Banking Deregulation, Entrepreneurial Risk, and the Predictability of US Stock Returns

Mathias Hoffmann, University of Zurich & CESifo[1,2]

This version: October 2011.

[1]Discussions with Sascha Becker, Monique Ebell, Nils Gottfries, Jim Nason, Michael Moore and Bent Sørensen have contributed substantially to shaping the ideas put forward in this paper. I would also like to thank two anonymous referees, seminar participants at Humboldt-University Berlin, Uppsala University, the Deutsche Bundesbank, the 2005 Society for Computational Economics Conference and at the CESifo Money, Macro and International Finance Conference 2006 for their comments and Jan-Egbert Sturm for a very helpful discussion of the paper.

[2]E-Mail:mathias.hoffmann@econ.uzh.ch Web: http://www.econ.uzh.ch/itf Postal address: University of Zurich, Department of Economics, Zuerichbergstrasse 14, CH-8032 Zurich.

Abstract

The owners of small noncorporate businesses face substantial and largely uninsurable entrepreneurial risk. They are also an important group of stock owners. This paper explores the role of entrepreneurial risk in explaining time variation in expected US stock returns in the period 1952–2010. It proposes an entrepreneurial distress factor that is based on a cointegrating relationship between aggregate consumption and income from proprietary and nonproprietary wealth. This factor, referred to here as the cpy residual, signals when proprietary income is low in relation to aggregate consumption and other forms of income in the economy. It is highly correlated with cross-sectional measures of idiosyncratic entrepreneurial and default risk, and it has considerable forecasting power for the expected equity premium. In line with the theoretical mechanism, the correlation between cpy and the stock market started to decline at the beginning of the 1980s, mainly because entrepreneurial risk became more easily diversifiable in the wake of US state-level bank deregulation.

KEYWORDS: UNINSURABLE BACKGROUND RISK, ENTREPRENEURIAL INCOME, EQUITY RISK PREMIUM, LONG-HORIZON PREDICTABILITY, STATE-LEVEL BANKING DEREGULATION, STOCK MARKET PARTICIPATION

JEL SUBJECT AREAS: E 21, E31, G12

1 Introduction

Households that bear substantial entrepreneurial risk in the form of private equity also hold a relatively large share of their wealth in the stock market. In a seminal paper, Heaton and Lucas (2000*a*) have shown that this empirical observation may single out entrepreneurs – the proprietors of noncorporate businesses – as a group of stock owners that may be particularly interesting from an asset pricing perspective.

The theoretical mechanism that links entrepreneurial risk to stock markets has two key elements. The first is limited participation in asset markets: a subgroup of the population bears most stock market risk. Indeed, over most of the postwar period, most US households did not hold any common stock, and equity ownership remains concentrated among wealthy, often entrepreneur, households. Mankiw and Zeldes (1991) showed that stockholders do indeed have much more volatile consumption, which may help solve the equity premium puzzle. Vissing-Jørgensen (2002) and Vissing-Jørgensen and Attanasio (2003) study the implications of limited participation for the size of the equity premium both theoretically and empirically. Polkovnichenko (2004) concludes that limited participation alone can only partially resolve the equity premium puzzle.

The second element is nondiversifiable idiosyncratic risk (Constantinides and Duffie (1996) and Heaton and Lucas (2000*b*,*a*)). Proprietary entrepreneurial activity is a prime example of such a risk: noncorporate businesses typically have no direct access to capital markets and small business access to bank credit is likely to be limited, in particular during recessions (see Gertler and Gilchrist (1994) and Hoffmann and Shcherbakova-Stewen (forthcoming)). This suggests that stocks are a bad hedge against shortfalls in business cash flow, illiquidity or even bankruptcy.[1] To the extent that entrepreneurs are an important group of stock holders, they will therefore require high expected returns in bad times.

This paper examines the empirical relevance of entrepreneurial risk in explaining stock returns *over time*, and it focuses on how the role of this mechanism has changed over the postwar period. I propose an entrepreneurial distress factor that is easily constructed from aggregate time series and therefore available over long time periods. This distress factor – that I call cpy – is the residual of a cointegrating relationship between consumption (c), proprietors' income (p) and other (labor) income (y) in the economy. Because it is derived from the log-linearization of the average household's budget constraint, cpy is based on minimal theoretical assumptions. Its economic interpretation is straightforward: When cpy is high, proprietary income is low in relation to other income and aggregate consumption in the economy and the average small business is relatively likely to face hard times. Consistent with this interpretation, cpy correlates negatively with cross-sectional measures of entrepreneurial risk and positively with the default spread. At the same time, cpy also has considerable predictive power for excess returns in the US stock market. During the first half of my sample, cpy easily outperforms a range of standard predictors of (excess) stock returns such as the dividend–price ratio and the payout ratio. It does so in and out of sample and at relatively short horizons. In fact, in terms of forecasting power, it compares favorably to cay, the approximation of the consumption–

[1]These risks are likely to be huge from the perspective of the average entrepreneur: Moskowitz and Vissing-Jørgensen (2002) estimate that 75 percent of private equity is held by households for which it accounts for at least half of their total net worth. Heaton and Lucas (2000*a*) highlight the empirical role of entrepreneurial background risk for stock returns by demonstrating that fluctuations in aggregate proprietary income help explain the *cross section* of stock returns in the period 1963–90.

wealth ratio suggested by Lettau and Ludvigson (2001, 2004).[2]

However, I also show that the link between cpy and excess returns on the US stock market is considerably weaker in the second half of my sample period. This decline in the correlation between cpy and stock market returns can be attributed to two developments that permanently affected the two key elements of the entrepreneurial risk mechanism that I discussed above. First, with the advent of employer-sponsored pension plans, stock ownership has widened to households with little or no entrepreneurial risk.[3] Second – and empirically more important – the deregulation of bank branching restrictions in US federal states during the 1980s has hugely facilitated credit market access for small firms, in particular during recessions.[4] My results in this paper are consistent with the view that better access to credit has made it less likely that entrepreneurs have to sell public stock in order to provide cash-flow for their business during aggregate downturns. This has severed the link between entrepreneurial risk and the stock market.

My empirical framework builds on the studies by Lettau and Ludvigson (2001, 2004), who have shown that an empirical approximation of the consumption–wealth ratio, the cay residual, has considerable forecasting power for excess equity returns. Like Lettau and Ludvigson, this paper pro-

[2]Clearly, because cpy measures cyclical proprietary income, it could be correlated with stock returns for reasons that have nothing to do with the uninsurable background risk mechanism that motivates the analysis here: profits by corporate and noncorporate businesses may comove over the business cycle. As financing conditions vary over the business cycle, so could corporate retained earnings (Gertler and Hubbard (1993)). However, the predictive power of cpy for excess returns on stocks far exceeds what can be explained by its covariation with corporate dividends and earnings – cpy seems related to variation in the discount factor, not to stock market cash flow.

[3]See Poterba (1994).

[4]Jayaratne and Strahan (1996) show that state-level banking deregulation lead to higher growth. Demyanyk, Ostergaard and Sørensen (2007) show that banking deregulation improved income insurance. Hoffmann and Shcherbakova-Stewen (forthcoming) show that inter-state consumption risk sharing improved in particular during recessions and this effect is present mainly in states with many small businesses. Park (forthcoming) finds that banks increased loan commitments following banking deregulation.

vides a consumption-based predictor of stock returns that is derived from the log-linearization of an intertemporal budget constraint. However, my indicator differs importantly from Lettau and Ludvigson in that it does not directly involve financial variables. Rather, cpy predicts stock markets from a linear combination of real flows. Also, while the results obtained by Lettau and Ludvigson are consistent with a wide class of theoretical models, they are silent about a particular theoretical explanation for the predictability of the equity premium. My interpretation of cpy as an entrepreneurial distress factor puts me in a position to explore empirically the extent to which the entrepreneurial risk mechanism discussed here can help explain *why* expected stock returns vary over time.[5,6]

The motivation behind cpy follows the logic of most current theories of consumption, including the permanent income hypothesis:[7] households' desire to smooth consumption implies that transitory shocks to income will not affect consumption very much. Conversely, permanent shocks to income cannot be smoothed and will affect consumption and income to an about equal extent. The consumption–income ratio should therefore signal bad and good times by indicating whether income is temporarily high or low in relation to the stochastic trend defined by consumption.

[5] I emphasize that it is not the purpose of this paper to suggest cpy as a new 'star' predictor variable. Since cpy is constructed from real flows whereas cay involves the stock of financial assets, it is almost necessarily the case that cay is the better predictor. However, as suggested above, the correlation between cpy and the stock market — and how it changes — is informative with respect to the underlying mechanisms that drive expected stock returns.

[6] My results are consistent with, and in fact strongly suggestive of a risk-based mechanism. However, they do not rule out that there are alternative 'irrational' mechanisms that explain why expected returns vary over time. Also, stock ownership by entrepreneur households could itself be irrational. The extreme exposure of entrepreneurs to (public and private) equity remains a puzzle as do the low returns on private equity (see Moskowitz and Vissing-Jørgensen (2002)).

[7] See Cochrane (1994) and Lettau and Ludvigson (2004) for empirical implementations.

Consistently with theory, I find movements in consumption to be largely permanent. Hence, cpy does indeed signal transitory variation in incomes. Still, this leaves open the question of how cpy, as a linear combination of *aggregate* variables, can be useful as a distress factor for a subgroup of the population – the owners of small businesses. The reason for this is empirical: the variation in cpy is in fact largely due to proprietary income, whereas labor income itself is close to a random walk. Hence, cpy signals when proprietary income is low in relation to the common trends that it shares with aggregate consumption and labor income. It is this feature of the data that allows me to refer to cpy as the *entrepreneurial consumption–income ratio* and that provides the foundation for its interpretion as a distress factor.[8]

The remainder of this paper is structured as follows. In the next section, I derive cpy and further motivate its interpretation as a proxy of entrepreneurial risk. In section three, I identify cpy in the data, show that it mainly captures variation in proprietary income and is indeed linked to measures of default and entrepreneurial idiosyncratic risk. Section four shows that, during the first half of the sample, cpy predicts fluctuations in the equity premium. In section five, I demonstrate that the correlation between cpy and aggregate stock markets has decreased due to widened stock market participation and better access of small firms to credit following banking deregulation. A final section summarizes and concludes.

[8] To underpin this interpretation theoretically, assume there are two types of households: entrepreneurs who receive proprietary income p, and workers who receive labor income, y. For each group, the log-consumption–income ratio should signal temporary fluctuations in their respective income streams, following the same intuition that was discussed for the aggregate consumption–income ratio above. While we cannot observe these consumption–income ratios separately – consumption time series for each of these groups do not exist – I show in section 2.2 that cpy approximately equals a weighted average of the consumption–income ratios of proprietors and workers. Specifically, if the temporary component in proprietary income is large relative to that in labor income and aggregate consumption – as is the case in the data – cpy will mainly reflect variation in entrepreneurs' consumption–income ratio.

2 The entrepreneurial consumption–income ratio

The analysis in this paper uses almost 60 years of *aggregate* time-series data to construct an entrepreneurial risk or 'distress' factor, cpy. My approach, which extends Lettau and Ludvigson (2001, 2004) and Campbell and Mankiw (1989), has the advantage that it rests on minimal theoretical identifying assumptions because it is based solely on the log-linearization of the average household's intertemporal budget constraint. The log-linearization allows me to derive a cointegrating relationship between the logarithms of consumption (c) and proprietary (p) and other forms of income (y); cpy is the residual from this cointegrating relationship. This section first derives cpy and then discusses the conditions under which it can be interpreted as an entrepreneurial risk factor.

2.1 A long-run relation between consumption and the dividends from wealth

For the purposes of this paper, I find it useful to state the budget constraint of the average household in present-value form. I write total wealth, Ψ_t, as the sum of proprietary wealth Π_t and other forms of wealth Θ_t:

$$\Psi_t = \Pi_t + \Theta_t \tag{1}$$

where total wealth Ψ_t is the present value of consumption expenditures and the right hand side gives the present value of all incomes, partitioned according to whether this income is derived from proprietary entrepreneurial or nonproprietary activity (wages and salaries). Letting lowercase letters denote logarithms, this identity can be rewritten as

$$\log\left(1-\frac{\Pi_t}{\Psi_t}\right)=\theta_t-\psi_t. \tag{2}$$

The share of proprietary wealth in total wealth is $\Pi_t/\Psi_t=\exp(\pi_t-\psi_t)$, and I denote the long-run mean of Π_t/Ψ_t with γ. Hence, I can write $\gamma = \exp(\overline{\pi-\psi})$, where $\overline{\pi-\psi}$ is the logarithm of the long-run mean of Π_t/Ψ_t. I now expand the left-hand side of (2) around $\overline{\pi-\psi}$ to obtain

$$\log\left(1-\frac{\Pi_t}{\Psi_t}\right)\approx\kappa-\frac{\gamma}{1-\gamma}\left[\pi_t-\psi_t\right]$$

where $\kappa=\log(1-\gamma)-\gamma\left(\overline{\pi-\psi}\right)(1-\gamma)^{-1}$ is a constant. Plugging this back into (2) and rearranging yields

$$\psi_t-\gamma\pi_t-(1-\gamma)\theta_t=-(1-\gamma)\kappa. \tag{3}$$

This equation is the basis for the long-run relationship between consumption and proprietary income that I am going to consider in this paper. The logarithms of total, proprietary and nonproprietary wealth are not directly observable. However, the long-run relation between wealth and its subcomponents can be made observable by acknowledging that proprietary income is the dividend to proprietary wealth as any other form of income must be the dividend from other (nonproprietary) forms of wealth. Ultimately, consumption is the dividend from total wealth. The gist of my argument is that consumption, proprietary income and labor income are all individually integrated of order one ($I(1)$), but all three variables differ from their respective permanent values (i.e., from the respective wealth components) only by an $I(0)$-process. Hence, replacing ψ, π, and θ in (3) with consumption (c), proprietary (p) and nonproprietary income (y) respectively, one must obtain an $I(0)$ process, so that

$$cpy_t = c_t - \gamma p_t - (1-\gamma) y_t \tag{4}$$

defines a cointegrating relationship. This cointegrating relationship is more formally derived in the appendix. By analogy to Lettau and Ludvigson, I refer to it by the abbreviation 'cpy'. An immediate empirical implication of (4) is that the coefficients of the cointegrating vector should correspond to the long-run shares of proprietary and other wealth in total wealth. I return to a discussion of this implication below.

By Granger's representation theorem, the fact that c, p, and y cointegrate also implies that their joint dynamics are captured by a vector error-correction mechanism. Stacking the three variables so that $\mathbf{x}_t = \begin{bmatrix} c_t & p_t & y_t \end{bmatrix}'$, one can then write the vector error-correction model (VECM) as

$$\mathbf{\Gamma}(L)\mathbf{\Delta x}_t = \boldsymbol{\alpha\beta}'\mathbf{x}_{t-1} + \varepsilon_t$$

where $\boldsymbol{\beta}' = \begin{bmatrix} 1 & -\gamma & 1-\gamma \end{bmatrix}$ is the cointegrating vector, $\boldsymbol{\alpha}$ is a vector of adjustment loadings, $\mathbf{\Gamma}(L)$ is a 3×3 matrix polynomial in the lag operator, $\Delta = 1 - L$ the first difference operator, and ε_t is a vector of disturbance terms. The error-correction mechanism implies that at least one of the three variables – consumption, labor, and proprietary income – has to adjust to restore cpy to its long-run mean. Hence, changes in at least one of the three variables will have to be predictable; i.e., at least one of the variables will have a statistically significant transitory component. I provide ample evidence that cpy can indeed mainly be associated with the temporary component of proprietary income.

2.2 Interpreting cpy as a proxy of entrepreneurial risk

The theoretical mechanism that I wish to investigate rests on the presence of nondiversified idiosyncratic risk at the household level. While household-level data are available from various sources, the sample period covered by these data sets at best reaches back to the early 1980s, and the data are typically at an annual frequency. To study the link between entrepreneurial risk and the stock market, however, it is clearly desirable to use long stretches of quarterly data. [9]

Therefore, I propose to use cpy as a distress factor for the owners of non-incorporated businesses. Because cpy is the linear combination of aggregate variables, it is easy to construct over long time periods. The fact that cpy is an aggregate indicator does, however, also raise the question under which conditions cpy can help identify distress risk for a subgroup of the population – entrepreneurs – and whether these conditions are fulfilled empirically. I turn to answering this question next.

Let us assume that there are just two types of households in the economy: proprietors, who only receive proprietary income, and workers, who only receive labor income. Then the budget constraints for each household type are

$$\begin{aligned} \Psi_t^p &= \Pi_t \\ \Psi_t^w &= \Theta_t \end{aligned}$$

where Ψ_t^p and Ψ_t^w are the present value of consumption of proprietors and workers respectively, and Π_t and Θ_t, as in the previous section, are the

[9]Furthermore – and the results of this paper underscore this point – the roles of nondiversified entrepreneurial income risk in explaining asset returns may have been more important in the distant past than recently, due to increased stock market participation and banking deregulation. Household-level data for the 1950s and 1960s are, however, not available.

present values of proprietary income and labor income respectively. Clearly, summing up the two group-specific budget constraints yields the aggregate constraint from the previous section, $\Psi_t = \Pi_t + \Theta_t$, where $\Psi_t = \Psi_t^p + \Psi_t^w$. Hence, the setup assumed here, with its two different household types, is nested in the aggregate framework discussed in the previous section.

In this scenario, under the assumption that both proprietors and workers attempt to smooth consumption so that their future consumption changes are not too predictable, it can be shown that

$$cpy_t \approx \gamma(c_t^p - p_t) + (1-\gamma)(c_t^w - y_t). \tag{5}$$

Here, c_t^p and c_t^w are the logarithms of proprietors' and workers' consumption respectively. See the technical appendix for details.

Equation (5) states that cpy approximately equals a weighted average of the consumption–income ratios of proprietors, $c^p - p$, and of workers, $c^w - y$. As was the case for the aggregate budget constraint above, intertemporal budget balance of each household group will impose stationarity on both $c_t^p - p_t$ and $c_t^w - y_t$. Also, if both household types smooth consumption, then $c^p - p$ will mainly capture transitory variation in p, whereas $c^w - y$ should capture transitory variation in y. Unfortunately, neither of these two type-specific consumption–income ratios is directly observable because long time series of data on proprietors' and workers' consumption do not exist. My suggestion to use cpy as a proxy for the entrepreneurial consumption–income ratio, $c_t^p - p_t$, is based on the empirical fact that cpy predominantly reflects temporary variation in p – exactly as $c^p - p$ should. In the data, both consumption, c, and labor income, y, have much smaller transitory components; i.e., are much closer to random walks. This suggests that it is indeed the variation in $c^p - p$ that drives most of the variability in cpy and explains why I refer to cpy as the *entrepreneurial* consumption–income ratio.

Hence, fluctuations in cpy identify the temporary deviation of proprietary income from its trend that, in turn, is defined by its long-run relation with aggregate consumption and labor income. My argument is now that if proprietary income is low in relation to aggregate consumption and other (i.e., nonproprietary, labor) income, then idiosyncratic risk for the average proprietor is high. When times are hard for the average small business, then its average proprietor faces a relatively high probability of having to liquidate private assets; e.g., in order to provide cash flow for the business and to avoid bankruptcy. Proprietors will be afraid of having to liquidate public equity holdings when prices are low. Therefore, to hold the outstanding stock willingly, they will have to be compensated with higher expected returns.

One key assumption underlying the identification of cpy is that the long-term share of proprietors' wealth in total wealth is constant; $\gamma = E(\Pi_t/\Psi_t)$. This assumption is analogous to the assumption in Lettau and Ludvigson (2001) that the shares of financial and human wealth are constant in the long run. I emphasize that this does not imply that the share of proprietor's wealth in aggregate wealth is actually constant in each period. In fact, this variable could move in long swings, as long as it is mean reverting. Certainly, changes in taxation or in organizational form (e.g., the trend towards S-corporations) could permanently affect the share of proprietary wealth relative to corporate dividends or labor income. Note, however, that the assumption of a reasonably stable γ is a precondition for the existence of a cointegrating relationship between consumption, proprietary and labor income. Hence, I will implicitly test this assumption when I identify a cointegrating relationship between the three variables in my empirical analysis below. Finally, theory imposes some discipline on the plausible magnitude of γ in my analysis here: proprietary wealth can be interpreted as the value of the capital stock held by nonpublicly listed firms. Under the same assumptions on technology (Cobb–Douglas) as in Lettau and Ludvigson (2001, 2004), γ should therefore approximately correspond to the long-term share of entrepreneurial capital in the economy; i.e., to the share of total capital income less dividends paid on the stock market in GDP. I further discuss this issue in my empirical analysis below.

3 Empirical implementation

3.1 Data

The data are quarterly data on personal income and its components from the US Bureau of Economic Analysis. Consumption data are from the same source. The data range is from 1952Q1 to 2010Q4. I express income and consumption in per capita terms and deflate both with the price index for personal consumption expenditure (PCE). Details on all data used in this paper and on their preparation are available in the data appendix.

An important issue in the preparation of the data used for analysis is the choice of consumption data. Most empirical asset pricing studies exclude durables consumption expenditure from the consumption concept. Because durables are consumed over several periods, total consumption is therefore likely to be a lot more variable than the true consumption stream (relevant for utility maximization), which should only include nondurables consumption and the stream of consumption services from the stock of durables. I follow this convention, and my main results are based on consumption measured as expenditure on nondurables and services, excluding shoes and clothing.[10] For robustness, I also report basic cointegrating results based on total consumption expenditure.[11]

3.2 Cointegration analysis

To identify the number of cointegrating vectors, I use Johansen's test procedures, the maximum eigenvalue and the trace test statistics. The results, provided in Table 1, clearly indicate the presence of one cointegrating relationship: the tests based on the specification with nondurables consumption (the conventional measure in the literature) are all highly significant well beyond the 95 percent level. The tests based on total consumption also signal cointegration, though somewhat less strongly, at the 90 percent level.

I estimate the cointegrating vector in two ways: first, based on the Johansen (1991) full information maximum likelihood approach, and second, based on a cointegrating regression in the spirit of Engle and Granger (1987). Again, the exercise is performed for both total consumption and nondurable consumption.

Shocks to income and consumption are likely to be correlated. Johansen's procedure implicitly takes care of this simultaneity. In the cointegrating regression, I account for it by applying the Stock and Watson (1993) dynamic OLS procedure, which amounts to adding leads and lags of first differences of the regressors.

Table 2 reports the estimated cointegrating vectors obtained from the different consumption data sets and based on both the full information and

[10] Lettau and Ludvigson (2004)show that this is justified under the assumption that true logarithmic consumption is a constant multiple of the logarithm of nondurables consumption.

[11] This is motivated by the argument put forward by Rudd and Whelan (2006), who argue that intertemporal budget balance requires the present value of *total* consumption expenditure to be equated to the present value of the dividends of wealth.

the regression-based method. The estimate of the cointegrating vector does not depend on the choice of estimation method, nor is it very sensitive to the choice of consumption data (nondurables vs. total consumption).

The estimated cointegrating vectors suggest that the present value of proprietary income amounts to about a quarter of the total present value of consumption. Lettau and Ludvigson (2004) find that household asset wealth – i.e., the present value of all household cash flow derived from capital – accounts for roughly one-third of total wealth in their cay relationship, in line with typical estimates from the real business cycle literature that put the capital share for the US economy at a very similar value. According to the National Income and Product Account (NIPA) Tables, proprietary income, rents and interest accounted on average for about three-quarters of household income from capital, with corporate dividends accounting for the last quarter. Hence an estimate for γ of around $0.25 = (3/4 \times 1/3)$ is consistent with both the evidence from US national accounting data and the findings in the earlier literature.

In the remainder of the paper, I now define cpy as the cointegrating residual

$$cpy = c_t - 0.2570 p_t - 0.7563 y_t, \tag{6}$$

which corresponds to the nondurables specification estimated based on Johansen's procedure (see Table 2).

The next section shows that cpy mainly captures transitory variation in proprietary income, justifying its interpretation as an entrepreneurial consumption–income ratio. I then provide first evidence on the link between cpy and background risk one the one hand and cpy and the stock market on the other.

3.3 cpy as the transitory component of proprietary income

The cointegrating relation between consumption and proprietary and other income allows us to identify permanent and transitory components of these variables without further identifying restrictions. To describe the joint dynamics of proprietary income, other income and consumption, I now estimate a cointegrated vector autoregression (VECM) in which I impose the cointegrating vector estimated before. I include one lagged difference of $\mathbf{x}_t$ in each equation, as suggested by standard information criteria. The results, reported in Table 3, are, however, not very sensitive to the number of lags chosen.

A first impression of the role of transitory components in explaining consumption, proprietary and other income can be gleaned from the coefficient on the lagged cointegrating residual: if this coefficient is zero, the respective variable does not contribute to the error-correction mechanism, implying that it does not contribute to the predictable dynamics that form the transitory component of the system.

Inspection of the VECM coefficients reveals that only the adjustment coefficients on proprietary income and consumption are significant. However, the coefficient on cpy in the equation for p is much bigger in absolute value, suggesting that the transitory dynamics in consumption and income is largely due to deviations of proprietary income from its long-run trend. I further examine this proposition by identifying the permanent and transitory components of c, p, and y more formally.

I do this in two ways. First, I build on recent literature inspired by Gonzalo and Granger (1995) and Proietti (1997) in which the permanent and transitory components of a cointegrated system are expressed as the linear combination of current levels. In this way, time series for the trend and cycles of c, p, and y are easily obtained. A second, alternative approach that allows me to obtain variance decompositions at different horizons, is to identify permanent and transitory shocks directly. Here, I build on Johansen (1995), Hoffmann (2001) and Gonzalo and Ng (2001). See the Technical Appendix for details of the two approaches.

Figure 1 plots the trend components of consumption, other income, and proprietary income obtained from the first approach, along with the variables themselves. As is apparent again, proprietary income is *the* variable in the system with a sizeable transitory component, whereas other components of income as well as consumption are always much closer to their random walk components. This message also transpires from the variance decompositions in Table 4. Transitory shocks play a small role in explaining consumption at short horizons. However, this component dies out very quickly. Conversely, transitory shocks explain more than 40 percent of proprietary income at short horizons and still almost 30 percent at the two-year horizon. This result is very much in line with the findings obtained by Cochrane (1994) as well as Lettau and Ludvigson (2001): both studies find consumption to be close to a random walk, whereas Cochrane finds that the consumption–income ratio predicts changes in income. The results here identify proprietary income as an important source of this predictability.

The result that proprietors' income is the component of personal income with the biggest transitory component is a first important ingredient of my interpretation of cpy as an entrepreneurial distress factor: if proprietary income is low in relation to other incomes in the economy and in relation to aggregate consumption, then times for the average entrepreneur are likely to be hard and entrepreneurial risk will be high. Conversely, levels of proprietary income above the long-run trends in income and consumption will reflect periods of low entrepreneurial risk. [12]

3.4 cpy, the stock market and background risk: first evidence

This sub-section provides graphical evidence that cpy is an indicator of fluctuations in entrepreneurial risk and that it is correlated with expected returns on the stock market. As entrepreneurial background risk has become more easily diversifiable in the wake of state-level banking deregulation and as more (nonentrepreneur) households have started to participate in the stock market (largely with the advent of employer-sponsored retirement schemes in the 1970s and 1980s), the correlation between cpy and the stock market has declined, however.

Figure 2 plots the cay residual from Lettau–Ludvigson along with the cpy residual estimated here. It is well known from the results in Lettau and Ludvigson (2001) that cay is essentially the cyclical component of the stock of financial wealth and notably of stock returns. Indeed, the two residuals share some of the major swings, but they are clearly not the same time series. Their correlation across the sample period is only 0.40. The comovement, however, seems a lot stronger in the earlier part of the sample: over the period from 1952Q1 to 1980Q4, the two time series have a correlation of 0.59. The early 1980s seem to mark a break in the correlation between cpy and cay. My suggestion is that the high correlation between cpy and the stock market in the first half of the sample is a reflection of entrepreneurial distress: entrepreneurs were the most important group of stock market par-

[12]Labor income may be insured through migration, through labor hoarding over the cycle or through unemployment benefit systems. Furthermore, there is evidence that corporations smooth dividend payments (Cochrane (1994), Lamont (1998)), implying a similar kind of smoothing for income derived from financial assets. Proprietary income, though, by its very nature is considerably harder to smooth.

ticipants, and they faced largely uninsurable risk because they could not borrow in times of recessions. As I will argue, the subsequent decline of the correlation of cpy with cay and the stock market is driven by banking deregulation and by the widening of stock market participation.

Figures 3 and 4 further buttress my interpretation of cpy as an entrepreneurial distress factor. As pointed out by Constantinides (2002), for idiosyncratic risk to explain asset returns, idiosyncratic shocks need to be persistent, and the cross-sectional variance of shocks needs to be negatively related to asset returns. Figure 3 plots cpy along with a measure of idiosyncratic entrepreneurial risk – the dispersion of proprietors' income growth across US federal states. To construct this measure, I used quarterly state-level per capita proprietary income from the Bureau of Economic Analysis from 1952Q1 to 2007Q4.[13] To capture risk at the business cycle frequency, I considered growth rates over two-year horizons (growth rates over longer or somewhat shorter horizons give very similar results, though). I then formed the idiosyncratic component of this growth rate for each state by deducting the growth rate of US-wide per capita proprietary income. The measure plotted in Figure 3 is the cross-sectional standard deviation of these state-specific growth rates for each period.

Visual inspection suggests an important link between these two time series. The correlation coefficient is 0.41, and the $t-$statistics of a regression of cpy on the cross-sectional standard deviation are higher than 6. This suggests that cpy captures an important element of the cross-sectional heterogeneity in the economic situation of proprietor-run businesses in the United States, and the correlation between this idiosyncratic risk and the stock market is negative as required by theory: if cross-sectional heterogeneity is high,

[13] After that date, no consistent state-level per capita data are as yet published by the BEA. These will only be released after they have been recalculated based on the 2010 census.

cpy is high.[14] Note also that idiosyncratic state-level proprietary income growth is highly persistent. Using the Im, Pesaran and Shin (2003) group mean panel unit root test, idiosyncratic proprietary income growth at the two-year horizon seems almost unpredictable: the average autoregressive coefficient is -0.08 with a mean t-value of -1.20. This suggests that relative state-level proprietors' income is close to a random walk for the average state.

Figure 4 plots cpy against the HP-filtered component of the yield spread between *Baa*- and *Aaa*-rated corporate bonds. This spread can be interpreted as the economy-wide price of default risk. The correlation between cpy and the default spread is 0.23 over the entire sample period, but, again, it is noticeably higher in the first half, at 0.38. Again, this pattern is explored in more detail in the last section of the paper. I note here that it is consistent with the view that banking deregulation has improved access to finance for small businesses, thus weakening the link between cpy and the price of default risk.[15]

4 Predicting stock market returns

This section explores the link between cpy and the stock market more formally, through long-horizon regressions of excess returns, earnings and dividends on cpy and a host of 'usual suspects' predictor variables. Specifically,

[14]As shown in the previous section, cpy is the transitory component of p, so that cpy is high when p is low. Hence, the positive correlation between cpy and idiosyncratic risk means that background risk is high when average proprietary income is low.

[15]I also compare cpy to quarterly bankruptcy rates obtained from the American Bankruptcy Institute at *http://www.aib.org*. These are available only from 1980. Consistent with the view that banking deregulation has mattered for small business access to credit, we see a downward trend in bankruptcy filings since the early 1980s that mirrors the narrowing of the default spread. Also, supporting my interpretation, cpy is still correlated with a coefficient of 0.23 with 4-quarter growth rates in bankruptcy filings in the post 1980 period (also see Figure A.1 in the *Technical Appendix*).

my regressions are of the form

$$x_{t+k} - x_t = \delta_k cpy_t + \phi_k' \mathbf{z}_t + u_t^k \tag{7}$$

where x_t stands in for returns, dividends or earnings, and $\mathbf{z}_t$ is a vector of predictor variables.

The distribution of standard t-values in long-horizon regressions may be considerably distorted at longer horizons: as the differencing horizon increases, the dependent variable behaves more and more like an integrated process, so that the regressions of the form (7) may signal a spurious rejection of the null of nonpredictability. Valkanov (2003) has derived an adapted statistics based on the t-value divided by the square root of the sample size that has a well-behaved limiting distribution and that can easily be simulated. I report the significance of the long-horizon regressions based on standard critical values for t-statistics as well as on Valkanov's correction.

The correlation between cay and cpy discussed in the previous section suggests that the link between the stock market and cpy should be stronger in the first half of the sample. In my analysis here, I therefore split the sample in two: the first half covers 1952Q4–1980Q4, whereas the second half covers the period 1981Q1 to 2010Q4.[16]

Table 5 provides regressions of excess returns on the Standard & Poor's (S&P) 500 index on cpy for the two subperiods. As a benchmark, for each subperiod, it also presents the results of similar regressions on Lettau and Ludvigson's cay residual and regressions in which both cpy and cay are included as regressors.

In the first half of the sample, cpy is a very potent and significant predic-

[16]As will become more apparent later on, the early 1980s saw state-level banking deregulation gathering pace. Equally, the advent of employer-sponsored retirement funds led to a huge increase in stock market participation during that same period.

tor of excess returns on the US stock market, with R^2 of 0.08 one-quarter ahead, increasing to more than 60 percent at horizons of more than four years. This compares very well with cay, and at short horizons, cpy even explains more variation in stock returns than does cay. In direct comparison, when both residuals are included in the regression, cpy absorbs most of the short-term predictability from cay at horizons below one year. Note also that all coefficients on cpy are positively signed, consistent with the uninsurable background risk story: cpy and expected returns are positively correlated, so that proprietary income below trend predicts rising expected returns. In the second half of the sample, however, this pattern of predictability seems completely absent: cpy has virtually no predictive power for excess stock returns. Conversely, cay (even though its forecasting ability also seems somewhat lower than in the first half of the period) holds up well as a predictor of stock market returns. In the following subsections, I first discuss what drives the correlation between cpy and the stock market in the first sample period. I then turn to analyzing the impact of banking deregulation and increased stock market participation on the decline of this correlation in the next section.

4.1 Comovement with corporate earnings and dividends

What can explain the performance of cpy in predicting stock markets in the first half of the sample period? Recall that cpy is essentially the transitory component of proprietary income, p. Of course, proprietary income could help to predict stock market returns for reasons that are unrelated to the background risk cum limited participation mechanism that provides the motivation for this paper. First, proprietary income could be correlated with corporate earnings over the business cycle. Second, proprietary income is the dividend from proprietary, noncorporate business wealth, so if both corporate and noncorporate firms are confronted with the same business cycle conditions, one may expect a correlation with corporate dividends. I explore both of these possibilities in turn.

The first two panels of Table 6 first repeat and extend my earlier results for long-horizon return predictability. For ease of comparison, Panel I just reproduces the results from the previous Table for the excess returns on the S&P500. Panel II provides results for the CRSP index of stock returns, which is much broader than the $S\&P500$. Again, there is significant evidence that cpy predicts excess returns, in particular at short to medium horizons.

Panel III of Table 6 then turns to whether cpy predicts business-cycle variation in earnings: though cpy is marginally significant for corporate earnings at horizons of eight quarters and beyond, this is true only for the uncorrected $t-$values but not generally based on the Valkanov correction. The adjusted R^2 measure remains very low at horizons below two years. Hence, corporate earnings are not very strongly predictable from cpy at business cycle frequencies.

The link between cpy and dividends is a bit stronger, though it cannot explain why cpy explains excess returns as well as it does: panel IV shows long-horizon regressions of dividend payments on the S&P 500. The cpy residual would appear to have predictive power for dividend growth at horizons from two years. There is a somewhat stronger link between cpy and dividends once I use the broader concept of personal dividend income from the BEA personal income tables rather than just dividends on the S&P 500 (panel V). In this set of regressions, dividends appear predictable at horizons as short as one year.[17] However, while cpy does have some predictive power for dividends at horizons from between one and two years ahead, the direct comparison with the excess-return predictability results in Panels I and II shows that the predictability of dividends cannot account for why cpy explains excess returns: first, the predictability of excess returns is, at all horizons, much stronger than that of dividends (in terms of R^2 and of the significance of the associated coefficients). Second, cpy does not have any predictive power for dividends at horizons below 1–2 years, but it *does* predict excess returns at such short horizons. The findings here therefore suggest that – in the first half of the sample considered here – cpy must have been linked to the stock market through a risk mechanism: cpy reflects variation in the discount factor of the average market participant.

4.2 cpy vs cay and financial predictors

The previous results are consistent with the view that cpy predicts stock markets because of a risk-based mechanism. This subsection explores how well cpy captures time variation in expected returns in comparison with a range of 'usual suspects' of economic and financial predictor variables. The results in Table 5 already showed that cpy absorbs a large part of the predictive power of cpy in the first half of the sample. Table 7 shows similar comparisons between cpy and other forecasting variables: the dividend–price ratio, the dividend–earnings (payout) ratio as well as the cyclical (HP-detrended) component of the T-bill rate and the default spread. I also include a variable

[17] The fact that cpy predicts *personal* dividend income so much better than dividends on the S&P 500 is at least in part likely to be due to the definition of the BEA data: personal dividend income includes dividends disbursed by nonlisted corporations, notably S-type corporations. For the purpose of our analysis here, such disbursements are certainly much closer in spirit to proprietary income than to dividends paid by stock-market-listed companies.

that I call res, the residual of the regression of cay on cpy. By construction, res is orthogonal to cpy, and taken together, cpy and res must explain expected returns at least as well as cay. In the comparison with each of the financial predictor variables, this allows me to query the extent to which the forecasting power of cay is explained by cpy and the extent to which it is explained by orthogonal factors (captured by res).

In all of these comparisons, cpy is highly significant. Conversely, res generally appears insignificant at short and business cycle horizons. The cyclical components of the T-Bill rate and the default spread are the only financial predictor variables that appear significant at short horizons (below four quarters). Furthermore, the default spread is the only variable that 'eats' somewhat into the individual significance of cpy at short horizons (vis-à-vis the specification in which cpy figures alone). This is in line with my earlier finding that the default spread is highly correlated with cpy in the first half of my sample and consistent with my interpretation of cpy as an indicator of entrepreneurial risk.[18] In the last panel of Table 7, I then let cpy compete against the cyclical components of the T-Bill rate and of the default spread and res. In this specification, res is significant but so remains cpy.

Table 8 presents the results of a forecasting exercise in which excess returns are predicted one period ahead out of sample. In the upper panel, I present forecast comparisons of a range of nested models (numbered as 1 and 2): a constant expected returns model and an AR(1) model of excess returns. Each of these models is nested into a richer model in which the constant or autoregressive term respectively figures together with cpy.

In the lower panel, I conduct pairwise nonnested comparisons between

[18]It is also consistent with my interpretation that the default spread is individually significant as a predictor variable – but mainly so in the first half of the sample period: better access to bank finance in the second half of the sample has weakened the impact of background risk on risk indicators.

cpy and a range of usual suspect forecasting variables: lagged returns (model 3), the dividend–price ratio (model 4), the dividend–earnings ratio (model 5), the default spread (model 6) and cay (model 7).

In the first two columns of each panel, I report the result for the case in which the cointegrating vector defining cpy is estimated from the whole sample. Results for the case in which the cointegrating vector is continually reestimated from within the forecasting sample are in columns 3 and 4.

Turning to the forecasting exercises in which the cointegrating vector is fixed first, we can see that the mean squared error of the model involving cpy is always smaller than the mean squared error of the alternative model. This difference is also statistically significant throughout: for the nested models, I report the McCracken (2007) out-of-sample F statistics, and for the nonnested comparisons, I report the modified Diebold–Mariano (MDM) test by Harvey, Leybourne and Newbold (1998). For both tests, the respective 90 and 95 percent critical values are provided in the last two columns of the table. In all cases, the tests are significant; at the 10 percent level for $d-p$ and the default spread, and at the 5 percent level for the other variables, including cay. This suggests that cpy adds significant out-of-sample forecasting power to the alternative model in all of these cases.

As a cointegrating residual, cpy is the deviation of proprietary income from a long-run equilibrium relationship. It is therefore preferable to estimate cpy from a long sample. Still, one may want to ask the question whether a forecaster during the first half our sample period could have used the information in c, p and y to do out-of-sample forecasts of excess returns. This is what is addressed by the forecast comparisons in which cpy is continually reestimated from the information that was available at the point at which the forecast is made. Again, the mean squared prediction error of the model involving cpy is always smaller than that of the alternative model. The difference is also generally significant (with the exception of the default spread). In general, however, even a reestimated cpy does at least as well as some of the best forecasting variables on this sample. Note also that cpy – in this first half of our sample, which stretches from 1952 to 1980 – even outperforms cay in its out-of-sample performance.

5 Changes in the link between entrepreneurial risk and the stock market

The results in the previous sections support the notion that fluctuations in the entrepreneurial consumption–income ratio are an indicator of entrepreneurial risk and that this factor is important in explaining time variation in the equity premium in the United States. However, the results also show that cpy has virtually no predictive power in the second half of the sample period. As I argue in this section, one would expect the impact of entrepreneurial risk on aggregate pricing relations to have changed over the sample period, because of at least two developments that should affect the very pillars of the transmission mechanism between cpy and the stock market.

The first of these developments is that household access to stock markets has been gradually widening since the early 1980s: the share of households owning stocks increased from 19% in 1983 to 49.5% in 2002. As argued by e.g. Poterba (1994), an important driver of the increase in household stock ownership was the growth of (employer-sponsored) 401(k) plans as a form of retirement savings. [19] Because by definition only employees are eligible

[19] According to Poterba, the number of households owning 401(k) plans rose from 4.4 million in 1983 to 20.4 million in 1993, and equity accounts for a large share of the assets under

for such plans, the role of proprietary income for the average stock-owning household is therefore likely to have declined. Unfortunately, detailed participation data on the share of households that participate in the stock market (such as the illustrative numbers just provided) are not available on a very regular basis. In addition, the share of households participating could be a poor measure for our purposes here if in fact new participants (who mainly own stocks through pension plans) accounted for a small share of the entire stock market. I therefore use the share of domestic stock market wealth that is held by pension funds as a value-weighted proxy for the growing nonentrepreneur participation in the stock market. These data are available on an annual basis from 1952 from the 2010 Statistical Abstract. They are plotted in the first panel of Figure 5. As is apparent, there is sharp increase in the share of domestic equity wealth held by pension funds in the 1970s and early 1980s.

A second, even more important development is that proprietary income has become more diversifiable through state-level banking deregulation. Small, owner-run businesses are particularly dependent on bank-intermediated finance. As argued in Jayaratne and Strahan (1996), the removal of branching and interstate ownership restrictions on banks has led to a more efficient allocation of capital and better risk sharing between banks. Jayaratne and Strahan present evidence that these efficiency gains are likely to have been passed on to banks' customers. Bank loans effectively work as a risk-sharing device because they allow the firm to smooth temporary fluctuations in cash flow. More recently, Demyanyk, Ostergaard and Sørensen (2007) have found that state-level personal income has indeed become less sensitive to

administration in 401(k) plans. According to the 2002 *Equity Ownership in America* survey, 33.2 million households owned stock mutual funds within employer-sponsored retirement plans in 2002.

state-level shocks after banking deregulation. They also show that this decline in the sensitivity of state-level personal income is driven mainly by a decline in the sensitivity of the proprietary income component of personal income. In a recent paper, Hoffmann and Shcherbakova-Stewen (forthcoming) show that state-level banking deregulation has improved risk sharing mainly during US-wide recessions. They document that these improvements in risk sharing during recessions are actually strongest in states with many small businesses. These findings in the literature suggest that small businesses particularly benefited from banking deregulation through better access to borrowing facilities – and that this access has improved particularly during periods of aggregate distress. As a consequence, entrepreneurial risk is likely to have become more diversifiable, and the risk for stock-owning entrepreneur households of having to liquidate a stock portfolio in bad times is likely to have decreased substantially. For the analysis in this paper, this should imply that in the wake of bank deregulation, entrepreneurial risk should have become less important in explaining stock market risk premia. The second panel of Figure 5 plots the number of states that had deregulated by a given year. The data capture the year of the lifting of intrastate bank branching restrictions and are from Demyanyk, Ostergaard and Sørensen (2007). Again it is apparent that there was a wave of deregulation in the early 1980s, exactly at the time when the predictive power of cpy starts to decline.

In Table 9, I show how banking deregulation and increased participation have affected the predictive power of cpy for excess returns. The table reports the results of long-horizon regressions – now estimated on the whole sample, 1952Q4–2010Q4 – in which, besides cpy, I include interaction terms with the banking deregulation and participation trend variables.

Specifically, I estimate regressions of the form :

$$\sum_{l=1}^{k} r_{t+l} - r^{f}_{t+l} = \delta_{1k} cpy_t + cpy_t \times \begin{bmatrix} BD_t & PR_t & ... \end{bmatrix} \times \boldsymbol{\delta}_{2k} + \begin{bmatrix} BD_t & PR_t & ... \end{bmatrix} \gamma + u^{k}_{t} \tag{8}$$

where $\begin{bmatrix} BD_t & PR_t & ... \end{bmatrix}$ is a vector containing, in turn, the factors that I expect to affect the predictive power of cpy_t over time: BD_t, my banking deregulation measure[20] and PR_t, the participation rate as measured by the share of equity held in pension funds. The dots are meant to indicate that this vector may also contain other potential trend variables, such as a linear trend. The vector $\boldsymbol{\delta}_{2k}$ stacks the coefficients on the respective interaction terms, and γ stacks the coefficients on the first-order terms.

Panel I of Table 9 presents results for the case when I control for deregulation, BD_t, on its own: the interaction term is highly significant and negative at all horizons. This suggests that banking deregulation has indeed contributed to the declining predictive power of cpy. In Panel II, I add the interaction with the participation rate, which, however, is not significant at any horizon and, if anything, has a positive coefficient. However, the coefficients on the interaction between cpy_t and BD_tremain negative and significant. This suggests that the link between cpy and expected stock returns is mainly affected by the impact of banking deregulation on the background risk faced by small businesses.

Clearly, the link between cpy and expected returns could have declined for a range of other reasons. For example, the rise of small incorporated firms (so-called S corporations) in the early 1980s or the increasing trend towards the incorporation of partnerships could have led to some forms of

[20]Specifically, BD_t is the fraction of states that had completely liberalized bank-branching regulation in the calendar year of period t.

proprietary income being registered as corporate dividends.[21] However, the first of these trends should not affect our measure of p, because proprietary income as recorded in the NIPA accounts also includes dividend payments from S-corporations. Furthermore, none of the results in this paper concerning the predictive power of cpy for stock markets and the decline of this predictive ability in the second half of the sample is sensitive to whether dividend income is included into the construction of cpy at all or whether proprietors' income is included in labor income, y, or in proprietors' income, p. This suggests that changes in the definition of dividend and proprietary income that were caused by changes in corporate status *per se* most likely cannot account for the decline in the predictive power of cpy. However, changes in taxation and asymmetries in the tax treatment of dividends and proprietary income could have affected the link between cpy and the stock market, as could have other latent changes in the economic environment.

To control for such developments, I include a linear trend in the interaction terms. When the trend alone is included (Panel III), it is indeed highly significant and negative at all horizons. However, once I include the trend along with the deregulation trend, BD_t,none of the two variables is significant at conventional significance levels. This should not be surprising, because one would expect a high degree of collinearity between any two such trend variables. What is telling however, is that the coefficient on the interaction between BD_t and cpy_t remains very stable – at all forecasting horizons – vis-à-vis the specification (in Panel I) in which only $BD_t \times cpy_t$

[21]For example, if a Wall Street partnership like Goldmann & Sachs incorporates, this may lead to payouts to partners being recorded as corporate dividends instead of proprietary income. Also, top labor incomes are highly exposed to the stock market, which could have worked against the decreasing correlation between cpy and returns. However, to the extent that top labor incomes are included in y, the findings here suggest that any cyclical component in y remains small compared to the one in p. I thank an anonymous referee for pointing out these possibilities.

was included. Conversely, the coefficients on the interaction between the linear trend and cpy becomes very unstable vis-à-vis the trend-only specification reported in Panel III.[22] These findings lend further support to the view that it is mainly the trend in state-level banking deregulation – and not some other concurrent development – that has affected the link between cpy and expected stock market returns.

I further examine changes in the link between cpy and cay, and between cpy and the default spread in Table 10. Again, I run regressions in which cpy is interacted with the banking deregulation and participation trends, as in (8) above, but now with cay and the default spread as dependent variables:

$$z = \theta_0 cpy_t + cpy_t \times \left[\begin{array}{ccc} BD_t, & PR_t & t \end{array} \right] \boldsymbol{\theta} + \left[\begin{array}{ccc} BD_t, & PR_t & t \end{array} \right] \boldsymbol{\gamma} + \mu + v_t$$

where z_t stands in turn for cay and the default spread.

The comparison between cpy and cay is motivated by the fact that cay is a general indicator of expected returns in the stock market that is consistent with a broad range of economic theories. To the extent that cay can be interpreted as a stand-in for the average market participants' stochastic discount factor, it is therefore interesting to know whether the correlation between cpy and cay has decreased mainly because entrepreneurial background risk has decreased (because of banking deregulation) or because cay signals fluctuations in the average discount factor of what has effectively become a much larger group of stock market participants (because of widening participation).

The results in the first four columns of Table 10 suggest that both developments seem to some extent able to account for changes in the correlation

[22]Furthermore, while t-statistics are not significant at conventional levels for both interactions, those on $BD_t \times cpy_t$ are all much higher than those on the interaction between cpy and the linear trend.

between cay and cpy: when BD_t and PR_t are included individually in the interaction regression, they are both individually significant and negatively signed, as theory would suggest. However, if both are included at the same time, it is only BD_t that remains significant. Note also that the size of the coefficient on BD_t is again stable across the two specifications, at around -0.7. Given that the coefficient θ_0 on cpy alone is of the same magnitude (around $0.7 - 0.8$) but with a positive sign, this suggests that the positive correlation between cpy and cay had virtually dropped to zero once banking deregulation was complete (so that $BD_t = 1$). Once, however, I also control for other, unobserved developments using an interaction between cpy and a linear trend (column 4), it is participation that becomes highly significant and negative. The interaction between time and cpy is significant but positive. Again, the coefficient on BD_t, though insignificant, stays remarkably stable vis-à-vis earlier specification. My reading of these results is that it banking deregulation seems to have a more consistent and robust impact on the decline in the correlation between cpy and cay than participation. However, both developments together seem to have mattered more for this correlation than other possible trends, consistent with the mechanism that I have been highlighting throughout the paper.

In the remaining columns of Table 10, I show that banking deregulation can also account for the decrease in the correlation between cpy and the default spread: BD_t is again significant and negative. The size of the coefficient on the interaction between cpy_t and BD_t (-0.14) is roughly equal to the negative of the coefficient θ_0 on cpy alone (0.12). Hence, by the time that banking deregulation was complete, the correlation between the default spread and cpy had vanished, very much as the correlation between cpy and cay. Conversely, the participation trend has no effect on the default spread by its own (column 6), different from the evidence provided for cay. This is consistent with the view that the default spread does indeed signal the risks faced by small businesses and not those faced by a wider group of stock market participants (as does cay). Finally, the regression in column 7 shows that the effect of banking deregulation on the correlation between cpy and the default spread is robust to the inclusion of an interaction of cpy with a linear trend, suggesting that other, concurrent developments are, again, not likely to account for the findings here.

6 Conclusion

This paper has proposed an entrepreneurial distress factor based on the residual of a cointegrating relationship between consumption, proprietors' income and other income in the economy. I call this residual cpy. While cpy is based on minimal theoretical assumptions because it is derived from the log-linearization of the average household's budget constraint, its economic interpretation is straightforward: when proprietors' income is low in relation to other income and aggregate consumption in the economy, the average small business entrepreneur is relatively likely to face hard times. Consistent with this interpretation, cpy correlates negatively with cross-sectional measures of entrepreneurial risk and positively with the default spread. At the same time, cpy also has considerable predictive power for excess returns in the US stock market.

Heaton and Lucas (2000*b*,*a*) prominently argued that proprietary income is an important source of uninsurable background risk for those households that participate in the stock market. To the extent that fluctuations in average proprietary income capture fluctuations in entrepreneurial risk, they are therefore likely to enter aggregate asset pricing relations. Their studies have shown that entrepreneurial risk does indeed matter for the cross section of stock returns. To my knowledge, the result here are the first to demonstrate that the same mechanism can help explain why expected returns vary *over time*.

The entrepreneurial risk mechanism rests on two pillars: limited participation and uninsurable background risk. Over my sample period, which ranges from 1952 to 2010, two developments in particular could therefore have affected these pillars. First, with the advent of 401 (k) defined contribution plans, stock ownership has widened to new household groups, making proprietary income and the associated entrepreneurial risk less important for the average stock holder. Second, state-level banking deregulation is likely to have facilitated small firms' access to credit, making it easier for small firms to alleviate the risks associated with liquidity shortages and thus effectively providing risk sharing for proprietors. In line with the theory, both developments have had a significant impact on the link between cpy and the stock market: while cpy has considerable forecasting power in the first half of the sample, its link with the stock market has declined as participation has widened to new household groups and – most importantly – as banking deregulation has facilitated the sharing of proprietary income risk. While these findings provide important evidence for the importance of the entrepreneurial risk mechanism in understanding postwar stock market dynamics, they also suggest that the mechanism may be considerably less important today than it used to be. My results constitute the first evidence on changes in the role of this important mechanism in postwar US stock markets. In future work, it will be interesting to explore the impact of cpy on the cross section of stock returns and how it has changed over time.

References

Campbell, John Y., and N. Gregory Mankiw. 1989. "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence." In *NBER Macroeconomics Annual 1989*. Vol. 4, , ed. Olivier Jean Blanchard and Stanley Fischer. MIT Press.

Cochrane, J. 1994. "Permanent and Transitory Components of GDP and Stock Returns." *Quarterly Journal of Economics*, 109(1): 241–266.

Constantinides, G.M. 2002. "Rational Asset Prices." *Journal of Finance*, 57: 1567–1591.

Constantinides, G.M., and D. Duffie. 1996. "Asset Pricing with Heterogeneous Consumers." *Journal of Political economy*, 104(2): 219–240.

Demyanyk, Yuliya, Charlotte Ostergaard, and Bent E. Sørensen. 2007. "U.S. Banking Deregulation, Small Businesses, and Interstate Insurance of Personal Income." *Journal of Finance*, 62(6): 2763–2801.

Engle, R.F., and C.W.J. Granger. 1987. "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica*, 55: 251–276.

Gertler, Mark, and R. Glenn Hubbard. 1993. "Corporate Financial Policy, Taxation, and Macroeconomic Risk." *The RAND Journal of Economics*, 24(2): 286–303.

Gertler, Mark, and Simon Gilchrist. 1994. "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms." *Quarterly Journal of Economics*, 109 (2): 309–40.

Gonzalo, J., and C. Granger. 1995. "Estimation of Common Long-Memory Components in Cointegrated Systems." *Journal of Business & Economic Statistics*, 13: 27–35.

Gonzalo, Jesus, and Serena Ng. 2001. "A systematic framework for analyzing the dynamic effects of permanent and transitory shocks." *Journal of Economic Dynamics and Control*, 25(10): 1527–1546.

Harvey, David I, Stephen J Leybourne, and Paul Newbold. 1998. "Tests for Forecast Encompassing." *Journal of Business & Economic Statistics*, 16(2): 254–59.

Heaton, J., and D. Lucas. 2000*a*. "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk." *The Journal of Finance*, 55(3): 1163–1198.

Heaton, J., and D. Lucas. 2000*b*. “Portfolio Choice in the Presence of Background Risk.” *The Economic Journal*, 110(460): 1–26.

Hoffmann, M., and I. Shcherbakova-Stewen. forthcoming. “Consumption Risk Sharing over the Business Cycle: the Role of small Firms’ Access to Credit Markets.” *The Review of Economics and Statistics*, , (0).

Hoffmann, Mathias. 2001. “Long Run Recursive VAR Models and QR Decompositions.” *Economics Letters*, 73(1): 15 – 20.

Im, K.S., M.H. Pesaran, and Y. Shin. 2003. “Testing for Unit Roots in Heterogeneous Panels.” *Journal of Econometrics*, 115(1): 53–74.

Jayaratne, Jith, and Philip E. Strahan. 1996. “The Finance-Growth Nexus: Evidence from Bank Branch Deregulation.” *Quarterly Journal of Economics*, 111 (3): 639–670.

Johansen, Søren. 1991. “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models.” *Econometrica*, 59: 1551–1580.

Johansen, Søren. 1995. *Likelihood-based Inference in cointegrated Vector autoregressive Models.* Oxford:Oxford University Press. ISBN 0-19-877449-4.

Lamont, O. 1998. “Earnings and Expected Returns.” *Journal of Finance*, 53: 1563–1587.

Lettau, Martin, and Sidney Ludvigson. 2001. “Consumption, Aggregate Wealth, and Expected Stock Returns.” *Journal of Finance*, 56(3): 815–849.

Lettau, Martin, and Sydney C. Ludvigson. 2004. “Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption.” *American Economic Review*, 94(1): 276–299.

Mankiw, N.G., and S.P. Zeldes. 1991. “The Consumption of Stockholders and Non-Stockholders.” *Journal of Financial Economics*, 29(1): 97–112.

McCracken, Michael. 2007. “Asymptotics for out of sample tests of Granger causality.” *Journal of Econometrics*, 140(2): 719–752.

Moskowitz, T., and A. Vissing-Jørgensen. 2002. “The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle.” *American Economic Review*, 92(4): 745–777.

Park, Ki Young. forthcoming. “Interstate Banking Deregulation and Bank Loan Commitments.” *The B.E. Journal of Macroeconomics (Advances)*.

Polkovnichenko, V. 2004. “Limited Stock Market Participation and the Equity Premium.” *Finance Research Letters*, 1(1): 24–34.

Poterba, J. 1994. “Individual Ownership of Corporate Stock: Aggregate and Survey Evidence.” *manuscript dated September*.

Proietti, T. 1997. “Short-Run Dynamics in Cointegrated Systems.” *Oxford Bulletin of Economics and Statistics*, 59(3): 405–422.

Rudd, Jeremy, and Karl Whelan. 2006. “Empirical Proxies for the Consumption-Wealth Ratio.” *Review of Economic Dynamics*, 9(1): 34–51.

Stock, J.H., and M.W. Watson. 1993. “A simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems.” *Econometrica*, 6: 783–820.

Valkanov, R. 2003. “Long-Horizon Regressions: Theoretical Results and Applications.” *Journal of Financial Economics*, 68(2): 201–232.

Vissing-Jørgensen, A. 2002. “Limited Asset Market Participation and the Elasticity of Intertemporal Substitution.” *Journal of Political Economy*, 110(4): 825–853.

Vissing-Jørgensen, A., and O.P. Attanasio. 2003. “Stock-market Participation, Intertemporal Substitution, and Risk-aversion.” *American Economic Review*, 93(2): 383–391.

Data Appendix

Consumption My source is NIPA table 1.5.5 ‘Gross Domestic Product: expanded detail’. I follow Lettau and Ludvigson in the construction of the consumption aggregate: nondurables consumption is constructed as consumption expenditure on nondurable goods (line 7) and services (line 12) less expenditure on shoes and clothing (line 9). Total consumption expenditure also includes expenditure on durables (line 3).

Proprietary and other income The data source is NIPA table 2.1 ‘Personal Income and its disposition’. My measure of entrepreneurial income is nonfarm proprietary income inclusive of rents (lines 11 and 12). Other (‘labor’) income includes compensation of employees (line 2) and transfers (line 16). The budget constraint underlying cpy should contain all forms of income, including asset income (line 13 of NIPA Table 2.1). I therefore have to allocate asset income to proprietary and labor income in some way. The results reported in the paper are based on a specification in which the fraction *(1 − share of equity in pension funds at time t)* of dividend income (line 15) is included in proprietary income and the remaining share of dividend income and all interest income (line 14) are included in labor income (see below for the source of the share of equity held in pension funds). This procedure implicitly accounts for the widening stock market participation, since entrepreneurs will obtain a smaller share of the dividends on public equity over time. Assigning all dividend income to proprietors or to labor income alternatively does not substantially change the results, though.

Finally, I obtain disposable income measures as follows: disposable proprietors' income = proprietors' income − (personal income − disposable income) × proprietors income / personal income; and disposable other income = other income − (personal income − disposable income) × other income / personal income.

Consumption deflator, population data All income and consumption data in the paper are deflated using the index of personal consumption expenditure PCE. Per capita values are obtained using population from NIPA table 2.1 'Personal Income and its disposition', line 41.

Stock market data, T-bill rates and the default spread The data on the S&P500 are from Robert Shiller's homepage:
`http://www.econ.yale.edu/~shiller/data.htm`.
As a broader index, I also use the CRSP's value-weighted index. Excess returns are constructed using the three-month T-Bill rate from the Federal Reserve Board. The default spread is the yield difference between the return on *Baa*- and *Aaa*-rated corporate bonds as published by the Federal Reserve Board.

Data on banking deregulation, household stock ownership and equity in pension funds Data on banking deregulation are obtained from Table 1 in Demyanyk, Ostergaard and Sørensen (2007). The illustrative data on the share of households owning stocks mentioned in the text are from various issues of *Equity Ownership in America* Survey . The share of public equity held in pension funds is available annually from 1952 to 2009, from the 2010 US Statistical Abstract (*Table 1200: Financial Asset Ownership by Type of Investor*). In the regressions based on quarterly data, the annual observation has been used for all four quarters of the respective year. Values from 2009 have also been used for 2010.

Consumption–wealth ratio I updated cay till 2010Q4, following Lettau and Ludvigson (2001, 2004). It is virtually identical to the cay kindly made available on Martin Lettau's web page (http://faculty.haas.berkeley.edu/lettau/data_cay.html) but which only extended to 2010Q2 at the time I completed this paper.

Table 1: Tests for cointegration

	Trace Test		Max.EigValue	
Hypothesis on number of cointegrating relations	statistics	95% CV	statistics	95% CV
Panel I: Non-Durables Consumption				
$H_0: 0 < h$ vs. $H_1: h \geq 1$	44.80	31.25	36.45	21.27
$H_0: 1 < h$ vs. $H_1: h \geq 2$	8.35	17.84	7.64	14.59
$H_0: 2 < h$ vs. $H_1: h = 3$	0.72	8.80	0.72	8.08
Panel II: Total Consumption				
		90% CV		90% CV
$H_0: 0 < h$ vs. $H_1: h \geq 1$	29.53	28.43	18.96	18.96
$H_0: 1 < h$ vs. $H_1: h \geq 2$	10.57	15.58	10.12	12.78
$H_0: 2 < h$ vs. $H_1: h = 3$	0.4491	6.69	0.4491	6.69

NOTES: the table provides Johansen's tests of the null $n-1 < h$ vs. $h \geq n$ (Trace) or $h = n$ (max. Eigenvalue) cointegrating relationships among the three variables c, p and y. Values significant at the 5%-level are in bold.

Table 2: estimated cointegrating vectors

	Non-Durables Consumption		Total Consumption	
	Johansen	Dynamic OLS	Johansen	Dynamic OLS
β_c	1.0000	1.0000	1.0000	1.0000
β_p	-0.2570	-0.2613	-0.2285	-0.2335
β_y	-0.7563	-0.7621	-0.7078	-0.7205

NOTES:Estimates of the cointegrating vector $\beta = \begin{bmatrix} 1 & \beta_p & \beta_y \end{bmatrix}'$. The estimate from the Johansen-procedure is based on a VECM with one lagged difference term and an unrestricted intercept. The Dynamic OLS regression is of the form $c_t = \mu - \beta_p p_t - \beta_y y_t + \sum_{l=-k}^{k} a_l \Delta p_{t-l} + b_l \Delta y_{t-l} + u_t$ where $k = 3$ leads and lags have been chosen.

Table 3: Estimated VECM

	Equation		
	Δc_t	Δp_t	Δy_t
Δc_{t-1}	**0.34**	0.43	**0.64**
	(5.11)	(1.73)	(4.82)
Δp_{t-1}	**0.04**	**0.21**	-0.05
	(2.12)	(3.12)	(-1.51)
Δy_{t-1}	0.01	-0.09	0.01
	(0.37)	(-0.61)	(0.13)
cpy_{t-1}	**-0.04**	**0.25**	-0.01
	(-2.56)	(4.15)	(-0.39)
$const$	**0.003**	0.002	**0.003**
	(8.09)	(1.21)	(2.75)
$\overline{R}^2$	0.23	0.15	0.11

NOTES: t-statistics in parentheses, coefficients significant at the 5% level are in bold-face. $cpy = c - 0.2570p - 0.7563y$

.

Table 4: Variance decompositions

Variance share of transitory component

	Horizon k in quarters						
	1	2	4	8	12	16	24
$c_{t+k} - \mathbf{E}_t(c_{t+k})$	0.17	0.12	0.07	0.05	0.03	0.02	0.02
	[0.02,0.37]	[0.01,0.27]	[0.01,0.18]	[0.01,0.11]	[0.01,0.08]	[0.00,0.06]	[0.00,0.05]
$p_{t+k} - \mathbf{E}_t(p_{t+k})$	0.45	0.40	0.35	0.29	0.24	0.19	0.13
	[0.25,0.72]	[0.21,0.64]	[0.16,0.57]	[0.13,0.49]	[0.10,0.39]	[0.08,0.31]	[0.06,0.22]
$y_{t+k} - \mathbf{E}_t(y_{t+k})$	0.00	0.03	0.04	0.03	0.02	0.01	0.01
	[0.00,0.12]	[0.01,0.13]	[0.01,0.12]	[0.01,0.09]	[0.01,0.08]	[0.00,0.08]	[0.00,0.06]

NOTES: Numbers in parentheses give the 90% confidence intervals obtained from a bootstrap with 250 replications.

Table 5: cpy,cay **and excess returns – pre- and post banking deregulation**

$$\sum_{l=1}^{k} \Delta q_{t+l} = \delta_k cpy_t + \gamma_k res_t + \mu_k + v_{kt}$$

$\mathbf{z}_t$		Horizon k in quarters						
		1	2	4	8	12	16	24
Panel I: 1952:Q1-1980:Q4								
cpy	δ_k	**1.29****	**2.55****	**4.54****	**7.87****	**10.27****	**11.29****	**16.00****
		(3.59)	(3.60)	(3.14)	(4.24)	(5.44)	(5.75)	(7.73)
$\overline{R^2}$		0.08	0.14	0.22	0.35	0.49	0.53	0.60
cay	δ_k	**1.03****	**2.34****	**4.84****	**8.51****	**10.63****	**12.36****	**17.63****
		(2.62)	(2.85)	(3.13)	(4.81)	(5.80)	(8.34)	(15.51)
$\overline{R^2}$		0.05	0.11	0.25	0.40	0.52	0.63	0.72
cpy	δ_k	**1.04****	**1.81****	**2.57****	**4.27****	**5.90****	**5.46****	**7.55****
		(2.24)	(2.37)	(2.03)	(3.06)	(7.42)	(3.93)	(5.95)
cay	γ_k	0.41	1.23	**3.22****	**5.81****	**6.74****	**8.63****	**12.44****
		(0.83)	(1.32)	(2.07)	(3.04)	(3.74)	(4.26)	(7.71)
$\overline{R^2}$		0.08	0.15	0.29	0.46	0.61	0.70	0.79
Panel II: 1981:Q1-2010:Q4								
cpy	δ_k	-0.79	-1.95**	**-4.56****	**-5.97**	-5.21	0.11	**9.95**
		(-1.14)	(-1.62)	(-2.18)	(-2.09)	(-1.28)	(0.03)	(2.18)
$\overline{R^2}$		0.00	0.03	0.09	0.09	0.04	-0.01	0.12
cay	δ_k	0.68**	**1.60****	**3.71****	**7.98****	**11.61****	**12.30****	**9.67**
		(1.67)	(2.41)	(2.96)	(4.46)	(5.50)	(3.64)	(2.28)
$\overline{R^2}$		0.01	0.05	0.14	0.37	0.54	0.45	0.21
cpy	δ_k	-0.85	-2.09**	**-4.88****	**-6.61****	**-5.86****	-2.36	**6.73**
		(-1.24)	(-1.83)	(-2.66)	(-3.68)	(-3.43)	(-0.93)	(2.06)
cay	γ_k	0.71**	**1.67****	**3.88****	**8.21****	**11.76****	**12.52****	**8.13**
		(1.71)	(2.35)	(2.92)	(5.20)	(5.81)	(3.82)	(2.17)
$\overline{R^2}$		0.02	0.08	0.24	0.49	0.60	0.45	0.25

NOTES: OLS regressions. $t-$statistics are based on heteroskedasticity and autocorrelation consistent standard errors based on Newey and West (1987), using a window width of $k+1$. Boldface coefficients are significant at the 95% level using standard critical values for the t-distribution, whereas a double asterisk ('**') indicates significance of the coefficient using Valkanov's (2003) $t/\sqrt{T}$ statistics, where T is the sample size of the respective regression. Small-sample critical values of the Valkanov-statistics have been simulated using 5000 replications.

Table 6: Long-horizon regressions on cpy: 1952:Q1–1980Q4

$$\sum_{l=1}^{k} \Delta q_{t+l} = \delta_k cpy_t + \mu_k + v_{kt}$$

	Horizon k in quarters						
	1	2	4	8	12	16	24
Panel I: excess returns on S&P500 - $\Delta q_{t+l} = r_{t+l} - r^f_{t+l}$							
δ_k	**1.29****	**2.55****	**4.54****	**7.87****	**10.27****	**11.29****	**16.00****
t-stat	(3.59)	(3.60)	(3.14)	(4.24)	(5.44)	(5.75)	(7.73)
$\overline{R^2}$	0.08	0.14	0.22	0.35	0.49	0.53	0.60
Panel II: excess returns on CRSP - $\Delta q_{t+l} = r_{t+l} - r^f_{t+l}$							
δ_k	**1.26****	**2.47****	**4.33****	**7.45****	**9.78****	**10.45****	**14.97****
t-stat	(3.03)	(3.16)	(2.82)	(4.12)	(5.29)	(5.78)	(7.39)
$\overline{R^2}$	0.05	0.10	0.17	0.30	0.46	0.50	0.55
Panel III: Earnings on S&P 500 - $\Delta q_{t+l} = \Delta e_{t+l}$							
δ_k	-0.27	-0.25	0.57	**2.61****	**4.38****	**5.26**	**4.77**
t-stat	(-1.07)	(-0.48)	(0.59)	(2.00)	(2.93)	(3.65)	(1.96)
$\overline{R^2}$	0.01	-0.00	-0.00	0.08	0.19	0.25	0.21
Panel IV: Dividends on S&P 500 - $\Delta q_{t+l} = \Delta d^{S\&P}_{t+l}$							
δ_k	0.07	0.18	0.60	**2.22****	**3.55****	**4.70****	**6.60****
t-stat	(0.61)	(0.83)	(1.52)	(3.63)	(5.76)	(5.67)	(4.95)
$\overline{R^2}$	-0.00	0.00	0.04	0.21	0.33	0.42	0.49
Panel V: Personal Dividend Income from BEA - $\Delta q_{t+l} = \Delta d^{BEA}_{t+l}$							
δ_k	0.12	0.25	**0.80****	**2.02****	**2.90****	**3.37****	**3.67****
t-stat	(0.95)	(1.16)	(2.22)	(4.78)	(5.79)	(3.97)	(2.67)
$\overline{R^2}$	-0.00	0.01	0.07	0.23	0.36	0.41	0.37

NOTES: OLS regressions. $t-$statistics are based on heteroskedasticity and autocorrelation consistent standard errors based on Newey and West (1987), using a window width of $k+1$. Boldface coefficients are significant at the 95% level using standard critical values for the t-distribution, whereas a double asterisk ('**') indicates significance of the coefficient using Valkanov's (2003) $t/\sqrt{T}$ statistics, where T is the sample size of the respective regression. Small-sample critical values of the Valkanov-statistics have been simulated using 5000 replications.

Table 7: Long-horizon regressions of returns on cpy, cay and 'usual suspects, 1952:Q1–1980Q4

$$\sum_{l=1}^{k} r_{t+l} - r^f_{t+l} = \mathbf{z}_t' \delta_k + v_{kt}$$

$\mathbf{z}_t$	Horizon k in quarters						
	1	2	4	8	12	16	24
cpy	**1.16****	**2.18****	**3.95****	**5.82****	**7.61****	**9.37****	**12.59****
	(2.85)	(2.83)	(2.90)	(4.63)	(5.06)	(5.99)	(7.28)
res	0.24	0.71	2.34	2.56	3.52	**7.19**	**9.47****
	(0.45)	(0.71)	(1.26)	(1.15)	(1.50)	(2.89)	(5.71)
$d-p$	0.02	0.06	0.09	0.35	0.39	0.19	0.37
	(0.65)	(1.05)	(0.91)	**(2.03)**	**(2.02)**	(1.03)	(3.36)
R^2	0.07	0.15	0.29	0.52	0.66	0.71	0.81
cpy	**1.15****	**2.24****	**3.64****	**6.18****	**8.33****	**9.40****	**14.30****
	(2.96)	(3.20)	(2.80)	(4.66)	(7.41)	(12.37)	(16.68)
res	0.51	1.49	**4.06****	**7.30****	**8.16****	**9.53**	**12.89**
	(0.96)	(1.45)	(2.45)	(4.27)	(4.88)	(4.78)	(7.79)
$d-e$	0.05**	0.10**	0.25**	0.45	0.49	0.39	0.21
	(0.79)	(0.96)	(1.45)	(1.76)	**(2.37)**	**(2.64)**	(1.35)
R^2	0.08	0.15	0.32	0.51	0.65	0.72	0.79
cpy	**1.33****	**2.61****	**4.50****	**7.59****	**9.83****	**10.57****	**14.90****
	(4.18)	(4.05)	(3.55)	(5.99)	(9.93)	(10.96)	(23.02)
res	0.39	1.12	**3.19**	**5.87****	**6.63****	**8.50**	**12.35**
	(0.83)	(1.24)	(2.09)	(3.38)	(4.30)	(4.27)	(7.24)
$T-Bill$	-6.34**	-7.46	-7.92	**21.27****	**24.90****	6.44	3.25
	(-1.89)	(-1.30)	(-1.04)	**(4.17)**	**(2.89)**	(0.78)	(0.31)
R^2	0.12	0.17	0.29	0.51	0.66	0.70	0.79
cpy	0.69**	**1.62****	**3.07****	**7.25****	**10.67****	**11.13****	**12.85****
	(1.74)	(2.33)	(2.30)	(5.26)	(9.16)	(9.59)	(13.14)
res	**1.29****	**2.58****	**5.17****	**6.42****	**5.69**	**7.86**	**14.90****
	(2.29)	(2.32)	(3.06)	(2.79)	(3.21)	(3.87)	(7.06)
Default Spread	**9.57****	**14.43****	**22.76****	7.19	**-14.28**	-10.71	**36.97****
	(3.03)	(2.61)	(2.59)	(0.64)	(-2.52)	(-1.28)	(3.74)
R^2	0.15	0.22	0.37	0.46	0.62	0.70	0.81
cpy	**0.77****	**1.74****	**3.15****	**6.87****	**10.31****	**11.09****	**12.72****
	(2.14)	(2.66)	(2.41)	(5.00)	(10.62)	(9.44)	(13.99)
res	**1.20****	**2.41****	**5.08****	**6.87****	**6.01****	**7.85**	**14.90****
	(2.30)	(2.32)	(3.14)	(3.32	(3.80)	(3.89)	(6.47)
TBill	**-5.52****	-5.99	-4.04	**23.17****	**23.23****	4.22	10.58
	(-1.94)	(-1.22)	(-0.60)	(5.09)	(2.62)	(0.42)	(0.97)
Default spread	**8.86****	**13.59****	**21.88****	11.71	-8.51	-9.63	**41.30****
	(3.07)	(2.75)	(2.75)	(0.99)	(-1.43)	(-0.98)	(4.61)
R^2	0.18	0.23	0.37	0.52	0.66	0.70	0.82

NOTES: res is the residual of a regression of cay on cpy. $d-p$ and $d-e$ denote the dividend-price and the dividend-earnings ratio respectively and T-bill is the cyclical component of the three months treasury-bill rate obtained through an HP-filter with $\lambda = 1600$. The Default spread is the HP-filtered cyclical component of the Baa-Aaa yield spread. For further notes see Tables 5 and 6.

Table 8: out-of-sample comparison of cpy with other forecasting variables 1952Q4:1980Q4

Model		CI-vector fixed		reestimated			
		$\frac{MSE_{cpy}}{MSE_{alt}}$	statistics	$\frac{MSE_{cpy}}{MSE_{alt}}$	statistics	crit. val. 95%	90%
Nested comparisons							
			OoS-F		OoS-F		
1	cpy and constant	0.91	3.42	0.95	2.04	2.06	1.12
2	cpy and AR(1)	0.92	3.21	0.95	1.83	2.06	1.12
Non-nested comparisons							
			MDM		MDM		
3	cpy vs. AR(1)	0.90	2.85	0.93	2.18	1.69	1.31
4	cpy vs. $d-p$	0.93	2.15	0.96	1.53	1.69	1.31
5	cpy vs $d-e$	0.87	1.39	0.90	1.36	1.69	1.31
6	cpy vs $Baa-Aaa$	0.93	1.42	0.96	0.94	1.69	1.31
7	cpy vs. cay	0.96	1.72	0.95	2.04	1.69	1.31

For the non-nested models (1 and 2), the table provides the McCracken (2008) out-of-sample F-statistics (OoS-F) and associated critical values, for the non-nested models (3-7), the Harvey et al. (1997) modified Diebold-Mariano (MDM) statistics. MSE_{cpy} is the mean squared prediction error for the model involving cpy andMSE_{alt}that of the alternative model. The reported results are based on recursive predictions, with the initial sample ranging from $1952Q1-1971Q4$. In the first two columns, results are for the case in which the cointegrating vector is fixed to the value estimated for the entire sample (1952Q1-2010Q4), in the third and fourth columns, the cointegrating vector for cpy and cay is re-estimated every period using Johansen's procedure.

Table 9: Factors affecting link between cpy and expected returns :1952:Q1–2010Q4

Long-horizon regressions of excess returns on the S&P 500 on cpy and interactions:

$$\sum_{l=1}^{k} \Delta q_{t+l} = \delta_{1k} cpy_t + \delta_{2k} BD_t \times cpy_t + \delta_{2k} PR_t \times cpy_t + \delta_3 t \times cpy_t + \left[\begin{array}{ccc} BD_t & PR_t & t \end{array} \right] \gamma + \mu_k + \upsilon_{kt}$$

where BD_t is the share of the number of states that had abolished state-level bank branching restrictions at time t and PR_t is the participation rate (as measured by the share of public equity held by pension funds). OLS regressions. $t-$statistics are based on heteroskedasticity and autocorrelation consistent standard errors based on Newey and West (1987), using a window width of $k+1$. Boldface coefficients are significant at the 95% level using standard critical values for the t-distribution, whereas a double asterisk ('**') indicates significance of the coefficient using Valkanov's (2003) $t/\sqrt{T}$ statistics, where T is the sample size of the respective regression. Small-sample critical values of the Valkanov-statistics have been simulated using 5000 replications.

	Horizon k in quarters						
	1	2	4	8	12	16	24
	Panel I: Banking Deregulation						
cpy	**2.03****	**3.97****	**7.08****	**11.46****	**14.61****	**14.50****	**15.02****
	(3.45)	(3.69)	(3.38)	(4.28)	(4.95)	(4.75)	(5.13)
$BD_t \times cpy_t$	**-3.44****	**-6.93****	**-12.88****	**-18.60****	**-21.16****	-14.79	-2.06
	(-2.70)	(-3.24)	(-3.14)	(-3.33)	(-2.81)	(-1.92)	(-0.35)
$\overline{R^2}$	0.04	0.07	0.13	0.17	0.22	0.25	0.44
	Panel II: Banking Deregulation and Participation						
cpy_t	**1.67****	**3.60****	**6.64****	**10.61****	**13.55****	**13.03****	**15.30****
	(3.06)	(3.45)	(3.20)	(3.98)	(4.64)	(4.43)	(4.93)
$BD_t \times cpy_t$	**-5.19****	**-8.72****	**-14.96****	**-22.79****	**-26.56****	**-22.90**	1.22
	(-2.54)	(-2.59)	(-2.46)	(-3.04)	(-2.83)	(-2.27)	(0.12)
$PR \times cpy_t$	7.04**	7.15	8.28	16.69	21.21	31.01	-10.65
	1.27	0.72	0.52	0.87	1.04	1.40	-0.34
$\overline{R^2}$	0.04	0.07	0.13	0.18	0.22	0.26	0.44
cpy_t	**2.01****	**4.02****	**7.24****	**11.81****	**15.50****	**15.04****	**15.34****
	(3.32)	(3.70)	(3.30)	(4.48)	(4.74)	(3.96)	(4.92)
$t \times cpy_t$	**-3.65****	**-7.55****	**-14.24****	**-20.85****	**-25.04****	-17.44	-3.11
	(-2.49)	(-3.22)	(-3.06)	(-3.61)	(-2.99)	(-1.71)	(-0.39)
$\overline{R^2}$	0.03	0.07	0.12	0.17	0.22	0.24	0.44
cpy_t	**1.81****	**3.57****	**6.07****	**8.89****	**11.47****	**10.92****	**12.43**
	(2.87)	(3.27)	(2.95)	(3.90)	(4.61)	(3.50)	(3.18)
$BD_t \times cpy_t$	-3.11**	-5.49**	-9.76	-14.06	-12.95	-11.49	0.40
	(-1.35)	(-1.34)	(-1.34)	(-1.59)	(-1.36)	(-0.97)	(0.03)
$t \times cpy_t$	-0.12	-1.22	-2.51	-2.43	-6.91	-1.79	-3.85
	(-0.05)	(-0.25)	(-0.27)	(-0.24)	(-0.83)	(-0.16)	(-0.26)
$\overline{R^2}$	0.04	0.08	0.15	0.26	0.36	0.39	0.53

Table 10: the changing link of cpy with cay and default spread

Regressions of the form:
$z = \theta_0 cpy_t + cpy_t \times \begin{bmatrix} BD_t & PR_t & t \end{bmatrix} \boldsymbol{\theta} + \begin{bmatrix} BD_t & PR_t & t \end{bmatrix} \boldsymbol{\gamma} + \mu + v_t$ where z_t stands in turn for $cay_t$and the $Baa - Aaa$ spread. Coefficients on first-order terms (γ) are not reported. Sample period is 1952Q1-2010Q4. t-statistics appear in parentheses, coefficients significant at the 95% level in bold.

	$z_t = cay_t$				$z_t = Baa_t - Aaa_t$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
cpy	**0.75**	**0.74**	**0.77**	**0.38**	**0.12**	**0.08**	**0.14**
	(6.09)	(5.50)	(5.68)	(2.83)	(4.91)	(2.80)	(5.41)
$BD_t \times cpy_t$	**-0.80**		*-0.70*	-0.58	**-0.14**		**-0.29**
	(-3.44)		(-1.93)	(-1.19)	(-2.98)		(-2.86)
$PART_t \times cpy_t$		**-1.99**	-0.42	**-4.77**		-0.23	
		(-2.79)	(-0.37)	(-3.33)		(-1.59)	
$t \times cpy_t$				**1.88**			0.15
				(2.31)			(1.29)
R^2	0.20	0.19	0.21	0.39	0.11	0.05	0.18

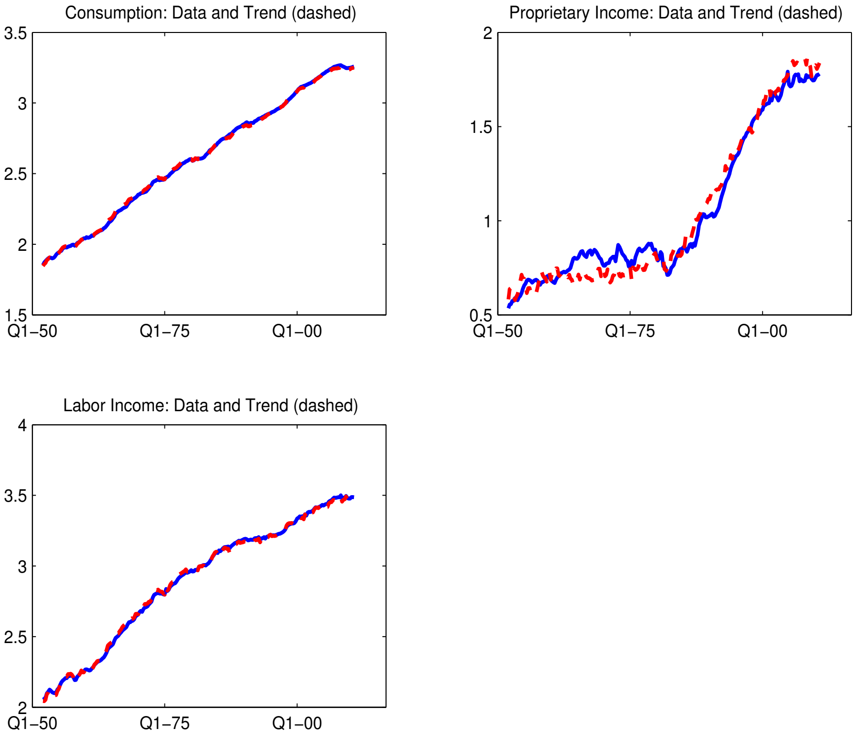

Figure 1: (Logarithmic) Data (solid line) vs their trend components (dashed) as identified from the VECM.

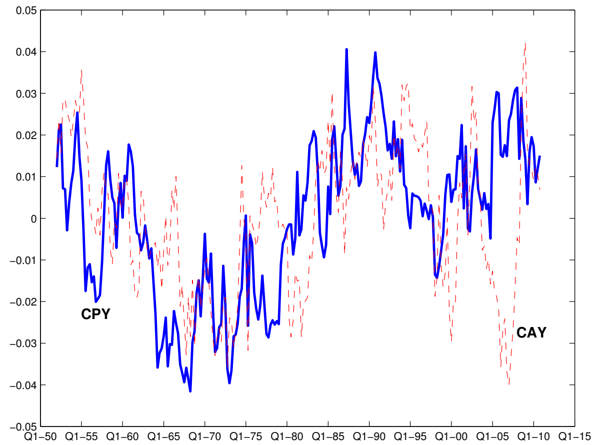

Figure 2: $cpy_t = c_t - 0.2570p_t - 0.7563y_t$ versus Lettau-Ludvigson (2004) cay.

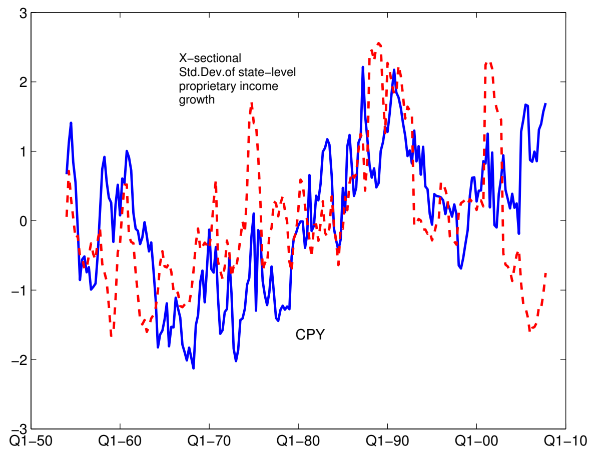

Figure 3: cpy (1952Q4-2007Q4) vs. the cross-sectional standard deviation of 8-quarter growth rates in proprietary income across U.S. federal states.

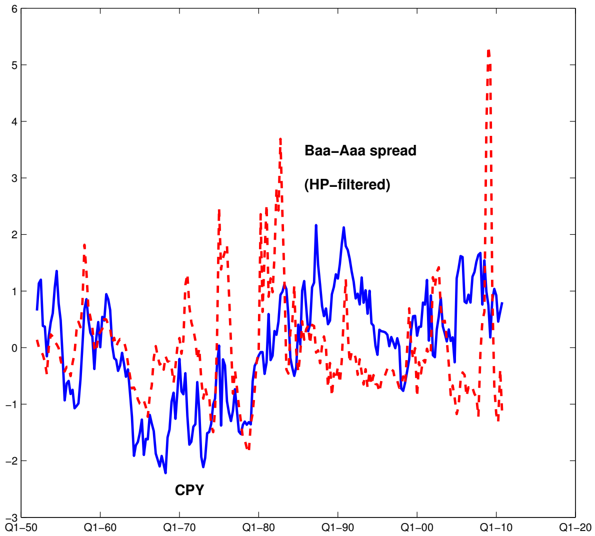

Figure 4: cpy vs. the cyclical component of the default ($Baa - Aaa$) spread. Both series are demeaned and standardized.

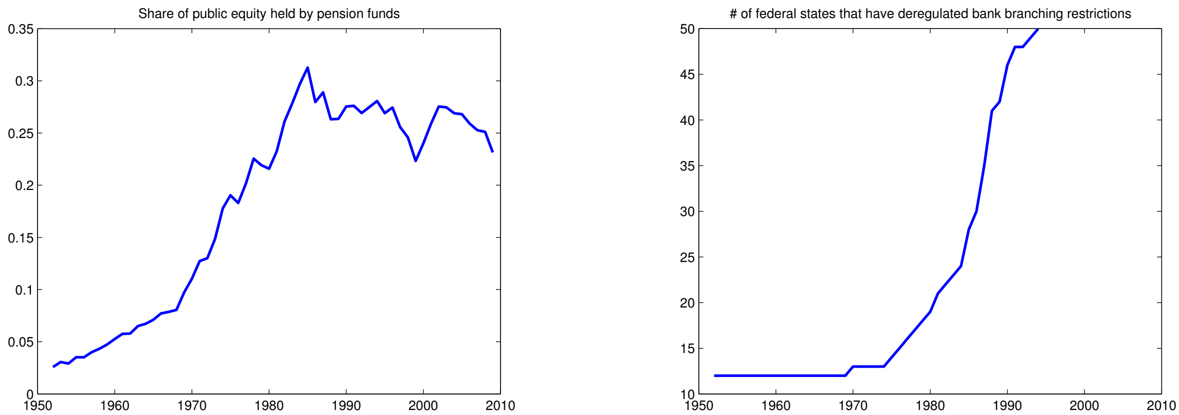

Figure 5: Share of public equity held by pension funds (left) and number of states that had dismantled intra-state branching regulation (right)

Technical appendix (only for publication as supplementary web-material)

Derivation of cpy from the aggregate budget constraint

Note that Ψ_t is the present value of all dividends, $\Psi_t = C_t + \sum_{k=1}^{\infty} \left[\prod_{s=1}^{k} R_{C,t+s} \right]^{-1} C_{t+k}$ where $R_{C,t+s}$ is the gross return on total wealth. This expression can be written recursively as

$$\Psi_{t+1} = R_{C,t+1}(\Psi_t - C_t),$$

which allows the use of the approach adopted by Campbell and Mankiw (1989) for the log-linearization of the consumption–wealth ratio:

$$\frac{\Psi_{t+1}}{\Psi_t} = R_{C,t+1}(1 - \exp(c_t - \psi_t)).$$

Taking logs yields

$$\Delta\psi_{t+1} = r_{c,t+1} + \log(1 - \exp(c_t - \psi_t)).$$

The logarithmic term can now be expanded around the long-run consumption–wealth ratio $\exp(\overline{c - \psi})$ so that

$$\begin{aligned} \log(1 - \exp(c_t - \psi_t)) &= \log(1 - \exp(\overline{c - \psi})) + \frac{-\exp(\overline{c - \psi})}{(1 - \exp(\overline{c - \psi}))} \left[c_t - \psi_t - \overline{c - \psi} \right] \\ &= \kappa_C - \frac{\exp(\overline{c - \psi})}{(1 - \exp(\overline{c - \psi}))} \left[c_t - \psi_t \right] \end{aligned}$$

where

$$\kappa_C = \log(1 - \exp(\overline{c - \psi})) + \frac{\exp(\overline{c - \psi})}{(1 - \exp(\overline{c - \psi}))} \overline{c - \psi}.$$

Write $\Delta\psi_{t+1}$ tautologically as

$$\Delta\psi_{t+1} = \Delta c_{t+1} - (c_{t+1} - \psi_{t+1}) + (c_t - \psi_t)$$

to obtain

$$\begin{aligned} &\kappa_C + r_{c,t+1} + \left[1 - \frac{1}{\rho} \right] \left[c_t - \psi_t \right] \\ &= \Delta c_{t+1} - (c_{t+1} - \psi_{t+1}) + (c_t - \psi_t) \end{aligned}$$

where $\rho_c = 1 - \exp(\overline{c - \psi})$. Then rearrange to obtain

$$\kappa_C + \frac{1}{\rho}\left[c_t - \psi_t\right] = r_{c,t+1} - \Delta c_{t+1} - (c_{t+1} - \psi_{t+1}),$$

which can be solved forward with $\rho_c^k(c_{t+k} - \psi_{t+k}) \to 0$ to get

$$[c_t - \psi_t] = \frac{\rho_c}{1-\rho_c}\kappa_c + \sum_{k=1}^{\infty} \rho_c^k \left[r_{c,t+1} - \Delta c_{t+1}\right].$$

If consumption and wealth are both integrated ($I(1)$) processes, then Δc will be stationary. Assuming that returns are also stationary, the right-hand side of this present-value relation reflects the discounted sum of stationary variables and will therefore be stationary. Hence, $c_t - \psi_t$ is stationary.

Applying the same log-linearization procedure to $p_t - \pi_t$, and $y_t - \theta_t$, I get

$$\psi_t = c_t + \mathbf{E}_t \sum_{k=1}^{\infty} \rho_c^k \left(\Delta c_{t+k} - r_{C,t+k}\right) \tag{9a}$$

$$\pi_t = p_t + \mathbf{E}_t \sum_{k=1}^{\infty} \rho_p^k \left(\Delta p_{t+k} - r_{P,t+k}\right) \tag{9b}$$

$$\theta_t = y_t + \mathbf{E}_t \sum_{k=1}^{\infty} \rho_y^k \left(\Delta y_{t+k} - r_{Y,t+k}\right) \tag{9c}$$

where ρ_x is the mean reinvestment ratio of the respective wealth component; e.g., $\rho_c = 1 - \exp(\overline{c - \psi})$. Plugging into (4), one then obtains the desired relation between c, p and y.

Note that $r_{C,t+k}$ can be interpreted as the return on total wealth, which is the weighted average of returns on proprietary and other forms of wealth, so that approximately

$$r_{C,t+k} \approx \gamma r_{P,t+k} + (1-\gamma) r_{Y,t+k}. \tag{10}$$

To see that cpy must be a cointegrating relationship, plug relations (9) and (10) into the linearized budget constraint (3). Then rearrange the forward-looking terms to the right so that

$$\begin{aligned} cpy = {} & -(1-\gamma)\kappa + \gamma E_t \sum_{k=1}^{\infty} \left(\rho_p^k \Delta p_{t+k} + \left(\rho_c^k - \rho_p^k\right) r_{P,t+k}\right) \qquad (11) \\ & + (1-\gamma) E_t \sum_{k=1}^{\infty} \left(\rho_y^k \Delta y_{t+k} + \left(\rho_c^k - \rho_y^k\right) r_{Y,t+k}\right) - E_t \sum_{k=1}^{\infty} \rho_c^k \Delta c_{t+k}. \end{aligned}$$

From this representation, it is apparent that cpy must be stationary: because c, p and y are all best characterized as individually $I(1)$, the present value of their changes must be stationary. If the returns on wealth are stationary, then their discounted sum must equally be stationary. This implies that cpy will be stationary. It therefore defines a cointegrating relationship that measures the temporary deviation of consumption, proprietary and other income from the common trends.

The deviation of the cointegrating relation from its long-run mean then predicts changes either in consumption or in one of the two components of income: away from the long-run trend, at least one of the three variables will have to adjust.

cpy as approximation of the entrepreneurial income ratio in the model with two household types

Start from the consolidated present values of consumption of proprietors and workers $\Psi_t = \Psi_t^p + \Psi_t^w$. Rearranging and taking logarithms on both sides, we get an equation analogous to (2) above:

$$\log\left(1 - \frac{\Psi_t^p}{\Psi_t}\right) = \psi_t^w - \psi_t. \tag{12}$$

Maintain the assumption from the previous section that the share of proprietary wealth in total wealth is constant in the long run, so that $\gamma = E(\Pi_t/\Psi_t)$ exists. It then follows from the budget constraint of the proprietors that $\gamma = E(\Psi_t^p/\Psi_t)$. Hence, log-linearizing (12) around γ we get

$$\psi_t = \gamma\psi_t^p + (1-\gamma)\psi_t^w + constant. \tag{13}$$

The stationarity of proprietors' and workers' respective consumption–wealth ratios allows us to obtain equations that are analogous to those obtained for the aggregate consumption–wealth ratio in (9a):

$$\begin{aligned}
\psi_t^p &= c_t^p + \mathbf{E}_t \sum_{k=1}^{\infty} \rho_p^k \left(\Delta c_{t+k}^p - r_{P,t+k}\right) \\
\psi_t^w &= c_t^w + \mathbf{E}_t \sum_{k=1}^{\infty} \rho_Y^k \left(\Delta c_{t+k}^w - r_{Y,t+k}\right)
\end{aligned}$$

where $r_{P,t}$ and $r_{Y,t}$ are the internal rates of return on proprietary and non-proprietary wealth from above. Substitute out for the ψ-terms in (13) and, ignoring constants, rearrange terms using $r_{C,t} = \gamma r_{P,t} + (1-\gamma) r_{Y,t}$:

$$\begin{aligned} c_t &= \gamma c_t^p + (1-\gamma) c_t^w \\ &\quad + \gamma \mathbf{E}_t \sum_{k=1}^{\infty} \left\{ \rho_p^k \Delta c_{t+k}^p + \left(\rho_C^k - \rho_P^k \right) r_{P,t+k} \right\} \\ &\quad + (1-\gamma) \mathbf{E}_t \sum_{k=1}^{\infty} \left\{ \rho_Y^k \Delta c_{t+k}^w + \left(\rho_C^k - \rho_Y^k \right) r_{Y,t+k} \right\} \\ &\quad - \mathbf{E}_t \sum_{k=1}^{\infty} \rho_C^k \left(\Delta c_{t+k} \right). \end{aligned}$$

Hence, if aggregate consumption is not very predictable (as is the case in the data) and under the assumption that proprietors' and workers' consumption growth are not too predictable either (or that aggregate consumption growth can be reasonably approximated by $\Delta c_{t+1} = \gamma \Delta c_{t+1}^p + (1-\gamma) \Delta c_{t+1}^w$), the approximation error is

$$\begin{aligned} cpy - \left[\gamma \left(c_t^p - p_t \right) + (1-\gamma) \left(c_t^w - y_t \right) \right] &= \gamma \mathbf{E}_t \sum_{k=1}^{\infty} \left(\rho_C^k - \rho_P^k \right) r_{P,t+k} \\ &\quad + (1-\gamma) \mathbf{E}_t \sum_{k=1}^{\infty} \left(\rho_C^k - \rho_Y^k \right) r_{Y,t+k}. \end{aligned}$$

Note that the terms on the right-hand side of this equation also figure on the right-hand side of (11) and that the approximation error is independent of expected growth rates in p or y. Hence, cpy and $\gamma \left(c_t^p - p_t \right) + (1 - \gamma) \left(c_t^w - y_t \right)$ contain the same information with respect to future changes of p and y. Because cpy mainly reflects expected changes in proprietary income, this means that it is dominated by variation in $c_t^p - p_t$, which (if c^p is not too predictable, as assumed) will also mainly predict changes in p.

Identifying permanent and transitory components

Specifically, Proietti (1997) proposes the following decomposition:

$$\begin{aligned} \mathbf{x}_t &= \mathbf{C}(\mathbf{1}) \mathbf{\Gamma}(\mathbf{1}) \mathbf{x}_t + \left[\mathbf{I} - \mathbf{C}(\mathbf{1}) \mathbf{\Gamma}(\mathbf{1}) \right] \mathbf{x}_t \\ &= \mathbf{x}_t^P + \mathbf{x}_t^T \end{aligned}$$

where $\mathbf{C}(\mathbf{1})$ is the long-run response of $\mathbf{x}_t$ to shocks; i.e., the loading associated with the random walk component in the Beveridge–Nelson–Stock–Watson decomposition of $\mathbf{x}_t$.

To identify permanent and transitory shocks directly, acknowledge that $\mathbf{C}(\mathbf{1})$ can be factored as $\mathbf{C}(\mathbf{1}) = \mathbf{A} \alpha'_\perp$ so that

$$\pi_t = \alpha'_\perp \varepsilon_t$$

can be interpreted as the vector of permanent shocks, the innovations to the random walk component of $\mathbf{x}_t$. By construction, shocks that are transitory with respect to all components of the vector $\mathbf{x}_t$ must be orthogonal to π_t so that these shocks must be given by[23]

$$\tau_t = \alpha' \mathbf{\Omega}^{-1} \varepsilon_t.$$

Collecting permanent and transitory shocks into one vector θ_t,

$$\theta_t = \begin{bmatrix} \pi_t \\ \tau_t \end{bmatrix} = \begin{bmatrix} \alpha'_\perp \\ \alpha' \mathbf{\Omega}^{-1} \end{bmatrix} \varepsilon_t = \mathbf{P} \varepsilon_t.$$

From the estimated VECM, it is possible to obtain the Wold representation

$$\mathbf{\Delta x}_t = \mathbf{C}(\mathbf{L}) \varepsilon_t$$

so that with

$$\varepsilon_t = \mathbf{P}^{-1} \theta_t$$

it is straightforward to identify the variance contribution of permanent and transitory shocks as well as impulse responses.[24]

[23] See Johansen (1991), Hoffmann (2001) and Gonzalo and Ng (2001)

[24] Note that the identification of the relative variance contributions of permanent and transitory shocks only requires knowledge of the (reduced-form) VECM parameters. The just-identifcation of the individual permanent and transitory shocks is not required. This will only be necessary once we are interested in conducting impulse response analysis. See e.g., Hoffmann (2001).

cpy **and bankruptcy filings**

FIGURE A.I:
The figure shows four-quarter growth rates in bankruptcy filings (dashed, red) and $cpy \times 10$ (blue, solid line). Data are obtained from the American Bankruptcy Institute at http://www.aib.org. Unfortunately, these are available only from 1980 onwards, so that a comparison with cpy in the early part of the sample is not possible. In addition, there seem to be changes in the definition of the AIB data that make it hard to interpret long time series of quarterly filings. Bankruptcy filings have generally trended downwards since 1980. Still there is a positive correlation with cpy at business cycle frequencies: bankruptcy filings are high when p is low, the correlation between the two lines in the figure is 0.23.