3.1. presents the model. It is a consumption-based asset pricing model with time separable utility, rational agents who make decisions based on their probability subjective measure and a stochastic capital tax on dividends and capital gains. In this sense, it is the Lucas (1978) model, generalized by a subjective probability measure as Adam et al. (2016) and a stochastic capital tax. Section 3.2 shows a set of analytical results. First, the Rational Expectations Equilibrium is exposed, characterized by a time-varying price-dividend ratio fully determined by the tax dynamics. Afterwards, the model is solved for the case in which agents have imperfect knowledge about the pricing function and learn about it. It is shown how this richer setup gives rise to an effect of the tax level on the volatility that is key for replicating the new fact reported in Section 2, but is missing under Rational Expectations.

3.1.- Model structure

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 that delivers dividends and whose market price changes each period giving rise to capital gains and losses.

Resource processes. This is a pure exchange economy. When the time starts, each agent is endowed with one unit of stock ($s_{-1} = 1$). The dividends D_t it pays are exogenous, obeying a random walk with drift process

$$\log D_t = \log a + \log D_{t-1} + \log \epsilon_t^d \quad (10)$$

[28]

with $a \geq 1$ being the permanent component and $\epsilon_t^d \sim \log \mathcal{N}(1, e^{s_d^2} - 1)$ the random unpredictable shock. Shocks are independent and identically distributed. Capital gains are endogenously determined (next subsection). Finally, she receives W_t units goods (which one might call wages). The introduction of exogenous wages captures the fact that the observed consumption dynamics are less volatile than the dividend dynamics. Instead of specifying a process for wages, following most of the literature I set up a specific aggregate consumption process which under the right calibration resembles the empirical aggregate consumption process. It shares the structure with the dividend process, but differ in the shocks standard deviation, which is lower. Therefore, aggregate consumption growth obeys the following random walk with drift

$$\log C_t = \log a + \log C_{t-1} + \log \epsilon_t^c \quad (11)$$

where $\epsilon_t^c \sim \log \mathcal{N}(1, e^{s_c^2} - 1)$ and are i.i.d. as well. Besides, $(\epsilon_t^d, \epsilon_t^c)$ are log-Normal jointly distributed with correlation ρ_{cd}. The process for W_t is implied by resource feasibility.

Markets. Financial markets are competitive but incomplete. A negative amount of stocks is allowed up to some point (specified below)[29]. Goods market behaves also competitively. Then, stock (relative) prices are such that markets clear, that is, $C_t = W_t + D_t$ holds in the goods market and there is no trade in the stock market ($S_t = 1$ for all t).

[28] The time subindex t is standing for quarterly periods along Section 3, unless clearly indicated otherwise. Note the contrast with Section 2, where t was used for years and q for quarters. In most of Section 3 I get rid of the subindex q to save notation whenever the distinction is not relevant.

[29] In this case, a negative position would be equivalent to the so-called covered short position, at which an investor borrows shares and pays a borrow-rate during the time the short position is hold.

Fiscal System. There is a synthetic capital tax τ_t on stock dividends and a fraction $\pi \in (0, 1)$ of capital gains. This fraction stands for the fraction of realized out of total capital gains[30]. The capital tax rate is stochastic, following a unit root process[31]

$$\tau_t = \tau_{t-1} + \epsilon_t^\tau \tag{12}$$

where $\epsilon_t^\tau \sim\ ii\mathcal{N}(0, s_\tau^2)$[32]. Tax shocks are assumed to be orthogonal to dividend and consumption shocks $(\epsilon_t^d, \epsilon_t^c)$ as well as to risk-adjusted capital gains (i.e. $\mathbb{E}_t^{\mathcal{P}}\left(\epsilon_t^\tau \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \frac{P_t}{P_{t-1}}\right) = 0$)[33]. On the other hand, all the revenues are transferred to the individuals as a lump-sum payment T_t[34].

Investors' information set. Investors do not know that they all are identical. Besides, it is assumed they know the stochastic processes for dividends, aggregate consumption, capital taxes and lump-sum transfer[35]. However, in general, they are not endowed with knowledge about the pricing function. Thus, instead of having an exact forecast of $P_t \mid (D, W, \tau, T)^t$, each investor has a probabilistic view of P_t conditional on the states history. This option gives rise to an additional uncertainty source and a larger sample space Ω, made of the realizations of the five internally exogenous processes (i.e., taken as given by investors), being $\omega = \{D_t, W_t, P_t, \tau_t, T_t\}_{t=0}^\infty$ a typical element of Ω. Therefore, the underlying probability space is given by $(\Omega, \mathcal{B}, \mathcal{P})$ with $\mathcal{B}$ denoting the corresponding σ-algebra of Borel subsets of Ω and $\mathcal{P}$ the agent's subjective probability measure over $(\Omega, \mathcal{B})$. Thus, the subjective probability measure $\mathcal{P}$ is a fundamental of the model.

Investors' program. Each investor faces a consumption-saving decision: she chooses sequences of

[30]In this model there is no trade in equilibrium, so there are capital gains and losses but none of them are realized. This fraction overcomes this problem. It is assumed to be exogenous: this a shortcut that rules out the decisions involving the optimal realization of capital gains, since there does not exist such concept in the model. In this way, it losses one channel through which taxes might affect volatility: regulating the volume of trading (high capital gains taxes may be associated with a lock-in effect). This issue is left for future research. Finally, π is assumed constant because its empirical time series for the US is trend-less and shows little variation (see Section 4). Then, it only introduces a noisy variability in the model that might hide the sources of variability I am interested in.

[31]When the observed tax time series is fit into an AR(1) model, the estimated coefficients are not statistically different from 0 (intercept) and 1 (slope). Thus, the unit root process constitutes a realistic representation of the tax process.

[32]This is in line with the empirical behavior of the residuals from an unit root model. In this case, the residuals behave as a Gaussian white noise. Normality has been tested via the Shapiro-Wilk Normality test.

[33]Note this assumption does not deny the covariation of expected future taxes with expected capital gains. It simply states that tax surprises are not expected to affect future capital gains.

[34]The assumption of a full tax rebate is irrelevant for the results of the model because it is solved with an approximation based on the exogenous aggregate consumption process. In other words, with or without tax rebates the income effect is missing.

[35]For now, it is enough to assume that transfers are a known sequence of quantities.

consumption and stock holdings $\{C_t^i, S_t^i\}_{t=0}^{\infty}$ by solving an optimization program that uses their subjective probability measure $\mathcal{P}$ (i.e., agents are internally rational). Thus, they solve the following problem

$$\max_{\substack{\{C_t^i, S_t^i\}_{t=0}^{\infty} \in \Gamma \\ S_{-1}^i = 1}} \mathbb{E}_0^{\mathcal{P}} \sum_{t=0}^{\infty} \delta^t \frac{(C_t^i)^{1-\gamma}}{1-\gamma} \tag{13}$$

where

$$\Gamma = \Big\{ \{C_t^i, S_t^i\} \;\mid\; C_t^i + P_t S_t^i \le W_t + (P_t + D_t) S_{t-1}^i + T_t - \tau_t (D_t + \pi (P_t - P_{t-1})) S_{t-1}^i; \\ \underline{S} \le S_t^i \le \bar{S} \Big\} \tag{14}$$

The utility function $U(C_t^i) = \frac{(C_t^i)^{1-\gamma}}{1-\gamma}$ is a time-separable continuous, increasing in consumption $U'(C_t^i) > 0$ but strictly concave $U''(C_t^i) < 0$ function. Inada conditions hold. This parametric specification of U represents a risk averse investor (specifically a CRRA investor), being γ her risk aversion level. Γ sets up the feasible set. Lower and upper bounds on S_t^i are assumed for convenience; mathematically, these bounds ensure that the feasibility set is compact and, given the continuity of $U(C_t^i)$, the existence of a maximum is guaranteed by the Bolzano-Weierstrass theorem; economically, the lower bound rules out Ponzi schemes, which are out of interest here.

3.2. Analytical Results

3.2.1. Solution Method

The concavity of the investor's program and the convexity of the feasible set guarantee the sufficiency of the First Order Conditions (FOC) for an optimal (interior) plan. The optimal conditions for consumption and stocks boils down to the following Euler Equation

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

The Euler Equation (15) shows that today stock price is the discounted one-period ahead expected after-tax stock value. Under homogeneous beliefs[36], expression (15) can be rewritten in

[36]Note that with heterogeneous beliefs the agent who prices the stock each period is different and then the Law

such a way that present prices emerge as the discounted sum of all future expected cash flows associated with the equity. Nevertheless, the Efficient Markets Hypothesis fails to hold under subjective beliefs. To see that, iterating (15) forward, applying a standard transversality condition and taking expectations with respect to taxes the stock price become

$$P_t = (1-\tau_t)\sum_{j=1}^{\infty}\delta^j(1-\pi\tau_t)^{j-1}\mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+j}^i}{C_t^i}\right)^{-\gamma}D_{t+j}\right]\prod_{l=0}^{j-1}\left(1-\delta\pi\tau_t\mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+l}^i}{C_t^i}\right)^{-\gamma}\right]\right)^{-1} \tag{16}$$

In this expression, P_t is equal to a discounted sum of fundamentals, where the discount factor is made of current taxes and future individual consumption. If agents belief $C_t^i = C_t$ holds at all periods, the knowledge of fundamentals is sufficient to determine current prices as given by equation (16). However, this is the case only when the homogeneity of beliefs is common knowledge; then, agents apply the equilibrium condition to themselves[37]. Consequently, in general, $C_t^i \neq C_t$. This is the case because future consumption choices depend on future prices (i.e., $C_{t+j}^i(\{D, W, \tau, T, P\}^{t+j})$) and then on current prices too (because in general $P_{t+j} = P(P^{t+j-1}, ...)$, where the super-index indicates the history of prices.). It follows that knowledge about fundamental processes (W, D, τ, T) is completely insufficient to determine future consumption plans. Consequently, expression (16) is not useful to determine P_t (the right hand side of (16) depends on P_t). It follows that the one-period ahead Euler Equation (15) -rather than the discounted sum (16)- is the optimal condition that will characterize equilibrium prices under general conditions. This fact implies that stock prices will not be fully efficient, in the sense that after-tax dividends are not discounted at the equilibrium discount factor.

The standard approach to solve the problem would be to derive a time-invariant stock policy function from which we can pin down stock prices. Nevertheless, Adam et al. (2016) posed an approximation that allows for obtaining a closed-form equilibrium prices directly from the Euler Equation. The idea is that under some conditions, agents expectations about individual consumption growth would be approximately equal to their beliefs about aggregate consumption growth, that is, $\mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+1}^i}{C_t^i}\right)^{-\gamma}\right] \simeq \mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]$. Since agents know the true aggregate consumption process, the subjective belief about it is equal to the objective one, that is, $\mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right] = \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]$.

of Iterated Expectations cannot be applied.

[37]Moreover, in equilibrium $T_t = \int_0^1 \tau_t r_t^{\tau} P_{t-1} S_{t-1}^i di$, so that taxes and transfers would cancel out and taxes would have no effect on stock prices at all as long as agents know that they are all identical. In this case, this scenario could be avoided simply by assuming that taxes are not rebated to investors.

The gap between subjective individual and aggregate consumption expectations is given by

$$\begin{aligned} &\mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+1}^i}{C_t^i}\right)\right] - \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)\right] = \\ &\frac{1}{\frac{W_t}{D_t}+1}\mathbb{E}_t^{\mathcal{P}}\left[\frac{P_{t+1}}{D_t}(1-S_{t+1}^i) + \frac{T_{t+1}}{D_t} - \tau_{t+1}\left(\frac{D_{t+1}}{D_t} - \pi\frac{(P_{t+1}-P_t)}{D_t}\right)\right] \end{aligned} \tag{17}$$

This expression is approximately zero under the following conditions. First, dividends must be a small part of investor's income (i.e., $W_t >> D_t$). Second, the subjective expectations of the bracket term must be bounded. Since $S_t^i, \tau_t, T_t, D_{t+1}/D_t$ are bounded, it just requires bounded price expectations (i.e., $\mathbb{E}_t^{\mathcal{P}}\frac{P_{t+1}}{D_t} < \bar{M}$) and a bounded price-dividend ratio (i.e. $\frac{P_t}{D_t} < \bar{PD}$). This conditions are assumed to be true. Therefore, aggregate consumption can be plugged in the Euler Equation (15) and equilibrium prices can be directly computed[38].

With the previous assumptions, the conditional expectation about risk-adjusted dividend growth reads as

$$\mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+1}^i}{C_t^i}\right)^{-\gamma}\left(\frac{D_{t+1}}{D_t}\right)\right] \approx \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{D_{t+1}}{D_t}\right)\right] = \mathbb{E}_t\left[(a\epsilon_{t+1}^c)^{-\gamma}a\epsilon_{t+1}^d\right] = a^{1-\gamma}\rho \tag{18}$$

where

$$\rho = \mathbb{E}_t\left[(\epsilon_{t+1}^c)^{-\gamma}\epsilon_{t+1}^d\right] = exp\{\gamma(1+\gamma)\frac{s_c^2}{2} - \gamma\rho_{cd}s_c s_d\}$$

Additionally, let β_t^p be the risk-adjusted capital gains expectations

$$\beta_t^p \equiv \mathbb{E}_t^{\mathcal{P}}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{P_{t+1}}{P_t}\right)\right] \tag{19}$$

Finally, the conditional tax expectations is simply

$$\mathbb{E}_t^{\mathcal{P}}(\tau_{t+1}) = \mathbb{E}_t(\tau_{t+1}) = \tau_t \tag{20}$$

Plugging the expectations (18), (19) and (20) in the Euler Equation (15), the equilibrium price-

[38] Notice, however, that if the equality between individual and aggregate consumption beliefs were to hold almost surely in each period and contingency in the future, then the equilibrium price will be equal to the tax- and risk-adjusted discounted sum of dividends given by (16). It will be shown that it is not the case, though: prices will differ substantially from its Rational Expectations counterpart. Then, it must be the case that agents have different beliefs and the approximation is not zero.

dividend ratio reads as

$$\frac{P_t}{D_t} = \frac{\delta a^{1-\gamma}\rho z(\tau_t)(1-\tau_t)}{1-\delta\beta_t^p z(\tau_t)(1-\pi\tau_t)} \tag{21}$$

where

$$z(\tau_t) \equiv \left(1 - \delta exp\left\{s_c^2 \frac{\gamma(\gamma+1)}{2}\right\}\pi\tau_t\right)^{-1}$$

From this expression it is clear that that capital taxes has an impact on the prices no matter how price expectations are modelled. Specifically, there is a negative relationship between taxes and prices as the following result makes explicit:

Result 1: A permanent tax change triggers a negative permanent effect on the price-dividend ratio.

The equilibrium price-dividend ratio is a monotonic negative function of the capital tax rate as long as $\delta a^{-\gamma} exp\left\{s_c^2 \frac{\gamma(\gamma+1)}{2}\right\} < 1$. *It follows that*

$$\tau \geq \tilde{\tau} \Rightarrow \frac{P_t}{D_t}\left(\tau\right) \leq \frac{P_t}{D_t}\left(\tilde{\tau}\right)$$

This result contrasts with what Sialm (2006) obtained in a similar framework, that is, that only tax changes had an effect on stock prices and that the sign might be positive under some conditions[39]. As opposed to that, *Result 1* shows that the tax rate level itself matters for determining prices and that the relationship between tax and prices is unambiguously negative. This result emerges exclusively from the way capital taxes have been modelled, namely as a levy on dividends and capital gains (instead of as a tax on stock purchases, as in Sialm (2006)). In other words, it holds irregardless of investors' stock price information endowment. The next subsections derive results for the role of taxes on price-dividend volatility for two different approaches to β_t^p: Rational Expectations and Internal Rationality.

3.2.2. Rational Expectations Equilibrium

In this section, it is assumed that agents' subjective probability measure $\mathcal{P}$ coincides with the objective probability measure, that is, investors use the objective conditional expectation on risk-adjusted capital gains. It follows that agents know the pricing function so that the knowledge of fundamentals is sufficient to pin down future consumption and then current prices. Thus,

[39] The conditions are orthogonal

iterating forward on the Euler Equation (15) and applying a standard transversality condition, the equilibrium price-dividend ratio becomes

$$\frac{P_t}{D_t} = \frac{\delta a^{1-\gamma}\rho z(\tau_t)(1-\tau_t)}{1-\delta a^{1-\gamma}\rho z(\tau_t)(1-\pi\tau_t)} \tag{22}$$

Therefore, in a world with taxes, the RE price-dividend ratio becomes a time-varying object whose dynamics are completely determined by taxes. This result has two implications. First, the only chance of RE to generate a volatile enough PD ratio is that taxes behave in a highly fluctuating manner. Second, constant taxes lead to a constant PD ratio, which implies that the RE approach is unable to generate permanent effects of tax changes on the PD volatility as the ones reported in Section 2. Section 3.2.2.B. elaborates more on this.

3.2.2. Internally Rational Equilibrium

In this subsection, agents have perfect knowledge about the dividend aggregate consumption and tax stochastic processes but they do not know the pricing function. Instead, they use the subjective probability measure $\mathcal{P}$ to form expectations about risk-adjusted capital gains. Thus, the price beliefs are set up, pointing out how agents adapt their expectations whenever new observed information comes up. This updating gives rise to a feedback effect between beliefs and realized capital gains.

The beliefs system is a fundamental of the model. Before, it was established that agents know the dividends, aggregate consumption and tax stochastic processes as determined by expressions (10), (11) and (12). Under Rational Expectations, this information was sufficient to infer the process for prices because investors knew the pricing function. As opposed to that, in this section it is assumed that investors have imperfect knowledge about how fundamentals map into prices. Hence, their view about prices must be set up in order to have a fully determined subjective probability measure $\mathcal{P}$.

Under Rational Expectations, risk-adjusted price growth expectations boil down to

$$\beta_t^{RE} = \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{P_{t+1}}{P_t}\right)\right] = a^{1-\gamma}\rho \tag{23}$$

given that $\mathbb{E}_t(\tau_{t+1}) = \tau_t$, that is, tax changes are unanticipated[40]. Thus, RE investors expect a constant price-dividend ratio. Contrarily, agents with price imperfect knowledge expect a price-dividend ratio that persistently deviates upwards and downwards from its long-run mean, in line with the observed booms and busts. Specifically, they view risk-adjusted price growth as being made of a persistent and unknown permanent component and a transitory shock that can be formalized through the following state-space model

$$\begin{aligned} \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{P_{t+1}}{P_t}\right) &= Ab_t + v_t \\ b_t &= b_{t-1} + w_t \end{aligned} \tag{24}$$

where A is a coefficient that maps the unobserved state b on capital gains, v_t is the price growth transitory component and w_t is the unobserved belief shock; measurement and state noises are jointly Normally distributed and uncorrelated, that is,

$$\begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim iiN\left[0, \begin{pmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_w^2 \end{pmatrix}\right]$$

Note that this setup encompasses Rational Expectations as a particular case (i.e, when $A = 1$ and $b_t = a^{1-\gamma}\rho$).

Altogether, investors' beliefs system is composed of the processes for dividends (10), aggregate consumption (11), taxes (12) and the capital gains perceived law of motion (24). Under such system, the equilibrium price-dividend ratio is obtained by substituting subjective capital gains expectations into the equilibrium price-dividend equation (21), that is, $\beta_t^p = A\beta_t$ (where β is the estimate of the unobserved permanent component b). It follows that under Internal Rationality the price-dividend ratio becomes a non-linear function of current tax rates taxes and beliefs $\frac{P_t}{D_t} = \frac{P}{D}\Big(\tau_t, \beta_t\Big)$ rather than just taxes as in the RE case.

In this case, actual risk-adjusted capital gains are given by

$$\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\frac{P_t}{P_{t-1}} = \left(1 + \frac{\delta A\Delta[z(\tau_t)(1-\pi\tau_t)\beta_t]}{1-\delta z(\tau_t)(1-\pi\tau_t)A\beta_t}\right)\frac{z(\tau_t)(1-\tau_t)}{z(\tau_{t-1})(1-\tau_{t-1})}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\frac{D_t}{D_{t-1}} \tag{25}$$

where Δ is the difference operator. Thus, in general, what agents expect (24) and what really

[40] The case of anticipated tax changes is explored in the robustness analysis.

happens (25) is in general different. This potential discrepancy between perceived and actual capital gains gives rise to a learning problem: investors do not keep making systematic errors but rather use the new capital gains information to gain knowledge about the true law of motion. In other words, they learn about beliefs unobserved coefficient β. The problem, though, is a bit more complex: agents observe noisy realizations of risk-adjusted price growth without knowing what is worth it to learn from (persistent changes) and what is mere noise (or transitory). Thus, they have to filter out the realizations to pin down the hidden persistent component. The most efficient way to do so is by using the Kalman filter, which recursively predicts next state based on current price signals. Then, internally rational agents use the state-space model (24) to update their expectations.

In order to be close to RE, it is assumed that investors' initial prior is centered around the Rational Expectations value, that is

$$b_0 \sim \mathcal{N}\Big(a^{1-\gamma}\rho, \sigma_0^2\Big)$$

where σ_0 is the initial prior error variance. Posterior beliefs at any time t are centered about the prior estimates β_t

$$b_t \sim \mathcal{N}(\beta_t, \sigma_t^2)$$

where Normality comes from the noise distributional assumptions. On the one hand, prior estimates evolve according to the following rule[41]

$$\beta_{t+1} = \beta_t + k_t \Bigg\{ \Big(\frac{C_t}{C_{t-1}}\Big)^{-\gamma} \Big(\frac{P_t}{P_{t-1}}\Big) - A\beta_t \Bigg\} \tag{26}$$

where k_t is the optimal Kalman gain which captures the intensity at which agents update their beliefs when new information is acquired. It is given by

$$k_t = \frac{A\sigma_t^2}{A^2\sigma_t^2 + \sigma_v^2} \tag{27}$$

[41] Note that in this rule, β_{t+1} is determined using information of risk-adjusted stock price up to period t. According to expression (24), β_{t+1} would be a function of price growth at t+1, such that the object within the curling bracket would be the prediction error. However, since price growth at t+1 is a function of β_{t+1} (as (25)), there is simultaneity. The restriction imposed in (26) is a shortcut to infuse recursiveness in the system: beliefs in t+1 are updated using price growth at t; ; β_{t+1} generates price growth at t+1, and so on. In this way, the object within the curling brackets is the deviation of expected future prices from actual prices. This information lag is commonly used in the learning literature. The reason is that simultaneity generates important difficulties in learning models (see Adam (2003)) and it seems to perform worst (see Adam et al. (2017)).

Note that when the optimal gain tends to zero, the beliefs permanent component is close to its RE value (i.e., $k_t \to 0 \Rightarrow \beta_{t+1} \to \beta_t$ and then, $\beta_t \to \mathbb{E}[b_0] = a^{1-\gamma}\rho$). On the other hand, σ_t^2 is the prior variance of the belief parameter at time t. It is a Riccati recursion with initial condition σ_0^2 that can be computed without using actual capital gains

$$\sigma_t^2 = \sigma_{t-1}^2 + \sigma_w^2 - \frac{(A\sigma_{t-1}^2)^2}{A^2\sigma_{t-1}^2 + \sigma_v^2} \tag{28}$$

All in all, a self-referential system made up of perceived and actual capital gains equations (i.e., expressions (25)-(26)-(27)-(28)) arises. In other words, agents expectations determine actual prices and the latter are used to adapt the former. Thus, there emerges a feedback effect between expected and actual capital gains that generates rich stock price dynamics. The mechanics are as follows. Given a prior for the learning coefficients, capital gains expectations are obtained. This expectations together with the risk-adjusted dividend growth realization and actual taxes determine realized capital gains. Then, the gap between expected and actual capital gains is used to correct next period learning coefficients, using the optimal Kalman gain as a weight. The system is graphically represented in Figure 4.

3.2.3.- The role of taxes in price-dividend booms and busts

The dynamics of the equilibrium price-dividend ratio under Internal Rationality are completely characterized by the tax history and the dynamics of the beliefs time-varying coefficient β_t. The former is exogenous while the latter is produced by the perceived-actual capital gains feedback described before, which can be summarized in the following Second Order Difference Equation

$$\Delta\beta_{t+1} = k_t \left\{ \left(1 + \frac{\delta A \Delta[z(\tau_t)(1-\pi\tau_t)\beta_t]}{1 - \delta z(\tau_t)(1-\pi\tau_t)A\beta_t}\right) \frac{z(\tau_t)(1-\tau_t)}{z(\tau_{t-1})(1-\tau_{t-1})} a^{1-\gamma}\epsilon_t^d(\epsilon_t^c)^{-\gamma} - A\beta_t \right\} \tag{29}$$

The basic properties of this equation can be illustrated in a two-dimensional phase diagram on the (β_t, β_{t+1}) plane (keeping consumption and dividend shocks and tax rates at their mean values; their role in β dynamics is explored below). The phase diagram is shown in the left graph of figure 5. The solid lines depict the two nullclines, that is, they show the points at which $\Delta\beta_t$ and $\Delta\beta_{t+1}$ equal to zero, respectively. Their intersection gives rise to four phases for expectatiosn dynamics. In area A, $\beta_t > \beta_{t-1}$ and $\beta_{t+1} > \beta_t$, that is, agents are optimistic today and they will be so tomorrow. Thus, there is an upward momentum typical from a bull market. Exactly the opposite holds true for area C, characteristic of a declining market. The end of the boom (burst)

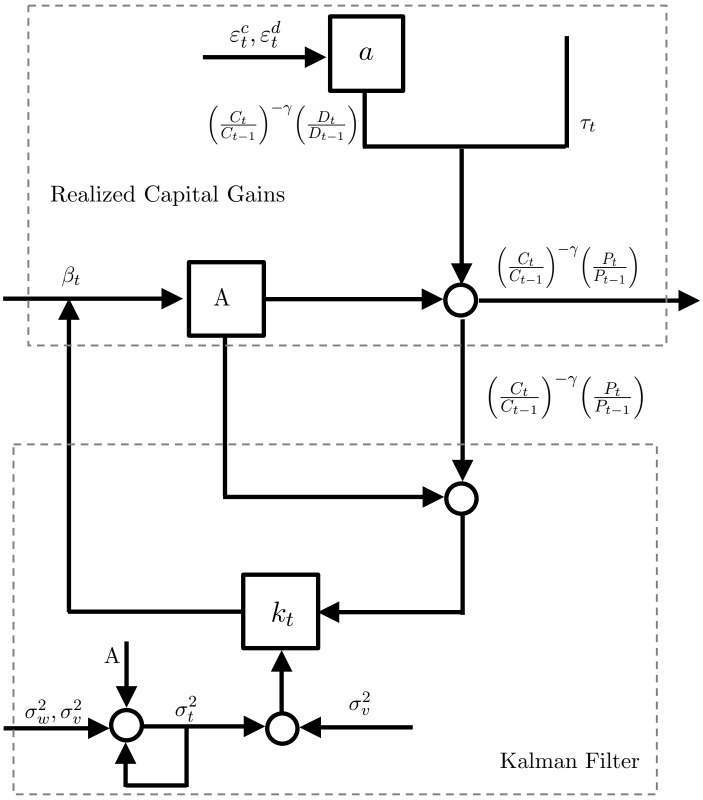

Figure 4: ***Block Diagram of the Model with Internal Rationality.*** *The top panel shows how risk-adjusted capital gains are generated, using consumption and dividend growth realizations, actual tax rates and capital gains expectations. The bottom panel depicts the Kalman predictive algorithm, that shows how beliefs are updated. Altogether, it illustrates the feedback mechanism that is at the core of the model.*

phase happens in area B (D), where $\beta_t > (<)\beta_{t-1}$ but $\beta_{t+1} < (>)\beta_t$, which implies that today's optimism (pessimism) is not strong enough for preventing a less optimistic (pessimistic) capital gains forecast. In other words, investor's optimism is defeated by a bad shock realizations (i.e., low dividends, a tax increase, etc.) that induces the investors to rectify her view. The role of an external factor (or "displacements") that comes to an end a self-fulfilling cycle is in line with the work of Minsky (1986) and Kindleberger (1978). In this case, there is reversion towards its fundamental value. Hence, beliefs orbit around the REE equilibrium without necessarily stopping at it[42].

Capital taxes play a key part in these expectations-driven booms and bust. Its role can be disentangled by distinguishing between the effects of tax changes and those triggered by the tax level itself. This distinction gives rise to the two main results.

Result 2: The direct effect of tax changes on price-dividend volatility is transitory.

[42]Momentum and mean reversion phenomena were formally proven by Adam et al. (2016).

Define the direct effect as the response of price-dividend volatility when all the remaining variables are kept constant (i.e., $\varepsilon_t^d(\varepsilon_t^c)^{-\gamma} = 1, \beta_t = b$). In this event and without loss of generality, consider the following tax history[43]*:*

$$\{\tau_t\} = \begin{cases} \tau_{q,t-j} = \tau_{1,t+1} & for \quad q \in [0,4], j = 0,1,...,t \\ \tau_{1,t+1} = \tau_{2,t+1} = \tau_{3,t+1} > \tau_{4,t+1} & \\ \tau_{q,t+j} = \tau_{4,t+1} & for \quad j = 2,3,... \end{cases}$$

Then, the long run price-dividend volatility only departs from its floor value ϕ_0[44] *in the K years after the tax cut:*

$$\{\mathrm{v}_t\} = \begin{cases} \mathrm{v}_{t-j} = \phi_0 & for \quad j = 0,1,...,t \\ \mathrm{v}_{t+j} > \phi_0 & for \quad j \in [1,K] \\ \mathrm{v}_{t+j} = \phi_0 & for \quad j = K+1, K+2, ... \end{cases}$$

where K is the number of annual lags of the realized volatility considered[45]*. The same result holds reverting the inequality direction. Moreover, the result also applies under Rational Expectations.*

Proof. Given the previous tax history, note that $\tau_{4,t+1} < \tau_{3,t+1}$ implies $z(\tau_{4,t+1}) > z(\tau_{3,t+1})$ and then,

$$\left\{\frac{P_{q,t}}{D_{q,t}}\right\} = \begin{cases} \frac{P_{q,t-j}}{D_{q,t-j}} = \frac{P_{1,t+1}}{D_{1,t+1}} & for \quad q \in [0,4]; j = 1,2,...,t \\ \frac{P_{2,t+1}}{D_{2,t+1}} = \frac{P_{3,t+1}}{D_{3,t+1}} < \frac{P_{4,t+1}}{D_{4,t+1}} & \\ \frac{P_{q,t+j}}{D_{q,t+j}} = \frac{P_{4,t+1}}{D_{4,t+1}} & for \quad q \in [0,4]; j = 2,3,... \end{cases}$$

From that, observe that the annual fixed window realized volatility is only different from zero

[43]The tax change is introduced in the last rather than in the first quarter to get an intra-annual variance different from zero. The result would apply as long tax rates change within years and not only between years. Otherwise, since the realized volatility uses an annual fixed window, constant taxes within each year lead to zero realized volatility. In this case, the absence of any effect would be an outcome of the fixed window assumption rather than of the asset pricing model, which is undesirable. Thus, the result is general for the model and particular for the realized volatility fixed window assumption.

[44]The long run volatility was defined in Section 2 as the permanent component of the conditional standard deviation of the price-dividend ratio, given by $v_t = \phi_0 + \phi_1 \sum_{j=0}^{K} \varphi_k(w) RV_{t-k}$, where RV is the realized volatility. Therefore, when $RV = 0$, $v_t = \phi_0$ (not necessarily zero).

[45]In the empirical analysis, it was set up K=10.

in the year of the tax cut, that is,

$$\{RV_t\} = \begin{cases} RV_{t-j} = 0 & for \quad j = 0, 1, ..., t \quad \text{and} \quad j = -2, -3, ... \\ RV_{t+1} \neq 0 \end{cases}$$

The result follows directly from that, given the definition of v_t. Observe that the persistence of RV_{t+1} over v_{t+j} for $j \in [2, K]$ depends on the lag weights $\varphi_k(w)$, which are monotonically decreasing on k. As a result, the tax shock has higher impact on contemporaneous volatility and then monotonically dies out. As for Rational Expectations, the result applies likewise since it does not rely on time-varying capital gains expectations (in fact, they are set constant by assumption). ■

Thus, the model displays a permanent effect of the PD ratio but a transient one of v as a response to a permanent tax change, no matter the expectation modelling choice. It follows that tax changes are not able to explain a permanent increase in volatility as the one documented for the US from mid 1990s on by themselves. However, when agents learn about capital gains there emerges an indirect effect of a tax change, coming from the fact that the tax level isolate prices from investors' beliefs changes. This is content of the following result.

Result 3: The indirect effect of a permanent tax change on price-dividend volatility is permanent. Define the indirect effect as the response of price-dividend volatility when the only variable kept constant is the tax rate (i.e. $\tau_t = \tau \quad \forall t$). In this event, the tax level determine the scale of price-dividend volatility, that is,

$$\tau < \tilde{\tau} \Rightarrow \{\mathrm{v}_t(\tau)\} > \{\mathrm{v}_t(\tilde{\tau})\}$$

This indirect effect is missing under Rational Expectations.

Proof. Let us proceed in two steps. First, note that the long run volatility v is a monotonic positive function of the realized volatility, which is the annual standard deviation of the price-dividend ratio. Then, it suffices in proving the effect of taxes on the variance of the price-dividend ratio and directly infer the qualitative results for v.

Second, evaluate the consequence of constant taxes under each expectations regime. On the one hand, under constant taxes the RE price-dividend ratio becomes a constant so that the PD variance is just zero. In other words, in a RE world the tax rate level is neutral with respect the PD ratio volatility. In the IR case, the tax level intermediate between expectations volatility and PD volatility. To see that, take a first order Taylor approximation around the RE value (i.e., $\beta_t = a^{1-\gamma}\rho$) and then take the variance of both sides such that

$$\mathbb{V}ar\Big[\frac{P}{D}\Big(\beta_t\Big)\Big] \approx \underbrace{\mathbb{V}ar\Big[\frac{P}{D}\Big(a^{1-\gamma}\rho\Big)\Big]}_{\text{RE var. }=0} + \underbrace{\omega(\tau)}_{\text{Tax wedge}} \times \underbrace{\mathbb{V}ar(\beta_t)}_{\text{Non-fundamental var.}}$$

where

$$\omega(\beta_t, \tau) = \Big(\frac{dP_t/D_t}{d\beta_t}\Big)^2\Big|_{\beta_t = a^{1-\gamma}\rho}$$

ω is a monotonically negative function of the tax rate level τ, that is,

$$\frac{\partial\omega(a^{1-\gamma}\rho, \tau)}{\partial\tau} < 0$$

with $\lim_{\tau\to 1}\omega = 0$ and $\lim_{\tau\to 0}\omega > 0$. In other words, the tax rate level acts as a (permanent) wedge between non-fundamental volatility and price-dividend volatility.

Furthermore, this result is amplified by the dependence of $\{\beta_t\}$ on the tax level. It can be observed directly from the Second Order Difference Equation governing the dynamics of β_t (expression (29)). Thus, for constant taxes it turns out

$$\Delta\beta_{t+1} = k_t\Bigg\{\Bigg(1 + \frac{\delta A z(\tau)(1-\pi\tau)\Delta\beta_t}{1 - \delta z(\tau)(1-\pi\tau)A\beta_t}\Bigg)a^{1-\gamma}\epsilon_t^d(\epsilon_t^c)^{-\gamma} - A\beta_t\Bigg\}$$

that is, taxes are a wedge that regulates to what extent actual beliefs change $\Delta\beta_t$ translate into one period ahead change $\Delta\beta_{t+1}$. Therefore, higher taxes would permanently make an increase in optimism less likely. ■

The second part of the proof is illustrated in the right graph of figure 5. As observed, the higher the tax rate the larger the reversion areas in the phase diagram of the beliefs coefficient β, making momentum more difficult. Consequently, expectations tend to depart less from their fundamental value and the PD ratio volatility is reduced. *Result 3* can be also stated in more intuitive terms: the tax rate level determines the sensitivity of price to expectations fluctuations. In other words, the

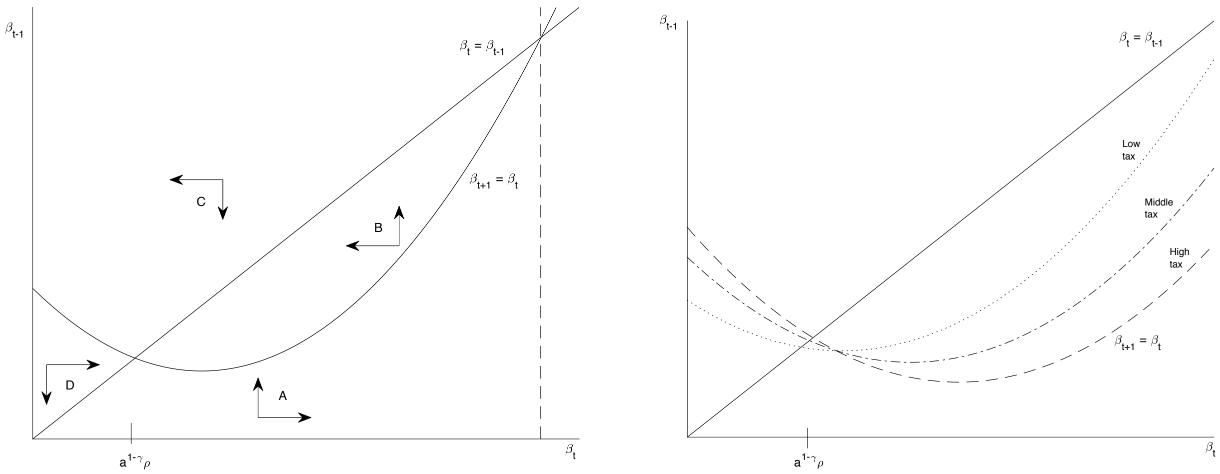

*Figure 5: **Phase Diagram of the belief coefficient** β. The diagram on the left illustrates the qualitative dynamic behavior of the learning coefficient β, given by the 2nd Order Difference Equation(29). It considers $\Delta\beta_{t+1}(\beta_t, \beta_{t-1})$. The rest of the variables are fixed at mean values. The solid lines pictures the nullclines that makes $\Delta\beta_t$ and $\Delta\beta_{t+1}$ equal to zero, respectively. Four phases emerges: boom (A), burst (C) and two reversion zones (B and D). The diagram on the right shows the movements in the $\Delta\beta_{t+1} = 0$ curve at different tax rate levels.*

beliefs elasticity of stock prices is monotonically decreasing on the tax rate[46]. Figure 6 illustrates this point.

The joint effect of results 2 and 3 implies that only a model in which agents learn about capital gains is able to generate a permanent response in the long run volatility to a tax change, in line with the facts documented in Section 2. The reason is the role the tax level plays, acting as a wedge between investors' sentiment fluctuations and prices. Particularly, the tax cuts increased the sensitivity of prices to beliefs as well as make beliefs more volatile. Altogether, it has been shown that the capital tax rate level matters not only for determining the level of the price-dividend but also its degree of volatility.

[46]The expectations elasticity of prices is given by

$$\frac{\partial P_t / P_t}{\partial \beta_t / \beta_t} = \frac{\delta z(\tau)(1-\pi\tau)\beta_t}{1-\delta z(\tau)(1-\pi\tau)\beta_t} \equiv \varepsilon^p_{\beta,t} = \varepsilon^p_\beta(\tau, \beta_t)$$

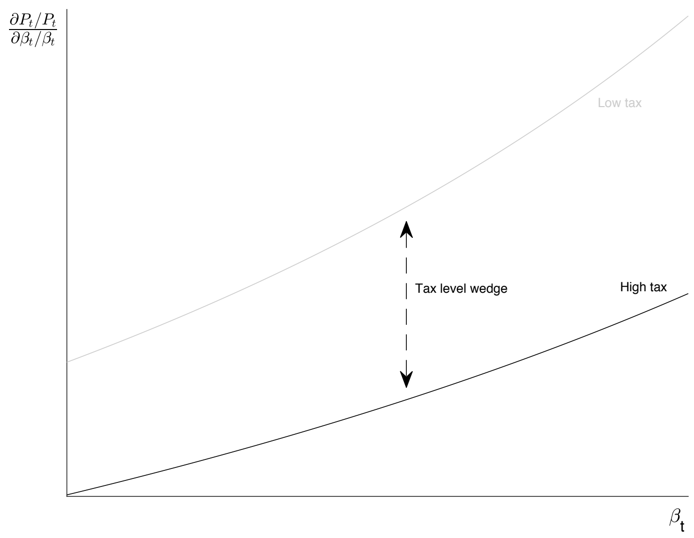

Figure 6: ***Beliefs elasticity of stock prices at different tax levels.*** *The curves plot the elasticity of stock prices to changes in capital gains expectations. That elasticity is a function of capital gains expectations and the tax rate level. Beliefs are plot in the x-axes and the effect of a tax level is plotted as a movement in the curve.*

4.- Quantitative Analysis

This section assess the quantitative performance of the model. The goodness of the model is evaluated in terms of its ability to match a set of empirical statistics, that includes standard asset pricing facts and additional statistics from Section 2. The basic set of empirical targets is established in section 4.1. A subset of the model parameters is calibrated from US data while the remaining ones are estimated through an extension of the Simulated Method of Moments (henceforth SMM) based on functions of moments (called 'statistics') instead of pure moments. Then, individual t-tests for checking the statistical equivalence between model implied statistics and empirical ones are established. The procedure is described in section 4.2. Additionally, section 4.3. reports the baseline results for both the Rational Expectations and Internal Rationality specifications.

Remarkably, the model is able to generate a high enough historical equity premium with low and stable risk-free rate, realistic consumption and dividend processes, positive discount factor and low risk-aversion. This fact is analyzed in depth in Section 4.3., which offers a decomposition of the stock returns geometric mean, pointing out the relevance of generating an upward trended price-dividend. Furthermore, section 4.4. runs a robustness analysis of the previous results along two key dimension: the case of fully anticipated tax changes and an alternative learning rule that uses the steady state Kalman gain. It turns out that the model performs equally well under those alternative specifications. Finally, section 4.5. poses a historical test to the model, feeding it with the observed US time series and showing its ability to replicate the observed time trajectory of the long run price-dividend volatility.

4.1. Asset Pricing Facts

The rightness of the model is assessed in terms of its ability to replicate a set of statistics. This set is chosen as to include standard asset pricing facts as well as the additional ones documented in Section 2. By doing so, the model test consists in reproducing the new facts without losing capacity to do well in the standard dimensions.

The literature has extensively reported a list of facts, characterizing the behavior of stock prices and returns and their relationship with interest rates and dividends. Beyond the high volatility characterizing dividends growth with respect to consumption growth, two main issues summarize the whole evidence (see, for instance, Cochrane (2017)). On the one hand, the equity premium, that is, the difference between the expected return on stocks and the risk-free rate. It is incorporated by considering the quarterly average return on stocks $\mathbb{E}(r^s)$ and the 3-month T bill interest rate $\mathbb{E}(r^b)$. The second issue is the large volatility of financial series. First, the cost of higher stock returns is a larger variance (against a much more stable risk-free rate), as captured by the standard deviation of equity returns $\sigma(r^s)$. Secondly, stock prices fluctuates much more than dividends. It follows that price-dividend ratios are highly volatile, showing persistent deviations from their long run average. This pattern is observed by looking at the PD ratio standard deviation $\sigma(PD)$ and its first order autocorrelation coefficient $corr(PD_t, PD_{t-1})$. A different way of looking at this matter is through the predictability of excess stock returns[47]. Returns are forecasted if the coefficient c_n^1 of the following regression turns out significant, with the predicted variability given by the R_n^2. Usually, the LHS variable is the acumulated excess return from quarter t to n years forward (I consider 5 years for the empirical analysis).

$$R_{t,n}^e = c_n^0 + c_n^1 \frac{P_t}{D_t} + u_{t,n} \tag{30}$$

[47] The argument is that if the covariance of the PD ratio with future dividend growth is low, it must be the case that the covariance with future returns is high, which implies that PD ratio must predict part of those future returns (as first mentioned by Cochrane (1992)).

Furthermore, this paper has highlighted some additional facts. As a complement for volatility statistics, Section 2 has shown the evolution of the long run price-dividend volatility, which can be captured by its first and second moments ($\mathbb{E}(\mathrm{v})$ and $\sigma(\mathrm{v})$, respectively). The main focus of the paper is the analysis of the relationship between taxes and stock prices. Traditionally, the tax capitalization has been presented as a level relationship (e.g. McGrattan and Prescott (2005); see Section 2). I take that into account by including the correlation between taxes and the PD ratio $corr(PD_t, \tau_t)$. Besides the level relationship, I have shown that taxes crucially affect the long run volatility of the PD ratio. This new fact is captured here by the correlation between the long run volatility and taxes $corr(\mathrm{v}_t, \tau_t)$. All statistics and their estimated standard deviation are reported in the first two columns of Table 1.

4.2. Estimation Procedure

This section explains how the model-implied statistics are obtained. On the one hand, the randomness of the variables is tackled through a Montecarlo experiment. On the other hand, the parameters value are pin down following a mixed strategy: a subset of parameters is calibrated from US data and the rest are estimated via SMM.

At this point, it must be noted that the equilibrium PD ratio faces a discontinuity. For this reason, simulation requires to set up the following modified belief updating equation to ensure non-negative prices

$$\beta_{t+1} = w\left(\beta_t + k_t\left\{\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\left(\frac{P_t}{P_{t-1}}\right) - A\beta_t\right\}\right) \tag{31}$$

where

$$w(x) = \begin{cases} x & \text{if} \quad x \leq \beta_t^L \\ \beta_t^L + \frac{x-\beta_t^L}{x+\beta_t^U-2\beta_t^L}(\beta_t^U - \beta_t^L) & \text{if} \quad x > \beta_t^L \end{cases} \tag{32}$$

and

$$\beta_t^q = \frac{PD^q - a^{1-\gamma}\rho z(\tau_{t+1})(1-\tau_{t+1})}{PD^q \delta A z(\tau_{t+1})(1-\pi\tau_{t+1})} \tag{33}$$

for $q = L, U$. Thus, this projection facility starts to dampen belief coefficients that imply a price-dividend ratio equal to PD^L and sets an effective upper bound at PD^U. Projection facilities are usual devices in this sort of algorithms (see Ljung (1977)); particularly, (32) is similar to the one used by Adam et al. (2016). It can be understood in a Bayesian sense, so that agents attach zero probability to beliefs coefficients implying a PD ratio higher than PD^U.

The model has a total of eleven parameters, namely, the realized to total capital gains ratio π, the coefficient of relative risk aversion γ, the discount factor δ, the mean of dividend and consumption growth a, the standard deviation of consumption and dividend growth s_c, s_d, the correlation of consumption and dividend growth processes ρ_{cd}, the unobserved state noise σ_w^2, which determines the Kalman gain, the scale factor A in capital gains perceived law of motion and the PD bounds for the projection facility (PD^L, PD^U). A subset of them is picked from US data and the rest is estimated. Specifically, the vector $\widetilde{\theta} = \{\pi, s_c, s_d, \rho_{cd}\}$ is calibrated as follows. The fraction of taxable capital gains is 0.10, which is the average of the realized capital gains (in adjusted gross income, from the IRS) over total capital gains (from the US Financial Accounts) for the 1956-2014 period[48]; dividend growth standard deviation s_d is set equal to 1.67, which is the value for the real quarterly series from the Global Financial Database[49]; consumption growth standard deviation s_c is 0.71, using the BEA real quarterly personal consumption series; consumption-dividend growth correlation ρ_{cd} is 0.13. The left parameters are $\theta = \{\delta, a, \gamma, \sigma_w^2, PD^L, PD^U, A\}$. The SMM estimate $\hat{\theta}$ is the vector that minimizes the weighted quadratic distance between observed and model-implied statistics as follows[50]

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[\hat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta)\right]' \hat{\Sigma}_{\mathcal{S}}^{-1} \left[\hat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta)\right] \tag{34}$$

In this expression, $\hat{\mathcal{S}}$ and $\widetilde{\mathcal{S}}(\theta)$ are the set of observed and simulated statistics respectively. It is implicit $\hat{\mathcal{S}} \equiv \hat{\mathcal{S}}(\theta_0)$, that is, empirical statistics are generated by the model at the true parameter value (θ_0). In turn, $\hat{\Sigma}_{\mathcal{S}}$ is the weighting matrix, which determines the relative importance of each statistic deviations from its target. A diagonal weighting matrix whose non-zero elements are the inverse of the Newey-West estimated variance-covariance matrix of the data statistics is used. Thus, the procedure minimizes the quadratic sum of t-statistics of the targeted objects, leading directly towards matching the individual moments. Model-implied statistics are generated through a Montecarlo experiment with 1000 realizations involving both consumption and dividends growth shocks. Contrarily, the historical time series of tax shocks is used, since the focus is to assess the consequences of the observed tax cuts.

[48]I thank Jacob Robbins for this data. Taxable capital gains are a time-varying object, with extreme movements around the Great Recession. However, I hold it fixed to focus on the tax variation due to both tax rates and ownership changes.

[49]See Appendix A.

[50]The procedure mimics Adam et al. (2016) and also uses their asymptotic results. They extended the standard SMM approach as outlined by Duffie and Singleton (1993) to deal with function of moments rather than pure moments using the delta method. See Adam et al. (2016) for more details.

4.3. Baseline Results

Table 1 summarizes the results. The second and third columns report the asset pricing statistics from the US data as described in Section 4.1. and their estimated standard deviation, respectively. The next columns show the simulated statistics and their associated t-statistics for the Internally Rational model and the RE version. Finally, the bottom panel presents the estimated parameter vector for each approach[51].

As observed, the IR model is able to pass all the t-tests, many of them with t-statistics below 1. Thus, the IR model replicates all the observed phenomena. On the one hand, it generates enough volatility, as already showed by Adam et al. (2016). This result is reinforced by showing the ability of the model to match the first and second moment of the long run volatility. This is not a mechanical consequence; for instance, there are ways for the model to match the standard deviation of the price-dividend ratio by producing frequent booms and busts that would tranlate into a too high standard deviation of the long run volatility. On the other hand, the IR model with stochastic taxes expands the matched statistics. First, it reproduces the equity premium (see next section). Second, it matches the connection between tax rates and stock valuations levels (in line with McGrattan and Prescott (2005) or Sialm (2009)). Finally, it is also able to produce a strong enough negative correlation between taxes and the long run volatility.

The empirical performance of the learning model looks even better when compared against the Rational Expectation model[52]. It shows able to replicate the PD level, signaling that RE manages to captures a sort of long run trend. However, it fails at all the other dimensions. Its inability to produce enough market volatility and a decent equity premium are well-known[53]. On top of that, these results point out that it is also unsuccessful in reproducing the observed connection between capital taxes and stock prices. On the one hand, it overstates the correlation between tax rates and the PD level, being close to -1. The reason is that the trajectory of the RE-PD ratio is uniquely determined by the dynamics of future taxes, as showed by expression (22). On the other

[51]The parameter space has been constrained as to ensure $\delta \in (0, 1]$.

[52]Observe, however, that RE have 3 free parameters less than IR.

[53]Note that the fact that RE model manages to produce an almost acceptable long run volatility mean is not a merit of RE but just a by-product of the fact that I use the empirical coefficients of the long run volatility (i.e., $\hat{\phi}_0, \hat{\phi}_1$) to compute the simulated v.

hand, RE produces a barely uncorrelated taxes and long run volatility due to the fact that tax changes trigger an immediate and transitory effect on volatility, as proven in Section 3. This issue is explored further in Section 4.5.

Table 1: ***Baseline Estimation Results.*** *This table reports data statistics, their estimated standard deviations as well as the full set of statistics and t-statistics (with the null of the equality between observed and simulated statistics) for the Internally Rational and Rational Expectations versions of the model. Furthermore, it reported the estimated parameter vector $\hat{\theta}$. The data sample is 1954:I-2012:I, except for stock and bond returns that starts in 1960:I*[54]

	US data		IR model		RE model	
	$\hat{\mathcal{S}}_i$	$\hat{\sigma}_{\hat{\mathcal{S}}_i}$	$\mathcal{S}_i(\hat{\theta})$	t-stat	$\mathcal{S}_i(\hat{\theta})$	t-stat
$\mathbb{E}(D_t/D_{t-1})$	1.0037	0.0021	1.0030	0.33	0.9985	2.42
$\mathbb{E}(r^s)$	1.64	0.56	1.92	-0.50	0.53	1.97
$\mathbb{E}(r^b)$	0.28	0.11	0.31	-0.22	0.41	-1.23
$\sigma(r^s)$	8.26	0.59	7.10	1.96	1.81	10.95
$\mathbb{E}(PD)$	148.99	15.32	142.47	0.43	162.34	-0.87
$\sigma(PD)$	64.07	12.17	74.97	-0.90	11.56	4.32
$corr(PD_t, PD_{t-1})$	0.95	0.05	0.97	-0.23	0.82	2.82
c_5^1	-0.0027	0.0017	-0.0060	1.88	0.0003	-1.79
R_5^2	0.14	0.17	0.34	-1.18	0.00	0.84
$\mathbb{E}(\mathrm{v})$	18.96	0.73	20.03	-1.46	17.14	2.49
$\sigma(\mathrm{v})$	1.80	0.53	2.64	-1.59	0.01	3.39
$corr(PD_t, \tau_t)$	-0.64	0.08	-0.61	-0.40	-0.99	4.51
$corr(\mathrm{v}_t, \tau_t)$	-0.75	0.13	-0.68	-0.50	-0.09	-5.07
Estimated parameters						
$\hat{\delta}$			1.0000		0.9932	
$\hat{\gamma}$			0.9341		2.0049	
$\hat{\sigma}_w^2$			4.14e-07			
$\hat{A}$			1.0309			
$\hat{PD}^L$			292.63			
$\hat{PD}^U$			326.20			

4.4. The Equity Premium

One of the appealing features of the model is its ability to generate a high enough equity premium together with a low and stable risk-free rate, realistic consumption and dividend growth processes, positive discount factor and low risk-aversion. In other words, it accounts for the (historical) equity premium puzzle as stated by Cochrane (2017)[55]. There are two reasons. On the one hand, the model-implied PD ratio matches well the first and second moments. On the other hand, the decline in taxes gives rise to an upward trend in the price-dividend ratio that naturally increases the sample stock returns mean. Both factors do not affect any of the factors determining the risk-free rate.

This claim can be proved by performing the following decomposition of the stock return geometric mean

$$\left(\prod_{t=1}^{N}\frac{P_t+D_t}{P_{t-1}}\right)^{\frac{1}{N}} = \underbrace{\left(\prod_{t=1}^{N}\frac{D_t}{D_{t-1}}\right)^{\frac{1}{N}}}_{R_1}\underbrace{\left(\frac{PD_N+1}{PD_0}\right)^{\frac{1}{N}}}_{R_2}\underbrace{\left(\prod_{t=1}^{N-1}\frac{PD_t+1}{PD_t}\right)^{\frac{1}{N}}}_{R_3} \tag{35}$$

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 linear 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. As seen in table 2, the learning model matches well R_1 and R_3. This is simply a by-product of the fact it replicates the mean dividend growth and first and second moments of the price-dividend ratio. Furthermore, it produces a trended price-dividend ratio, resulting from the direct and indirect effects of the decline in tax rates. The component R_2 turns out to be a bit too high, which explains why the model-implied $\mathbb{E}(r^s)$ is a slightly above the observed one. However, this overstatement of R_2 is very dependent on when the initial and terminal periods are set; for instance, if the data sample starts at 1954:I instead 1960:I, the observed R_2 becomes 1.0043 and the simulated one just 1.0059, that is, they become much closer.

Crucially, the learning model produces this high mean stock return together with a low and stable risk-free rate. It is given by

$$r_t^b = \left(\frac{1}{\delta a^{-\gamma} exp\left\{\frac{s_c^2}{2}\gamma(\gamma+1)\right\}} - 1\right)\left(\frac{1}{1-\tau_t^b}\right) \tag{36}$$

Note that the volatility and the rising trend were generated by factors which are orthogonal to the determination of the risk-free rate, evading then the problematic routes of increasing the risk

[55] 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

Table 2: ***Decomposition of the stock return geometric mean.*** *The table shows the decomposition raised in expression 35 for the US as well as for the estimated IR and RE models. Simulated statistics are generated via a Montecarlo experiment. The data covers the 1960:I-2012:I period.*

	US data		IR model		RE model	
	$\hat{\mathcal{S}}_i$	$\hat{\sigma}_{\hat{\mathcal{S}}_i}$	$\mathcal{S}_i(\hat{\theta})$	t-stat	$\mathcal{S}_i(\hat{\theta})$	t-stat
R_1	1.0033	0.0015	1.0028	0.40	0.9997	3.39
R_2	1.0025	-	1.0057	-	1.0010	-
R_3	1.0075	0.0012	1.0081	-0.40	1.0061	1.24

aversion parameter (which leads to the risk-free rate puzzle) or increasing consumption volatility (which leads to an unrealistic consumption process). Thus, non-fundamental volatility turns out a powerful way to solve the equity premium. However, the model without taxes only generates half the required premium (Adam et al. (2016)). Then, the second element is the role of tax cuts in generating a trended price-dividend ratio, which contrasts with the trendless PD ratio typical of many asset-pricing models.

Note, though, that bond interest tax rates τ^b also determine the risk-free rate. Thus, since bond taxes experienced a decline similar to the one document for dividends and capital gains taxes, r_t^b exhibits this declining trend too. However, the effect of bond tax breaks on the risk-free rate is tiny[56]. Thus, the decline in bond taxes does not cause an undesirable highly volatile risk-free rate, that is, capital taxes helps to solve the mean stock return without generating a volatile risk-free rate. This fact is related to an observation aforementioned: the success of the model is not due to the mechanic effect of a tax cut but to the indirect effects that it triggers[57].

4.5. Robustness Analysis

In this section, the robustness of the previous results is analyzed. Particularly, the sensitivity of the outcomes is explored across two dimensions: the case of anticipated tax changes and an alternative learning rule using the steady state Kalman filter.

[56]The observed $\sigma(r_t^b)$ is equal to 0.84 and the estimated one is 0.01, with a t-statistic of 9.47.

[57]In this line, it was shown that the direct effect of taxes (the only existent in RE) was completely unable to produce enough volatility.

4.5.1. Anticipated tax changes

Up to now, it has been assumed that $\mathbb{E}_t^{\mathcal{P}}(\tau_{t+1}) = \tau_t$, that is, tax changes surprised agents. Nevertheless, it can be argued that financial markets are forward looking and try to anticipate them. In fact, Kueng (2014) showed that investors do so remarkably well. Thus, in this section the model is solved and estimated for the case that $\mathbb{E}_t^{\mathcal{P}}(\tau_{t+1}) = \tau_{t+1}$ and $\mathbb{E}_t^{\mathcal{P}}(\tau_{t+n}) = \tau_{t+2}$ for $n \geq 2$. In other words, it is assumed that investors know the tax changes coming in the near future[58].

Key equations of the model must be changed as a result. To begin with, the general form of the equilibrium price-dividend ratio becomes a function of one-period ahead taxes

$$\frac{P_t}{D_t} = \frac{\delta a^{1-\gamma}\rho z(\tau_{t+1})(1-\tau_{t+1})}{1-\delta\beta_t^p z(\tau_{t+1})(1-\pi\tau_{t+1})} \tag{37}$$

Consequently, the RE expectations about capital gains become a function of one period ahead tax changes, that is,

$$\beta_t^{RE} = \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{P_{t+1}}{P_t}\right)\right] = \underbrace{\frac{(1-\tau_{t+2})z(\tau_{t+2})(1-\delta a^{1-\gamma}\rho(1-\pi\tau_{t+2}))}{(1-\tau_{t+1})z(\tau_{t+1})(1-\delta a^{1-\gamma}\rho(1-\pi\tau_{t+1}))}}_{\equiv h(\tau_{t+1},\tau_{t+2})} a^{1-\gamma}\rho \tag{38}$$

As before, investors subjective model for prices is set up as to generalize the RE version, that is,

$$\begin{gathered}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{P_{t+1}}{P_t}\right) = Ah(\tau_{t+1},\tau_{t+2})b_t + v_t \\ b_t = b_{t-1} + w_t\end{gathered} \tag{39}$$

In other words, investors consider future tax changes when forming expectations about capital gains[59]. In this case, the effect of a tax change is amplified through this expectations channel. The estimation results of this specifications are reported in Table 3. It turns out that the model with anticipated tax changes performs even slightly better than the benchmark model[60]; all the individual significance test are passed too.

[58] The case of perfect foresight about the whole tax history was also explored. On top of being informationally super demanding, it performs poor in quantitative terms. The reason is that future tax cuts cause an increase in volatility way before they took place while volatility goes down to its floor level when they are implemented. As a result, the correlation between long run volatility and taxes becomes even positive.

[59] Many equations change as a consequence: the actual price growth as well as the updating rule, the Kalman gain and the state variance recursion. Since the change is easy to infer, I do not include them here to save space.

[60] The value of the criterion function becomes 9.75 vs. 15.31 in the baseline estimation.

*Table 3: **Robustness Estimation Results.** This table reports observed and simulated statistics and t-statistics (with the null of the equality between observed and simulated statistics) for the Internally Rational model with anticipated tax changes and steady state Kalman gain (constant gain learning). Furthermore, it reported the estimated parameter vector $\hat{\theta}$.*

	US data	Anticipated Tax Changes		Constant Gain Learning	
	$\hat{\mathcal{S}}_i$	$\mathcal{S}_i(\hat{\theta})$	t-stat	$\mathcal{S}_i(\hat{\theta})$	t-stat
$\mathbb{E}(D_t/D_{t-1})$	1.0037	1.0023	0.63	1.0026	0.51
$\mathbb{E}(r^s)$	1.64	1.98	-0.61	1.84	-0.37
$\mathbb{E}(r^b)$	0.28	0.23	0.49	0.30	-0.14
$\sigma(r^s)$	8.26	7.93	0.56	7.10	1.97
$\mathbb{E}(PD)$	148.99	135.73	0.87	142.45	0.43
$\sigma(PD)$	64.07	66.57	-0.21	73.43	-0.77
$corr(PD_t, PD_{t-1})$	0.95	0.96	-0.15	0.97	-0.22
c_5^1	-0.0027	-0.0056	1.70	-0.0059	1.83
R_5^2	0.14	0.36	-1.29	0.33	-1.12
$\mathbb{E}(\mathrm{v})$	18.96	19.60	-0.87	20.10	-1.55
$\sigma(\mathrm{v})$	1.80	2.30	-0.94	2.72	-1.74
$corr(PD_t, \tau_t)$	-0.64	-0.69	0.64	-0.57	-0.86
$corr(\mathrm{v}_t, \tau_t)$	-0.75	-0.62	-1.01	-0.67	-0.61
Estimated parameters					
$\hat{\delta}$		1.0000		1.0000	
$\hat{\gamma}$		0.9105		1.0497	
$\hat{\sigma}_w^2$		4.05e-07		3.61e-07	
$\hat{A}$		1.0304		1.01858	
$\hat{PD}^L$		237.78		292.85	
$\hat{PD}^U$		338.14		332.78	

4.5.2. Steady State Kalman Filter

In the baseline specification the Kalman algorithm was used to predict the expectations coefficient β_t, which was observed. Then, the algorithm used a time-varying gain to update the beliefs. The Kalman gain was computed by using the time-varying variance of the coefficient. Alternativelyl, it is known that the Riccati recursion for σ_t^2 converges to its steady-state value σ^{2*} provided the vector of measurement and state coefficients $(A, 1)$ is observable and the vector of state coefficients and the variance matrix of state error $(1, \sigma_w^2)$ is controllable. These conditions are immediately verified in this univariate case. Therefore, the steady state variance of the unobserved state can be

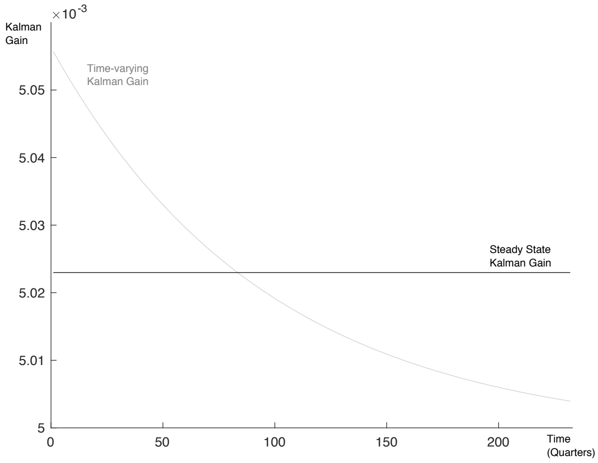

Figure 7: ***Time-varying vs. steady state Kalman gain.*** *The graph plots the estimated Kalman gain for both the baseline estimation and the steady state Kalman filter version, using the estimated coefficient from Table 3.*

obtained by solving the following discrete-time Algebraic Riccati Equation

$$\sigma_{ss}^2 = \sigma_{ss}^2 + \sigma_w^2 - \frac{(A\sigma_{ss}^2)^2}{A^2\sigma_{ss}^2 + \sigma_v^2} \tag{40}$$

Since all quantities are scalars, equation (40) can be analytically solved. Thus, the steady state variance becomes

$$\sigma_{ss}^2 = \frac{A^2\sigma_w^2 \pm \sqrt{A^4\sigma_w^4 + 4A^2\sigma_v^2\sigma_w^2}}{2A^2} \tag{41}$$

From that, the steady state Kalman gain reads as

$$k_{ss} = \frac{A\sigma_{ss}^2}{A^2\sigma_{ss}^2 + \sigma_v^2} \tag{42}$$

Altogether, the steady state Kalman filter gives rise to a constant gain learning scheme. Figure 7 compares the evolution of the estimated Kalman gain in and out of the steady state. It turns out that the steady state gain is a sort of average of the non-steady gain. Furthermore, the fourth and fifth columns of Table 3 reports the estimation results for the constant gain learning case. The model continues to perform well.

4.6. Historical Test

This section test the model ability to match the particular time trajectory followed by the long run volatility along the 1954-2012 period. So far, the test has consisted in running t-tests for the set of simulated statistics generated by a Montecarlo experiment over consumption and dividends shocks. Instead, now the model is asked to replicate both the statistics and the trajectory of v by using the particular realization the was observed for the 1954-2012 period.

Specifically, the model is fed with the observed time series of dividend and consumption growth. In other words, the observed series of shocks is used rather than a number of parallel random realizations[61]. Besides, the mean dividend growth (parameter a) is restricted to be at their historical mean. Dividend and consumption growth determine price growth and beliefs dynamics, entering equations (25) and (26). Apart from that, the model works exactly as before, with beliefs and prices being endogenously determined.

The model is re-estimated as to match the set of targeted statistics. Additionally, the criterion function has been modified with a penalization term standing for the square distance between the simulated and observed time series of the long run volatility too. In this way, the optimization program select parameters not only to pass the t-tests but also to replicate the time trajectory of v. The results are reported in Table 4. As observed, the model generates statistics very close to the observed ones. The only exception is the stock returns standard deviation, which lacks a bit of volatility. Besides, it is worth mentioning that the estimated upper bound for the PD ratio PD^U is 356.70 while the observed one is 356.31. The fact that the two of them are basically identical seems to endorse the projection facility used along the quantitative analysis.

Furthermore, as observed in figure 8, the historical model manages to generate a long run volatility in line with the observed one. Particularly, it replicate fact that the long run volatility only went up from mid 1990s on despite there was a first round of tax cuts in the early 1980s. In Section 2, this mismatch was mostly explained by the level effect coming from the non-linear tax-volatility relationship but also by the structural change in the coefficients governing that relationship. Now it is clear that the model is able to produce it without any further ingredient[62].

These results strongly contrast with the poor performance of the RE model[63], which completely

[61] Note that this was the approach used for stochastic taxes in the baseline model: instead of simulating tax shocks, the particular observed history of shocks was used. Now the historical approach is used for the other exogenous variables.

[62] The structural break in the coefficients λ_0, λ_1 characterizing the log-log relationship between taxes and volatility would mechanically be replicated by the historical model given the good fit of the v trajectory and the use of the observed tax history.

[63] The RE model is not flexible enough to incorporate directly neither dividends nor price growth. The only thing

Table 4: ***Estimation Results for the Historical test.*** *This table reports observed and simulated statistics and t-statistics (with the null of the equality between observed and simulated statistics) for the Internally Rational model estimated by using the observed dividends and consumption shocks history and the historical mean dividend growth. Furthermore, it reported the estimated parameter vector $\hat{\theta}$.*

	US data	IR model with Historical Data	
	$\hat{\mathcal{S}}_i$	$\mathcal{S}_i(\hat{\theta})$	t-stat
$\mathbb{E}(D_t/D_{t-1})$	1.0037	1.0037	0.00
$\mathbb{E}(r^s)$	1.64	1.89	-0.44
$\mathbb{E}(r^b)$	0.28	0.43	-1.39
$\sigma(r^s)$	8.26	6.16	3.57
$\mathbb{E}(PD)$	148.99	131.88	1.12
$\sigma(PD)$	64.07	79.65	-1.28
$corr(PD_t, PD_{t-1})$	0.95	0.98	-0.56
c_5^1	-0.0027	-0.0061	1.96
R_5^2	0.14	0.20	-0.35
$\mathbb{E}(\mathrm{v})$	18.96	18.48	0.66
$\sigma(\mathrm{v})$	1.80	2.12	-0.61
$corr(PD_t, \tau_t)$	-0.64	-0.71	0.95
$corr(\mathrm{v}_t, \tau_t)$	-0.75	-0.79	0.33
Estimated parameters			
$\hat{\delta}$		1.0000	
$\hat{\gamma}$		1.03205	
$\hat{\sigma}_w^2$		1.0266e-07	
$\hat{A}$		1.1687	
$\hat{PD}^L$		348.74	
$\hat{PD}^U$		356.70	

misses both the scale and timing of the long run volatility dynamics. This can be appreciated further by looking at empirical and model-implied IRFs (right graph on figure 8). The RE-implied volatility response to a tax shock is barely distinguishable from zero in spite of the latter being uniquely determined by taxes. Altogether, it signals the good empirical performance of the learning model and the relevance of the indirect tax level effect that comes up with it.

that can be done is to restrict the mean dividend growth a at their observed value and re-estimate the model. The v time series in Figure xx comes from that estimation. However, this v is indistinguishable from the one obtained using the baseline estimated parameters.

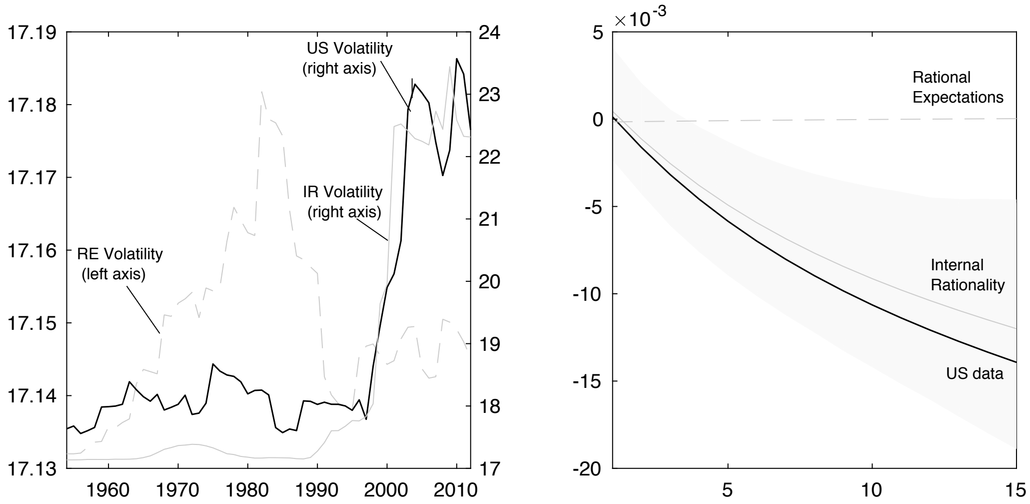

Figure 8: ***Historical fit of the models.*** *The graph on the left plots the volatility time series for the US data, the IR historical model and the RE model with the historical mean growth. The graph on the right plots the correspondent Impulse-Response Functions. The gray area indicates bootstrap confidence bands at 68% confidence with 1000 repetitions are computed using the bootstrap after bootstrap method of Kilian et al. (1998).*

5.- Extension: Heterogeneous taxes

In this section, the synthetic capital tax rate that levies on dividends and realized capital gains likewise is replaced by two rates, each one taxing one source of stock income differently. In other words, the previous model is generalized, bringing it closer to the existent fiscal system[64]. First, effective rates are computed, highlighting the different trajectories followed by each one. Then, the analytical results are revisited, pointing out the different role of each tax. Finally, the results of the estimated model are reported.

As before, effective average marginal rates for each source of capital income are considered, being defined as

$$\tau_t^D = \tau_t^d(1-\eta_t) \tag{43}$$

for the effective dividends tax rate and

$$\tau_t^{KG} = (\chi\tau_t^{skg} + (1-\chi)\tau_t^{lkg})(1-\eta_t) \tag{44}$$

for the realized capital gains tax rate, where χ stands for the fraction of short over total capital

[64]Even though the distinction between long and short capital gains is still ruled out.

gains; the other variables are as defined in Section 2. Figure 9 graphs their evolution. Dividends tax rates exhibit a trajectory fairly similar to the overall capital tax while the capital gains tax experienced larger ups and downs, ending at a level 4 times lower than its maximum reached in 1972. Consequently, the relationship of both rates with the price-dividend volatility is significantly negative but the one between capital gains taxes and volatility turned almost flat before the 1980s.

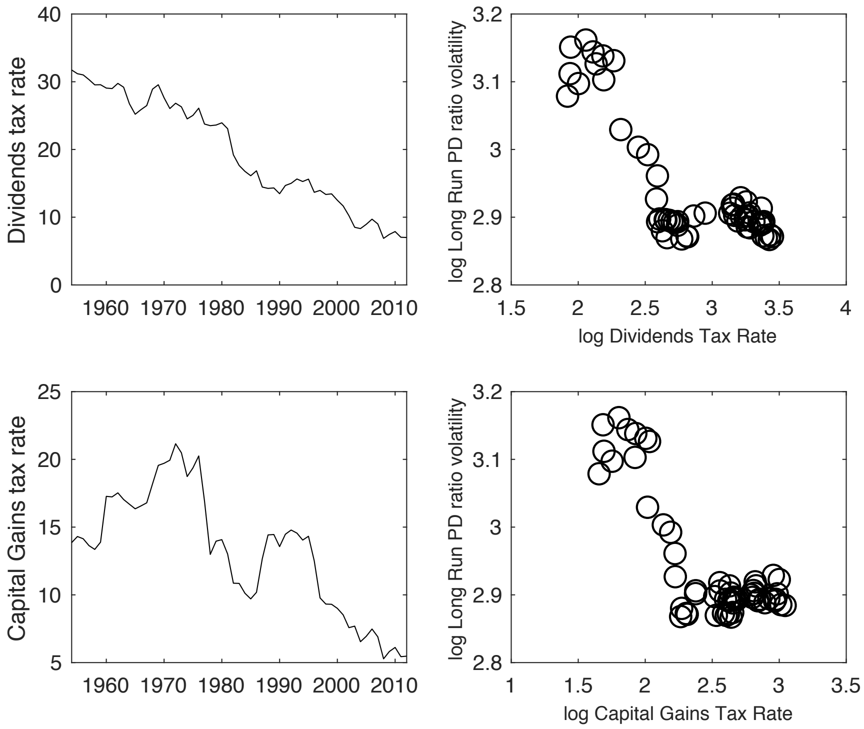

Figure 9: ***Dividends and Capital gains taxes time series and relation with volatility.*** *The top (bottom) panel pictures the time series of the effective average marginal tax rate on dividends (capital gains) and its comovement with price-dividend long run volatility. Data sample is 1954-2012 for the US, annual frequency.*

The distinction between dividends and capital gains taxes slightly modifies the analytical results. Thus, the equilibrium price-dividend ratio becomes

$$\frac{P_t}{D_t} = \frac{\delta a^{1-\gamma} \rho z(\tau_t^{KG})(1-\tau_t^D)}{1-\delta \beta_t^p z(\tau_t^{KG})(1-\pi \tau_t^{KG})} \tag{45}$$

where β_t^p can be replaced by its RE or IR version.

The role of taxes in the price-dividend ratio booms and busts is exactly as described in Section 3: tax changes triggers a transitory effect on volatility regardless expectations modelling while the tax level is non-neutral only under Internal Rationality. The only difference is that dividends and capital gains tax levels have a hierarchy in terms of their power to determine the degree of volatility, as showed in the following result:

Result 4: The capital gains tax level is more important than the dividend tax level to regulate volatility.

Under Internal Rationality, the indirect effect of permanent tax changes on the dividends and capital gains rate is permanent. Hence, dividends and capital gains tax levels jointly determine the scale of price-dividend volatility, that is, $\{v_t(\tau^D, \tau^{KG})\}$. *However, the expectations elasticity of stock prices only depends on the capital gains tax level, that is,*

$$\frac{\partial P_t/P_t}{\partial \beta_t/\beta_t} \neq f(\tau^D)$$

This is true irregardless of agents learning only about capital gains or dividends too.

Proof. The first part of the proposition follows from a Taylor first order condition around the RE value. Proceeding as in *Result 3*,

$$\mathbb{V}ar\left[\frac{P}{D}\Big(\beta_t\Big)\right] \approx \underbrace{\mathbb{V}ar\left[\frac{P}{D}\Big(a^{1-\gamma}\rho\Big)\right]}_{\text{RE var. } = 0} + \underbrace{\omega(\tau^D, \tau^{KG})}_{\text{Tax wedge}} \times \underbrace{\mathbb{V}ar(\beta_t)}_{\text{Non-fundamental var.}}$$

where $\omega(\tau^D, \tau^{KG})$ is defined as in *Result 3*. Thus, both rates determine the extend to which beliefs volatility translate into price fluctuations.

However, if the previous result is conditioned on the variables level, as an elasticity does, the role of the dividend tax level disappears. The beliefs elasticity of prices is given by

$$\frac{\partial P_t/P_t}{\partial \beta_t/\beta_t} = \frac{\delta z(\tau^{KG})(1-\pi\tau^{KG})\beta_t}{1-\delta z(\tau^{KG})(1-\pi\tau^{KG})\beta_t} \equiv \varepsilon^p_{\beta,t} = \varepsilon^p_\beta(\tau^{KG}, \beta_t)$$

This is true irregardless of agents learning or not about dividends. If they do, the dividends beliefs elasticity of prices is just 1 $\left(\text{i.e., } \frac{\partial P_t/P_t}{\partial \beta^D_t/\beta^D_t} = 1\right)$, without any dependence on dividend taxes whatsoever. ■

The results of the estimation with heterogeneous taxes are in Table 5. Again, the model manages to perform well in passing all the t-tests. Although it passes the individual significance test for correlations involving capital gains taxes, those ones shows bit too weak. That happens because of the time trajectory of capital gains taxes, that is, there are some periods in which they went up and it does not translate into lower volatility in the simulations, specially in the second tax raise in the mid 1980s.

Table 5: ***Estimation Results for the model with heterogeneous stochastic taxes.*** *This table reports observed and simulated statistics and t-statistics (with the null of the equality between observed and simulated statistics) for the Internally Rational model. Furthermore, it reported the estimated parameter vector $\hat{\theta}$.*

	US data	IR model with τ^D, τ^{KG}	
	$\hat{\mathcal{S}}_i$	$\mathcal{S}_i(\hat{\theta})$	t-stat
$\mathbb{E}(D_t/D_{t-1})$	1.0037	1.0027	0.47
$\mathbb{E}(r^s)$	1.64	2.01	-0.67
$\mathbb{E}(r^b)$	0.28	0.27	0.11
$\sigma(r^s)$	8.26	7.13	1.92
$\mathbb{E}(PD)$	148.99	137.43	0.75
$\sigma(PD)$	64.07	75.65	-0.95
$corr(PD_t, PD_{t-1})$	0.95	0.97	-0.27
c_5^1	-0.0027	-0.0062	2.04
R_5^2	0.14	0.37	-1.36
$\mathbb{E}(\mathrm{v})$	18.96	19.92	-1.31
$\sigma(\mathrm{v})$	1.80	2.59	-1.49
$corr(PD_t, \tau_t^D)$	-0.63	-0.68	0.67
$corr(\mathrm{v}_t, \tau_t^D)$	-0.71	-0.69	-0.13
$corr(PD_t, \tau_t^{KG})$	-0.58	-0.51	-0.71
$corr(\mathrm{v}_t, \tau_t^{KG})$	-0.73	-0.55	-1.29
Estimated parameters			
$\hat{\delta}$		1.0000	
$\hat{\gamma}$		0.9245	
$\hat{\sigma}_w^2$		4.10e-07	
$\hat{A}$		1.0305	
$\hat{PD}^L$		284.78	
$\hat{PD}^U$		319.72	

6.- Conclusions

This article shows that capital taxes play a crucial role not only in determining the stock price level but also its low-frequency volatility. In other words, it raises a tax theory of the aggregate long run financial volatility.

The hypothesis was backed with both empirical and theoretical evidence. First, using US data for the 1954-2012 period, it has been shown that capital taxes are a primary candidate to explain the sharp and permanent rise in long run price-dividend volatility experienced since the 1990s. Second, taxes have appeared as a powerful determinant of stock prices in a general equilibrium model. On the one hand, stochastic capital taxes are unambiguously negatively related to asset valuations. A corollary of this monotonic relationship is that tax volatility capitalizes into stock price volatility. However, this tax fundamental channel gives rise to a mere transitory response of volatility, completely unable to explain the empirical evidence.

It has been shown, though, that a new channel comes up whenever rational investors face uncertainty about how fundamentals map into prices. In this case, the tax level becomes a wedge between non-fundamental volatility and price fluctuations. In other words, lower taxes increase the sensitivity of prices to beliefs cycles, giving rise to a higher degree of volatility that persists even when taxes stabilizes at this lower level. Indeed, the quantitative evaluation of the model with the two channels reveals that it is indeed able to replicate a large set of asset pricing facts well. As opposed to that, the same model with Rational Expectations -and then, without the second channel- performs poorly.

One of the dimensions at which the model with stochastic taxes and rational learning investors excels is in generating a high enough equity premium with low risk aversion, low and stable risk-free rate, high stock returns volatility and realistic consumption and dividend processes. The reason is twofold. On the one hand, the model matches well the first and second moments of the price-dividend ratio, due to the volatility induced by investors' beliefs fluctuations. On top of it, the decline in taxes gives rise to a trended PD ratio -in line with the empirical observation- that naturally increases the stock returns means. The trended price-dividend ratio resulting from the decline in taxes strongly contrasts with the trendless ratio obtained by many asset pricing models. Importantly, these two factors are (almost) orthogonal to the determination of the risk-free rate, avoiding a too high or too volatile rate.

The empirical and theoretical arguments put forward here are important for policy debates about financial stability. In the context of the model, high tax rates make stock booms and busts less likely. In other words, capital taxes act as stabilizers of asset price cycles. This opens the door to explore the optimal use of capital income taxes to reduce asset price volatility. This possibility seems particularly relevant given the evidence about the lack of effectiveness of a financial transactions tax in doing this job.

References

Adam, K. (2003). Learning and equilibrium selection in a monetary overlapping generations model with sticky prices. *The Review of Economic Studies*, 70(4):887–907.

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.

Ang, A. and Bekaert, G. (2007). Stock return predictability: Is it there? *The Review of Financial Studies*, 20(3):651–707.

Asgharian, H., Hou, A. J., and Javed, F. (2013). The importance of the macroeconomic variables in forecasting stock return variance: A garch-midas approach. *Journal of Forecasting*, 32(7):600–612.

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.

Barro, R. J. and Sahasakul, C. (1983). Measuring the average marginal tax rate from the individual income tax.

Beltratti, A. and Morana, C. (2006). Breaks and persistency: macroeconomic causes of stock market volatility. *Journal of econometrics*, 131(1-2):151–177.

Bernanke, B. (2004). The great moderation. *Washington, DC*.

Blanchard, O. and Simon, J. (2001). The long and large decline in us output volatility. *Brookings papers on economic activity*, 2001(1):135–174.

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*.

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.

Campbell, S. D. and Diebold, F. X. (2009). Stock returns and expected business conditions: Half a century of direct evidence. *Journal of Business & Economic Statistics*, 27(2):266–278.

Christiansen, C., Schmeling, M., and Schrimpf, A. (2012). A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics*, 27(6):956–977.

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.

Conrad, C. and Loch, K. (2015). Anticipating long-term stock market volatility. *Journal of Applied Econometrics*, 30(7):1090–1114.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.

Corradi, V., Distaso, W., and Mele, A. (2012). Macroeconomic determinants of stock market volatility and volatility risk-premiums. *Swiss Finance Institute Research Paper*, (12-18).

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.

Diebold, F. X. and Yilmaz, K. (2008). Macroeconomic volatility and stock market volatility, worldwide. Technical report, National Bureau of Economic Research.

Duffie, D. and Singleton, K. J. (1993). Simulated moments estimation of markov models of asset prices. *Econometrica*, 61(4):929–952.

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*.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

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

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

Errunza, V. and Hogan, K. (1998). Macroeconomic determinants of european stock market volatility. *European Financial Management*, 4(3):361–377.

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.

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

Iraola, M. A. and Santos, M. S. (2017). Asset price volatility, price markups, and macroeconomic fluctuations. *Journal of Monetary Economics*, 90:84–98.

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

Kilian, L. et al. (1998). Pitfalls in constructing bootstrap confidence intervals for asymptotically pivotal statistics. Technical report.

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

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

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.

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

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

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

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419.

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.

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

Minsky, H. P. (1986). Stabilizing an unstable economy.

Officer, R. R. (1973). The variability of the market factor of the new york stock exchange. *the Journal of Business*, 46(3):434–453.

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

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

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

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

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

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

Sims, E. (2011). Graduate macro theory ii: Notes on time series. *Rap. tech., University of Notre Dame, URL http://www3. nd. edu/˜ esims1/time_series_notes. pdf.*

Stock, J. H. and Watson, M. W. (2002). Has the business cycle changed and why? *NBER macroeconomics annual*, 17:159–218.

Taylor, S. (1986). Modeling financial time series. *John Wiley & Sons. Great Britain.*

Appendix A: Data sources

Appendix B: Empirical volatility measure

(...)