HCEO WORKING PAPER SERIES

Sectoral Labor Reallocation and Return Predictability

Esther Eiling
Raymond Kan
Ali Sharifkhani

Working Paper

2018-006

02/2018

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Sectoral Labor Reallocation and Return Predictability

Esther Eiling, Raymond Kan and Ali Sharifkhani[1]

February 2, 2018

Abstract

Sectoral labor reallocation shocks change the optimal allocation of workers across industries. We find that a proxy for this type of labor market shocks has very strong and robust predictive power for future stock market returns. In predictive regressions, the one-year out-of-sample R^2 is as high as 14.88%. We propose a production-based asset pricing model that links the return predictability to time-varying labor adjustment costs. When human capital is tied to the industry, hiring workers from other industries involves more search and training costs. Hence, sectoral reallocation shocks lead to lower returns to hiring and therefore lower future stock returns.

JEL Classification: G12, G17

Keywords: Financial Markets and the Macroeconomy, Return Predictability, Sectoral Shifts, Production-Based Asset Pricing

[1]Eiling is with the University of Amsterdam, Kan and Sharifkhani are with the Rotman School of Management, University of Toronto. Corresponding author: Esther Eiling, e.eiling@uva.nl. We thank Frederico Belo, Alexander David, Luis Garicano, Ralph Koijen, Francisco Santos, Mikhail Simutin, Xiaofei Zhao, conference participants at the 2017 UBC Winter Finance Conference, the 2016 meeting of the Northern Finance Association, the 2016 Australasian Finance and Banking Conference, the 2016 Financial Econometrics and Empirical Asset Pricing Conference at Lancaster University, and seminar participants at Amsterdam University and the University of Toronto for helpful discussions. Research support from the Social Sciences and Humanities Research Council is gratefully acknowledged.

Many of the fundamental shocks that drive economic growth also lead to the reallocation of workers across industries.[1] For instance, technological development and changes to consumer preferences can affect the match between wants and resources across different sectors in the economy (Black, 1995). This changes the optimal allocation of human capital across sectors. We show, both empirically and theoretically, that these shocks to the labor market lead to very strong and robust predictability of future stock market returns.

A well-known proxy for sectoral labor reallocation shocks is the cross-sectional dispersion in industry stock returns (e.g., Loungani et al., 1990; Brainard and Cutler, 1993; Loungani and Trehan, 1997). A higher dispersion in returns across industries implies a greater need for sectoral labor reallocation. We use the cross-sectional return volatility (CSV) of past 12-month industry-specific returns for 49 industries as a proxy for sectoral labor reallocation shocks. Extensive analysis displays the link between CSV and labor market conditions, which confirms the validity of this proxy. When comparing this measure to other stock return disperison measures, we find that the industry dimension is a key driver of its predictive power.

The predictive power of CSV is striking. We predict k-month ahead excess market returns, where $k = 1$, 3, 12, 24 and 36 months. First, the predictive regression coefficient is always significantly negative. Second, for one-quarter up to three-year ahead market returns, the in-sample and out-of-sample R^2s are substantially higher than those of eleven well-known alternative predictive variables, including three hiring-related variables from Chen and Zhang (2011).[2] For instance, when predicting one-year ahead market returns, CSV has an out-of-sample R^2 of 14.88%, while those of the alternative variables range between −17.97% (log net payout yield) and 1.00% (term spread). When comparing CSV to the recently introduced short interest index (SII) (Rapach et al., 2016), we find that SII outperforms at horizons shorter than 12 months, but this reverses for longer horizons. Although at the one-year horizon the out-of-sample R^2 of SII (13.03%) is only slightly below that of CSV, the difference increases for the 24-month horizon (19.90% for CSV versus 1.54% for SII). Finally, trading strategies based on CSV suggest that the predictive power of CSV is also economically important.

[1] As argued by, among others, Dixit and Rob (1994).

[2] The other alternative predictors are the log dividend-price ratio, the one-month T-bill rate, the log price-earnings ratio, the log net payout yield (Boudoukh et al., 2007), the default spread, the term spread, the inflation rate, and the consumption-wealth ratio (Lettau and Ludvigson, 2001).

Next, we show that labor adjustment costs serve as an economic channel that generates the observed stock return predictability. When labor adjustment is costly, a firm's market value includes the human capital of its workforce (Merz and Yashiv, 2007). As human capital is to some extent tied to the industry (e.g., Katz and Summers, 1989; Neal, 1995), search and training costs are likely to be higher when hiring from a different industry compared to hiring from within the same industry. In other words, when there is a greater need for sectoral labor reallocation, labor adjustment costs are higher, as argued by Weiss (1986). Consequently, a dollar invested in hiring a worker obtains less workforce, and hence lowers the return to hiring. Put differently, higher adjustment costs induced by sectoral labor reallocation shocks result in lower future returns.

We formalize this intuition in a production-based asset pricing model where sectoral reallocation shocks lead to time-varying and asymmetric labor adjustment costs. The asymmetry prevents workers from declining industries from instantaneously moving to expanding industries. We argue that industry labor adjustment costs are a function of the industry-specific productivity shock, leading to asymmetry across industries, and of the time-varying dispersion in productivity shocks across industries, which can be related to CSV. With a reasonable choice of parameters, we demonstrate numerically that labor adjustment costs and CSV are negatively related to future stock returns.[3]

Our empirical tests confirm the labor adjustment cost channel. We construct two labor skill-based CSV measures where we only include industries that rely mostly on high (low) skilled labor. Assuming that high skilled labor is costlier to adjust, the high skill CSV measure should have more predictive power for future stock returns. In line with this hypothesis, we show that the predictive power of CSV hinges completely on the industries that rely most on high skilled labor. In contrast, when we only include industries that depend more on low skilled labor, the predictive power of CSV disappears.

We perform a battery of tests to verify that CSV proxies for sectoral labor reallocation shocks. First, past industry idiosyncratic returns (measured with respect to the CAPM) help predict industry-level employment changes. Low performing industries tend to slightly lower their employment, while high performing industries significantly increase employment. However, when the

[3]Note that our proposed economic mechanism (and production-based model) links CSV to total rather than excess returns. However, we find that our results are quantitatively and qualitatively very similar when predicting total or excess returns. We focus on predicting excess market returns in order make our results comparable to existing studies on return predictability.

need for labor reallocation across sectors is high (i.e., when CSV is high), winner industries no longer significantly increase their employment. As a result, aggregate unemployment increases, which is exactly what we find: CSV predicts significantly higher aggregate unemployment growth, in line with the sectoral shifts hypothesis of Lilien (1982). The predictive power of CSV is even stronger for mismatch unemployment growth that is driven only by sectoral misalignment between unemployed workers and vacant job positions (Şahin et al., 2014). Furthermore, CSV strongly predicts Şahin et al.'s sectoral mismatch index that can be interpreted as an ex-post measure of sectoral labor reallocation. In sum, by linking CSV to several labor market variables, we confirm that it can be used as a proxy for sectoral labor reallocation shocks.

Several papers use other measures of cross-sectional return dispersion to predict market returns. Goyal and Santa-Clara (2003) use the dispersion in individual stock returns as a proxy for idiosyncratic volatility.[4] They find that it positively predicts future market returns. However, when we construct CSV using individual stock returns, we find negative or insignificant predictability, depending on the weighting scheme of CSV. This is in line with Bali et al. (2005) and Wei and Zhang (2005) who also find that the results are sensitive to the weighting scheme and sample period. Maio (2015) uses return dispersion of 100 size and book-to-market (BM) ranked portfolios and links it to heterogeneous beliefs.[5] We find that our industry-based CSV measure has stronger predictive power for future stock market returns than the size-BM based CSV, especially for longer horizons. Moreover, the size-BM based CSV does not predict future unemployment growth. This suggests that the predictive power of cross-sectional dispersion measures is, to a large extent, due to the industry dimension.[6]

We rule out three alternative explanations for the predictive power of our CSV measure. We first test whether CSV proxies for capital rather than labor adjustment costs. To this end, we use the industry-level asset redeployability index of Kim and Kung (2017) as a measure of the ease by which capital can be moved across industries. When we condition CSV on asset redeployability we find no effect on return predictability, which contrasts the capital adjustment cost channel. Second,

[4] See also Garcia et al. (2014).

[5] Stivers and Sun (2010) find that the latter measure positively predicts the value premium and negatively predicts the momentum premium. Other related papers are Jorgensen et al. (2012) and Kalay et al. (2014) who show that stock market returns help predict future dispersion in firm-level earnings growth.

[6] Loungani et al. (1991) suggest that their proxy for sectoral shifts negatively predicts future market returns. However, the analysis is very preliminary and the paper does not explain why there is a lagged response of stock market returns to sectoral shifts.

our predictability results at the industry level are inconsistent with investor over- or underreaction to industry-specific shocks. Third, our analysis using analyst coverage data allows us to reject slow information diffusion across industries (as in Hong et al. (2007)) as a potential explanation.

Two related papers also analyze the asset pricing consequences of labor adjustment costs. Merz and Yashiv (2007) first incorporate labor adjustment costs in the production-based asset pricing model of Cochrane (1991). Labor now becomes a "quasi-fixed" factor. Firms are compensated for the costs involved in hiring new workers and the resulting rents are included in the market value of the firm. This makes firms' hiring decisions forward looking, as argued by Belo et al. (2014). They show that hiring rates help predict firm-level stock returns in the cross-section. Instead, we focus on predictability in the time series by linking time-varying labor adjustment costs arising from sectoral labor reallocation shocks to future market returns.

The paper is structured as follows. Section I discusses the data. Section II explains the construction of CSV and verifies its use as a proxy for sectoral reallocation shocks. Our key empirical results on the predictive power of CSV for future market returns are presented in Section III. Section IV presents a production-based asset pricing model with asymmetric and time-varying labor adjustment costs, along with calibration and simulation results. Section V discusses robustness tests, which are presented in a separate Internet Appendix. Section VI concludes. The Appendix provides details on the labor market data used and the computational algorithm used to numerically solve our model.

I. Data

We use monthly returns on 49 industry portfolios in excess of the one-month T-bill rate to construct our main proxy for sectoral reallocation shocks, CSV. The data are from Kenneth French's website. The full sample period runs from January 1952 to December 2013, a total of 744 monthly observations. In addition, we evaluate the robustness of our results by performing our analysis over various subsample periods. As a proxy for the market portfolio we use monthly returns on the value-weighted CRSP market index. In robustness tests, we construct CSV based on monthly returns on 100 size and book-to-market ranked portfolios (also from French's website). Further, we construct CSV based on individual stock returns for all stocks traded on the NYSE, AMEX and

Nasdaq from CRSP.

We compare the performance of CSV to twelve alternative predictive variables that have been proposed in the literature. First, we use the log dividend-price ratio on the CRSP value-weighted market index (logDP), where dividends in a certain month are calculated as the sum of the past 12 months of dividends (following Fama and French, 1988). Next, we consider the yield on one-month T-bill relative to its previous 3-month moving average (RF). Third, we consider the cyclically adjusted log price-earnings ratio (logPE) from Robert Shiller's website. Fourth, we consider the log net payout yield (logNPY) from Boudoukh et al. (2007), which is provided by Michael Roberts. Fifth, we consider the default spread (DEF), calculated as the difference between the yield on Moody's Baa and Aaa rated corporate bonds. Next, we include the term spread (TERM), calculated as the yield difference between 10-year government bonds and 3-month T-bill rates. The seventh alternative predictor is the inflation rate (INFL), calculated as the log growth rate of the Consumer Price Index. The data for these latter three variables are from the Federal Reserve Bank in St. Louis. We also include the consumption-wealth ratio (CAY) from Lettau and Ludvigson (2001), which is provided by Martin Lettau. Following Vissing-Jørgensen and Attanasio (2003), we interpolate monthly values from quarterly values. This may give the monthly CAY measure some look-ahead bias, but as we will discuss later, it does not lead to an outperformance of this variable. In addition, we include three labor market related variables used by Chen and Zhang (2011): Payroll growth (PYRL), the Net hiring rate (NetHR) and the Net job creation rate in manufacturing (NetJC). As a final predictive variable we consider the short interest index (SII) of Rapach et al. (2016) from David Rapach's website. A number of alternative predictors are not available for the full sample period.[7] We perform the analysis for each variable based on the maximum number of observations available.

In addition, we use several labor market variables. Monthly aggregate US unemployment rates (UN) are from the Current Population Survey, provided by the Bureau of Labor Statistics (BLS). Following Loungani et al. (1990) we use a log transformation, where $un = \log(UN/(1-UN))$. We also consider short term (0–5 weeks) and long term (27+ weeks) unemployment rates (BLS Table A-12). Industry-level employment data are from the Current Employment Statistics survey. Lastly, to measure labor skill, we combine data from the BLS Occupational Employment Statistics,

[7]The logNPY ends in December 2010, TERM starts in April 1953, CAY starts in April 1952 and ends in September 2013, NetHR starts in March 1977, NetJC ends in May 2005 and SII starts in January 1973.

the Census Current Population Survey – Merged Outgoing Rotation Group and the Dictionary of Occupational Titles. Further details about this labor market data (including the predictive variables from Chen and Zhang (2011)) can be found in Appendix A.

II. Cross-Sectional Return Volatility and Sectoral Labor Reallocation

Following, among others, Loungani et al. (1990) and Brainard and Cutler (1993), we use the cross-sectional volatility (CSV) of industry-specific equity returns as a proxy for sectoral shifts.[8] The idea is as follows. If certain industries are affected by adverse shocks while others are hit by positive shocks, the industry-specific returns presumably incorporate these shocks instantaneously. Hence, the cross-sectional dispersion of industry returns increases. The increase in CSV reflects the mismatch between taste and technology across industries and induces a need for labor reallocation.

Lilien (1982) considers the cross-sectional dispersion in industry unemployment growth as a proxy for sectoral shifts.[9] However, Abraham and Katz (1986) show that this measure is more driven by aggregate demand shocks than by sectoral reallocation shocks. The advantage of using the dispersion in industry-specific stock returns as a proxy is that we can take out aggregate demand shocks by considering industry idiosyncratic returns. Also, while sectoral shifts are typically reflected in employment data with a lag, stock returns are expected to respond instantaneously. Another advantage of using a stock-return based proxy is that data is available at high frequencies and with a long history. In Section II.B we verify the validity of our proxy by linking CSV to various labor market variables.

A. Construction of the CSV Measure

Our CSV measure is based on industry-specific returns of 49 industries. Using industry-specific rather than total industry returns removes the effect of aggregate shocks which do not increase the need for sectoral reallocation. Following Brainard and Cutler (1993), we first run the following

[8]Throughout the paper, we use sectoral labor reallocation shocks and sectoral shifts interchangeably.

[9]Other employment-based measures are based on long term unemployment growth (Rissman, 1993) and the correlation between industry-level employment growth rates during and after a recession (Groshen and Potter, 2003).

regression using the data from the past 36 months:

$$R_{i,s} = \alpha_i + \beta_i R_{M,s} + \varepsilon_{i,s}, \quad s = t-35, \ldots, t, \tag{1}$$

where $R_{i,s}$ and $R_{M,s}$ are the month s excess returns of industry i and the market portfolio respectively, in excess of the 30-day T-bill rate. We then estimate the industry-specific returns for industry i at month s as its abnormal return from the CAPM, which is measured by

$$\eta_{i,s} = \hat{\alpha}_i + \hat{\varepsilon}_{i,s}, \tag{2}$$

where $\hat{\alpha}_i$ and $\hat{\varepsilon}_{i,s}$ are the OLS regression estimates of α_i and the fitted residuals obtained from (1). We then compute CSV at the end of month t as the cross-sectional standard deviation of the industry-specific returns from the past 12 months:[10]

$$CSV_t = \left[\frac{1}{48} \sum_{i=1}^{49} \left(\eta_{i,t-11:t} - \bar{\eta}_{t-11:t} \right)^2 \right]^{\frac{1}{2}}, \tag{3}$$

where

$$\eta_{i,t-11:t} = \prod_{s=t-11}^{t} (1 + \eta_{i,s}) - 1, \tag{4}$$

and

$$\bar{\eta}_{t-11:t} = \frac{1}{49} \sum_{i=1}^{49} \eta_{i,t-11:t}.$$

Our main CSV measure puts equal weights across the 49 industry-specific returns.[11] The impact of an industry-specific shock on the need for labor reallocation may depend on, for instance, the presence of unions in the industry (which makes layoffs more difficult), the presence of more industry-specific human capital (which makes labor less mobile), and the total employment in the industry (shocks in industries that are a large part of the labor market are expected to have a stronger effect on future aggregate unemployment). Most of these variables are unobserved, especially for a large cross-section of industries with a long history. While industry-level employment data is available, there are important limitations as discussed in Appendix A. Therefore, we use equal

[10] Sectoral reallocation shocks are permanent. Using a longer return horizon to calculate CSV helps to capture permanent shocks. In Section V, as a robustness test, we consider past 3-month and past 24-month industry-specific returns.

[11] At the beginning of the sample period, a few of the industry portfolios have missing returns, so the CSV is computed based on the industries with non-missing return data.

weights in our main CSV measure. Section V discusses a robustness test with employment-based weights.

[Insert Table I about here]

Table I reports summary statistics of CSV and the twelve alternative predictors for future equity market returns that we consider. CSV varies substantially over time; the average is 0.155 and the standard deviation is 0.044. The first and second order autocorrelation coefficients are 0.91 and 0.82 respectively, which is lower than those for most of the alternative predictors. The final column shows the correlation between the alternative predictors and CSV. Correlations range from −0.42 (log net payout yield) to 0.24 (logPE). While several correlations are significantly different from zero, the magnitudes are modest. This suggests that CSV captures a new aspect of equity return predictability compared to existing variables, which is confirmed by our empirical analysis in Section III.

B. CSV as a Proxy for Sectoral Shifts

We verify that CSV proxies for sectoral labor reallocation in five different ways. First, we show that industry-specific stock returns predict industry-level employment changes. Second, we show that CSV predicts aggregate unemployment growth. Third, we use CSV to predict a direct measure of the mismatch between job seekers and job vacancies across industries. Fourth, CSV predicts the part of unemployment growth that can be attributed to the sectoral mismatch between job opportunities and job seekers. Finally, in the robustness section (Section V) we show that modifications of the CSV measure that are less in line with the economic channel of sectoral shifts also weaken its predictive ability.

Table II shows the link between industry equity returns and subsequent employment changes at the industry level. This analysis helps us rule out the possibility that human capital and equity returns are inversely related, by which a negative shock to industry-level equity returns would increase the demand for labor in the industry.

[Insert Table II about here]

To this end, we use industry-level employment data that are available for 35 industries. We construct 35 industry equity portfolios for matched industry codes using CRSP individual stock data.

Each month, we sort industry equity portfolios into five quintiles, based on their past 12-month industry-specific returns. Then, we calculate the continuously compounded average employment growth for each quintile over the following k months. Throughout the paper we consider $k = 1$, 3, 12, 24 and 36 months. Table II panel A shows the results. We can see that the "loser" industries with the lowest past equity returns decrease employment, while the "winner" industries increase employment. The employment changes for expanding industries are statistically significant. The difference between winners and losers (WML) is always statistically significant and positive. These results show a link between past industry equity returns and subsequent industry-level employment changes, confirming results in Brainard and Cutler (1993) and Shin (1997).

However, these findings by themselves do not yet indicate that aggregate unemployment increases, as the workers that are laid off in low performing industries could be the ones that are hired immediately in the top performing industries. Table II Panel B reports the same analysis, except that we condition on months in which CSV is in the top 10% of all values of CSV over the sample period. In other words, these are months when labor adjustment costs are particularly high. Then, we calculate the average employment growth in subsequent months. While loser industries now significantly reduce their workforce, winner industries no longer significantly increase employment, except for $k = 3$. In other words, during times when labor adjustment costs are high (i.e., CSV is high), workers who are laid off are not immediately rehired. This indeed leads to higher aggregate unemployment, which we confirm next.

Figure 1 shows the time series of CSV as well as the time series of the aggregate unemployment rate (in levels). The shaded areas correspond to NBER recession dates.

[Insert Figure 1 about here]

We can see that CSV fluctuates substantially and while several peaks correspond to NBER recessions (e.g., 2008), others do not (e.g., 1966).

An expected consequence of sectoral labor reallocation shocks is an increase in future aggregate unemployment (e.g., Lilien, 1982; Şahin et al., 2014).[12] We explicitly test whether CSV has predictive power for aggregate unemployment changes. To this end, we run the following predictive

[12] The debate on the relative importance of sectoral reallocation versus aggregate demand shocks as drivers for unemployment is still ongoing (e.g., Groshen and Potter, 2003; Aaronson et al., 2004 – for an overview, see Gallipoli and Pelloni, 2013). Overall, while the impact of aggregate demand shocks may not be completely ruled out, the evidence suggests that sectoral shifts are an important determinant of aggregate unemployment growth.

regression

$$\Delta un_{t:t+k} = un_{t+k} - un_t = b_0 + b_1 CSV_t + \varepsilon_{t:t+k}, \tag{5}$$

where $\Delta un_{t:t+k}$ is the log unemployment growth from the end of month t to month $t+k$, as defined in Section II. Table III Panel A reports the results.

[Insert Table III about here]

The table shows the OLS estimates of b_1, Newey-West (1987) adjusted t-ratios (based on $k-1$ lags) and the R^2s. Our analysis confirms the sectoral shifts hypothesis: CSV predicts higher future aggregate unemployment growth. In line with our expectations, the coefficient estimate is positive and significant for all values of k. Also, the final column reports the contemporaneous correlation between unemployment growth and continuously compounded excess market returns. The correlation is negative (and statistically significant for $k > 3$), confirming that the stock market declines during periods of lower economic activity.

To further test the validity of CSV as a measure of sectoral reallocation shocks, we separately consider growth in short term unemployment rates (i.e., workers who are unemployed for a period between 0–5 weeks) and in long term unemployment rates (i.e., workers who are unemployed with a duration of more than 27 weeks). A key difference between the impact of aggregate demand shocks and sectoral shifts on unemployment growth is that the effect of aggregate demand shocks (i.e., business cycle changes) is temporary, while the effect of structural reallocation changes is permanent. Hence, we would expect long-term unemployment growth to be driven more by CSV, while short-term unemployment changes are more driven by aggregate demand shocks and less by CSV. This is confirmed by Panels B and C in Table III.[13]

The R^2s in Panel A do not exceed 5.35%, which suggests that besides sectoral shifts, other factors, including aggregate demand shocks, play a role in determining the aggregate unemployment. In our next analysis we therefore link CSV to unemployment measures that directly capture the misalignment of job seekers and job opportunities across sectors. To this end we first construct the mismatch index of Şahin et al. (2014), which measures the fraction of hires lost due to the job seeker misallocation across sectors. The mismatch index essentially is an ex-post measure of sectoral shifts. In contrast, CSV, which is based on industry returns, is an ex-ante measure of

[13] Blanchard and Diamond (1989), Brainard and Cutler (1993) and Loungani and Trehan (1997) show similar results.

sectoral shifts. We show that the two measures are strongly related. Table IV Panel A reports the results of the following predictive regression:

$$\mathcal{M}_{t+k} = b_0 + b_1 CSV_t + \varepsilon_{t+k}, \tag{6}$$

where $\mathcal{M}_{t+k}$ is the level of mismatch index at time $t+k$ assuming heterogeneity in labor productivity across industries, as discussed in Şahin et al. (2014). The analysis confirms the existence of a predictive relationship between CSV and the mismatch index. The coefficient associated with CSV is significant at 1% and 5% level for $k = 12$ and $k = 24$, respectively. Moreover, in terms of in-sample R^2, CSV demonstrates a strong predictive power with values up to 20.58% for $k = 12$.

[Insert Table IV about here]

Next, we examine if CSV can predict the component of the unemployment growth that is attributed to sectoral mismatch. Şahin et al. (2014) refer to this component as the mismatch unemployment. It is defined as the difference between the aggregate unemployment rate and a counterfactual unemployment rate where there is no impediment to the optimal allocation of job seekers across sectors. We repeat the exercise reported in Table III by replacing the aggregate unemployment rate with the mismatch unemployment rate. The results are reported in Table IV Panel B. We observe a considerable improvement in the predictive power of CSV for mismatch unemployment growth compared to aggregate unemployment growth. The coefficient associated with CSV is positive and significant for all horizons and the in-sample R^2 for this predictive regression ranges from 20.50% for $k = 36$ to 46.28% for $k = 12$.

In sum, the above findings show a strong link between CSV, future unemployment growth, future unemployment growth due to sectoral mismatch and an ex-post measure of sectoral shifts. These results validate CSV as a proxy for sectoral reallocation shocks.

III. Stock Market Return Predictability

We now turn to our main empirical analysis where we study the ability of CSV to predict future stock market returns. We hypothesize that an increase in CSV leads to higher labor adjustment costs as hiring workers from another industry is more costly than hiring from within the same industry. As a result, the return to a dollar invested in a firm's workforce is lower. Therefore, we expect that CSV negatively predicts future stock returns.

A. Predictive Regressions

We start by running the following predictive regression of the k-month excess return on the market:

$$r_{t:t+k} = \alpha + \beta z_t + \varepsilon_{t:t+k}, \tag{7}$$

where $r_{t:t+k} = r_{t+1} + \cdots + r_{t+k}$ is the continuously compounded excess return of the market from the end of month t to month $t+k$, and z_t is the value of a predictive variable observed at the end of month t. We calculate standard errors of the OLS estimates of α and β, following Hodrick (1992) as well as following Newey-West (1987) with $k-1$ lags. We use $k = 1$, 3, 12, 24 and 36 months.

Next, we test for out-of-sample predictability, following among others, Campbell and Thompson (2008). Using all returns up to month t with a minimum of 20 years of monthly data, we estimate the above regression. Then, we use the estimated parameters to construct a forecast of the k-month excess return from month t to month $t+k$:

$$\hat{r}_{t:t+k} = \hat{\alpha}_t + \hat{\beta}_t z_t, \tag{8}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are estimated using data from the beginning of the sample period to month t. In addition to reporting the in-sample R^2, we also report the out-of-sample R^2 for the predictive regressions using the historical average excess market return (calculated over all months up to t) as a benchmark. The out-of-sample R^2 is calculated as:

$$R^2_{OOS} = 1 - \frac{\sum_{t=240}^{T-k}(r_{t:t+k} - \hat{r}_{t:t+k})^2}{\sum_{t=240}^{T-k}(r_{t:t+k} - k\bar{r}_{1:t})^2}, \tag{9}$$

where $\bar{r}_{1:t}$ is the average excess market return computed using data up to month t, and T is the length of the return series. The summation is over all months for which returns are forecasted (i.e., starting in month 241). Note that the out-of-sample R^2 can be negative in case the predictive variable has poor out-of-sample predictive ability.[14]

[Insert Table V about here]

[14]Following suggestions by Campbell and Thompson (2008), we also try predicting simple arithmetic returns and restricting $\hat{r}_{t:t+k}$ to be positive (or else we replace the forecast by zero). Our results remain quantitatively and qualitatively similar.

Table V reports the results. Each panel is based on a different horizon k. First, we see that CSV negatively predicts future market returns for all k, which is in line with the labor adjustment cost channel. For $k = 1$ and $k = 36$, the regression coefficient is statistically significant at the 5% level, for all other k it is significant at the 1% level.

In-sample R^2s range from 0.80% for $k = 1$ to 20.45% for $k = 24$. The impressive performance of CSV extends to our out-of-sample analysis as well. Out-of-sample R^2s are all positive and range from 0.47% ($k = 1$) to 19.90% ($k = 24$). We should be careful however in comparing the results across different horizons k. For $k > 1$, longer horizon returns are overlapping. The increase in overlap for higher k could lead to an upward bias in the R^2. Hence, an increase of R^2 for higher k may simply be a statistical artifact rather than a sign of true improved performance. Therefore, in the following section we analyze trading strategies based on CSV. The utility gains from these trading strategies can be directly compared across horizons.

To put the predictive ability of CSV in perspective, we compare the performance of CSV to alternative predictive variables for a given k. The results are included in Table V. We first compare CSV to eleven well-known predictive variables (see Section I for a description), including three labor market related variables proposed by Chen and Zhang (2011). Then we specifically compare CSV to the short interest index (SII), which was recently proposed by Rapach et al. (2016) as "the strongest predictor of the equity risk premium identified to date." Given the impressive performance of SII, it poses the main hurdle for any new predictive variable.[15]

The first eleven alternative predictors do not always have significant coefficients. The log price-earnings ratio, the log net payout yield and the inflation rate are often insignificant based on Hodrick (1992) standard errors. The coefficient estimate of the default spread is insignificant in all cases. When $k = 1$, the demeaned risk-free rate, PYRL and NetJC have somewhat higher in-sample R^2s compared to CSV. Also, the demeaned risk-free rate has a slightly higher out-of-sample R^2. However, for $k > 1$, CSV shows an impressive outperformance compared to these eleven existing predictive variables. The differences are often remarkable. For example, when $k = 12$, CSV has

[15]We consider the Variance Risk Premium (VRP) as an additional predictive variable. Data (from Hao Zhou's website) is available only from January 1990 onwards, which leaves too few observations for an out-of-sample analysis. Based on in-sample analysis only, we find that the VRP and CSV excel at different horizons. Unreported results show that over the 1990–2013 period, the VRP has a higher in-sample R^2 at the monthly horizon (5.14% versus 1.96%), but at the three-month horizon the R^2s are similar (10.08% versus 9.86%). For longer horizons, CSV displays strong outperformance. For instance, for $k = 12$ the VRP has an R^2 of 3.42%, while CSV has an R^2 of 35.91%.

an in-sample R^2 of 14.24%. The alternative variables have R^2s ranging from 0.71% (NetHR) to 6.86% (NetJC). The out-of-sample performance of CSV is even better. At $k = 12$, CSV has an out-of-sample R^2 of 14.88%, while other variables have out-of-sample R^2s ranging between -17.97% (log NPY) and 1.00% (TERM). In fact, consistent with Welch and Goyal (2008), we find that the alternative variables often have negative out-of-sample R^2s, indicating their poor out-of-sample performance.

Based on the above horse race, CSV easily outperforms the eleven alterative predictors. However, the main hurdle is the short interest index proposed by Rapach et al. (2016). The bottom rows of panels A and B in Table V show that SII outperforms CSV for 1- and 3-month horizons in terms of the in-sample and out-of-sample R^2s. For example, for $k = 3$ the out-of-sample R^2 of SII is 7.23%, while that of CSV is 3.87%. However, CSV shows superior performance for horizons of 12 months and up. When $k = 12$, the out-of-sample R^2s are 13.03% (SII) and 14.88% (CSV). The difference is larger for $k = 24$ (1.54% for SII versus 19.90% for CSV) and for $k = 36$, when SII even has a slight negative out-of-sample R^2 of -2.93% versus 15.05% for CSV. The correlation between the two variables is low at 0.11, which suggests that they capture different aspects of stock return predictability.

[Insert Table VI about here]

We further examine how CSV stacks up against other variables in predicting excess market return by running a series of multiple regressions in which we use a subset of predictors as independent variables. That is,

$$r_{t:t+k} = \alpha + \sum_{i \in S} \beta_i z_{i,t} + \varepsilon_{t:t+k}, \tag{10}$$

where S represents the index of the subset of predictors used in multiple regression and $z_{i,t}$ is the forecasting variable i observed at the end of month t. Panels A and B of Table VI show the in-sample and out-of-sample R^2s respectively for each specification. We start with a specification where all of the above mentioned alternative predictors are included. Except for $k = 1$, adding CSV to the set of predictors improves the predictive power in terms of both the in-sample and out-of-sample R^2s. Not surprisingly, the out-of-sample R^2s for such specifications with many predictors are negative due to overfitting. Therefore, we suggest two parsimonious specifications in which CSV and SII, or CSV and PYRL are used as independent variables. With CSV and SII we obtain an out-of-sample R^2 of as much as 34.63% for $k = 12$. Using CSV and PYRL generates comparable results with an out-of-sample R^2 of 19.28% for $k = 12$ and 24.39% for $k = 24$. We present the estimated values of the parameters as well as the in-sample and out-of-sample R^2s for the recommended regression in Panel C. Using CSV, PYRL and SII as predictive variables in a multiple regression setting, we obtain an impressive out-of-sample R^2 of 36.26% for $k = 12$. It is important to note that the coefficient associated with CSV is largely unchanged compared to the univariate regression and it remains significant in all of these specifications.

In summary, the results in Tables V and VI suggest that CSV is an important variable for forecasting market returns, both by itself as well as in the presence of other predictive variables previously proposed in the literature.

B. Trading Strategies

In order to assess whether the superior predictive performance of CSV can actually translate into higher utility for an investor, we construct trading strategies based on CSV. An additional advantage is that we can now directly compare the utility in terms of certainty equivalents across different horizons k. Follow Rapach et al. (2010), we take the perspective of a mean-variance investor who allocates between the stock market portfolio and the risk-free asset. At each month t, the weight allocated to the market portfolio is determined by

$$\hat{w}_t = \frac{1}{\gamma} \frac{\hat{r}_{t:t+k}}{\hat{\sigma}^2_{t:t+k}}, \tag{11}$$

and $1 - \hat{w}_t$ is allocated to the risk-free asset. The coefficient of risk aversion is denoted by γ, for which we use a value of three following Rapach et al. (2010). $\hat{r}_{t:t+k}$ is the predicted k-month ahead continuously compounded excess market return at month t and is defined in (8). The forecast of the market excess return variance is denoted by $\hat{\sigma}^2_{t:t+k}$. We follow Campbell and Thompson (2008) and estimate it as k times the sample variance of the monthly excess market return over a rolling window of the past five years. As z_t we use our proxy for sectoral shifts (CSV) as well as the alternative predictors. To increase the power of our tests, the strategies that we examine include portfolios with overlapping periods when $k > 1$. Therefore, in any given month t, the strategies hold a series of portfolios that are selected in the current month as well as in the previous $k - 1$ months.

[Insert Table VII about here]

Table VII reports the results. Panel A shows, for different return horizons, the annualized sample mean and standard deviation of the excess return of portfolio strategies where CSV is used as the predictive variable. These are compared to a benchmark strategy where the market excess return is predicted using the historical monthly mean excess market return, i.e., $\bar{r}_{1:t}$. We find that the CSV-based trading strategies always lead to higher mean returns than the benchmark strategy, but also with slightly higher standard deviations (except for $k = 36$). For example, for $k = 12$, the CSV-based strategy leads to a mean return of 10.43% per annum and a standard deviation of 17.01%. In comparison, the benchmark strategy leads to a mean return of 3.99% and standard deviation of 15.79%.

To assess the economic and statistical significance of this outperformance and to compare a CSV-based strategy to strategies based on alternative predictors, we calculate the certainty equivalent (CE). The CE is calculated as

$$CE = \bar{R}_p - \frac{\gamma}{2}\hat{\sigma}_p^2, \tag{12}$$

where $\bar{R}_p$ and $\hat{\sigma}_p^2$ are the annualized sample mean and sample variance of the excess portfolio return associated with each trading strategy, calculated using the out-of-sample excess returns. We calculate the difference between the CE of the CSV-based trading strategy and the CE of the benchmark trading strategy, which is based on the historical mean. The CE difference provides a measure of additional risk-free return that the strategy earns compared to the benchmark strategy. Similarly, we calculate CE differences for strategies based on alternative predictors. We always use the historical mean-based strategy as a benchmark. To assess the statistical significance, we compute the t-ratio associated with the CE differences as derived in the Internet Appendix. Note that although the magnitude of the certainty equivalent depends on the coefficient of risk aversion (we use $\gamma = 3$), the t-ratio of the CE difference is independent of γ.

The results are reported in Panel B. The CSV-based strategy leads to a positive and significant CE difference for all horizons. The difference is economically important. For instance, for $k = 12$, the CSV-based strategy leads to an additional risk-free return of 5.83% per annum compared to the benchmark strategy. The CE differences are statistically significant at the 10% level for $k = 1$, at the 5% level for $k = 3$, and at the 1% level for $k = 12$, 24, and 36.[16] As we can directly compare CE across different horizons k, we conclude that the predictive power of CSV is economically more

[16] Note that this is a one-sided test.

important from one-quarter to one-year ahead market returns.

In comparison, ten out of the twelve alternative predictors do not show any CE differences that are positive and significant. Often, the estimated difference is negative, suggesting the strategy underperforms the benchmark strategy. The only exceptions are the strategies based on the term spread for $k = 24$ and 36 and based on the SII. Consistent with our predictive regression results, the short interest index is the only variable that outperforms CSV in terms of certainty equivalent for horizons up to one year. For longer horizons, however, CSV clearly dominates this variable in terms of certainty equivalent.

In sum, the strong predictive power of CSV for future market returns shown in predictive regressions is also economically meaningful. When using CSV in out-of-sample trading strategies, the portfolio performance is significantly improved.

C. The Role of Labor Adjustment Costs

We hypothesize that the observed predictive relationship between CSV and future stock market returns can be explained by costly labor adjustment. When CSV is high, more workers need to be reallocated between industries. Hiring workers from another industry rather than from the same industry involves more search and training costs. Higher labor adjustment costs decrease the future return on a firm's investment in its workforce and therefore lowers future stock returns. Before we formalize the intuition by deriving a production-based asset pricing model in Section IV, we first establish an empirical link between labor adjustment costs, CSV and future stock returns.

A natural prediction of our hypothesis is that the effect of CSV on future market returns should be relatively more pronounced when considering industries in which labor adjustment costs are intrinsically higher. Industries where the type of labor employed is generally easy to replace would find it comparably less costly to adjust their labor force during periods when the need for reallocation is high. We test this hypothesis by using industry-level labor skills as a proxy for the intrinsic labor adjustment costs. Filling vacancies for high skilled workers is expected to be costlier than filling vacancies for low skilled workers. As a result, a higher CSV would imply higher labor adjustment costs only when the type of labor employed in the industry is more difficult to replace.

To this end, we classify occupations into high and low skill based on the level of Specific Vocational Training (SVP) index required for the job, extracted from the Dictionary of Occupational Titles (DOT). This index serves as a proxy for the level of skill required for each occupation. Following Belo et al. (2017), we consider an occupation as being high skilled if the value of SVP is greater than six (corresponding to occupations requiring over two years of preparation), and low skilled otherwise. We define industry-level skill following Belo et al. (2017) as the percentage employment in high skill occupations in the industry. We also construct an alternative skill measure based on the ratio of total wages paid to the high skilled workers relative to the total wage expenditure in the industry. This is in line with the notion that wages better reflect the extent to which an industry depends on its skilled workers in the production process. Next, we identify the High Skill (HS) and Low Skill (LS) industries each year as those that belong to the highest and lowest terciles of industries in terms of the industry skill measure. We then construct CSV for each set of industries using the procedure explained in Section II.A with industries being ranked each year in June.

Table VIII shows the results of the predictive regression for future excess stock market returns using CSV_{HS}, which is the CSV constructed based on the High-Skill industries, and CSV_{LS} which is based on the Low-Skill industries. Since the labor skill data becomes available in 1990, we can perform this part of our analysis only in-sample. Panel A presents the results when the industry-level skill measure is defined based on the percentage of high-skilled workers in the industry. Consistent with our hypothesis, we observe that the predictive power of CSV is concentrated among industries that predominantly depend on high skill workers, i.e., industries that are likely to face higher labor adjustment costs. The predictive regression using CSV_{HS} generates a coefficient that is negative and highly statistically significant for all horizons. In sharp contrast, the coefficient associated with the CSV_{LS} is never significantly different from zero. For all k, we find that the in-sample R^2 is substantially higher for CSV_{HS} than for CSV_{LS}. For example, for $k = 12$, CSV_{HS} leads to an R^2 of 14.81%, while that of CSV_{LS} is only 2.09%. The difference between high skill and low skill CSV becomes even more evident once we include both as independent variables in a multiple regression setting. The estimated coefficients of CSV_{HS} are negative and significant at the 5% level for all horizons, while the coefficients for CSV_{LS} are not significantly different from zero. Panel B shows similar results when we define the industry-level skill measure as the percentage of wages associated with high skill worker.

[Insert Table VIII about here]

Overall, our results suggest that CSV predicts excess market returns only when labor is intrinsically costly to adjust due to the high level of skill that is required. This is in line with costly labor adjustment as a potential channel behind the observed predictive relationship between CSV and future market returns.

IV. A Production-Based Asset Pricing Model

We propose a production-based asset pricing model with time-varying labor adjustment costs that generate a link between CSV and future market returns. Our model belongs to the class of models that incorporate labor adjustment costs in a neoclassical framework, such as Merz and Yashiv (2007) and Belo et al. (2014).

Our setup is as follows. We assume that the economy consists of N industries, each of which is represented by a single representative firm. At the beginning of period t, each firm (industry) faces two types of productivity shocks: x_t, which is the aggregate productivity shock affecting all industries, and $z_{i,t}$, which is the idiosyncratic productivity shock affecting only industry i. We define $z_{i,t}$ as the product of two random variables:

$$z_{i,t} = S_t \cdot \tilde{z}_{i,t}, \tag{13}$$

where $S_t > 0$ represents reallocation shocks and therefore drives the cross-sectional dispersion in industry productivity. Each firm generates operating profits $Y_{i,t}$ according to a non-increasing return-to-scale production function

$$Y_{i,t} = f(x_t, z_{i,t}) N_{i,t}^{\alpha}, \quad 0 < \alpha \leq 1, \tag{14}$$

where α is the labor share in production and $N_{i,t}$ is the size of the firm's workforce. The dynamics of $N_{i,t}$ are determined by the firms' optimal hiring decisions. We assume that a firm's workforce has the following law of motion

$$N_{i,t+1} = (1 - \delta) N_{i,t} + H_{i,t}, \tag{15}$$

where δ is the total separation rate. Similar to Chen and Zhang (2011) we specify $H_{i,t}$ as the gross number of hires during period t. We follow among others, Merz and Yashiv (2007), Chen and Zhang (2011) and Belo et al. (2014) by assuming that hiring of workers is costly, due to for instance search costs and resources and time spent on training. Moreover, consistent with the labor search literature we assume that total labor adjustment costs have a quadratic functional form:

$$\frac{c_{i,t}}{2}\left(\frac{H_{i,t}}{N_{i,t}}\right)^2 N_{i,t}.$$

This specification has the desired property of being convex and increasing in the number of new hires and decreasing in the size of the firm's workforce, as suggested by intuition. However, while in Chen and Zhang (2011) the per-unit labor adjustment costs are constant, we allow $c_{i,t}$ to vary over time and across industries. Specifically, we assume that $c_{i,t}$ is a function of S_t and the normalized industry-specific productivity shock

$$c_{i,t} = \kappa_h S_t \Phi(\tilde{z}_{i,t}), \tag{16}$$

where κ_h is a constant and $\Phi(\cdot)$ is the standard normal CDF function.

Expression (16) allows us to incorporate sectoral shifts in the production-based asset pricing model. First, adjustment costs are industry-specific and asymmetric; $\Phi(\tilde{z}_{i,t})$ is higher for top performing industries that look to hire more workers after having received positive idiosyncratic productivity shocks, than for underperforming industries that tend to lay off workers. Note that $H_{i,t}$ is defined as the gross number of hires. Layoffs are incorporated by the total separation rate δ, which is exogenous. Therefore, an underperforming industry may decide to fire workers by hiring less than the total separation $\delta N_{i,t}$. This is similar to Chen and Zhang (2011). On aggregate, when hiring is more expensive than firing, workers who are laid off in losing industries will not be re-hired in winning industries instantaneously. As a result, future unemployment increases, which is the key mechanism of the sectoral shifts hypothesis (Lilien, 1982). Weiss (1986) emphasizes the importance of allowing for industry-specific productivity shocks and asymmetric labor adjustment costs when modelling the impact of sectoral shifts on future aggregate unemployment.

Second, (16) assumes that labor adjustment costs are time-varying and increasing in S_t. As mentioned earlier, a positive shock to S_t implies a higher dispersion across industries in terms of idiosyncratic productivity shocks, and potentially a higher need for the reallocation of labor across sectors. The labor adjustment costs firms face when hiring new workers limit worker mobility from contracting to expanding industries. A possible micro foundation for this is industry specificity of human capital. This in turn increases the hiring costs, mainly due to higher search costs resulting from a tighter labor market, as well as higher training costs when hiring from a pool of workers with unmatched skills.

Assuming that firms are purely equity financed, the firm's dividend will be equal to

$$D_{i,t} = Y_{i,t} - W_{i,t}N_{i,t} - \frac{c_{i,t}}{2}\left(\frac{H_{i,t}}{N_{i,t}}\right)^2 N_{i,t}, \tag{17}$$

where $W_{i,t}$ is the wage rate.

Each period, firms choose the number of workers to hire with a goal to maximize their discounted future cash flows, i.e., $D_t + E_t[\sum_{j=1}^{\infty} M_{t+j}D_{i,t+j}]$, where M_{t+j} is the stochastic discount factor at time $t + j$. As a result, we can write the firms' maximization problem as

$$\max_{H_{i,t+j},N_{i,t+j+1}} E_t\left[D_{i,t} + \sum_{j=1}^{\infty} M_{t+j}\left[f(x_{t+j}, z_{i,t+j})N^{\alpha}_{i,t+j} - W_{i,t+j}N_{i,t+j} - \frac{c_{i,t+j}}{2}\left(\frac{H_{i,t+j}}{N_{i,t+j}}\right)^2 N_{i,t+j}\right]\right] \tag{18}$$

subject to (15).

We obtain the firm's optimal hiring decision by numerically solving this optimization problem. As such, the firms' optimal hiring decisions are used to determine the corresponding optimal dividend and market value, which are in turn the determinants of the firm (industry) stock returns.

A. A Special Case

Prior to numerically solving the firm's optimization problem, we find it instructive to consider a special case of this problem where we analytically show the link between the cross sectional dispersion in productivity shocks and future stock market returns. Here, α is assumed to be equal to 1, implying a production function with constant returns to scale. In this special case, we further assume that wages are exogenous. However, in the next section where we numerically solve a general version of the model, we allow wages to be endogenously determined.

Denoting $q_{i,t}$ as the Lagrange multiplier, the first order conditions will be

$$q_{i,t} = c_{i,t}\frac{H_{i,t}}{N_{i,t}}, \tag{19}$$

and

$$q_{i,t} = E_t\left[M_{t+1}\left[f(x_{t+1}, z_{i,t+1}) - W_{i,t+1} + \frac{c_{i,t+1}}{2}\left(\frac{H_{i,t+1}}{N_{i,t+1}}\right)^2 + (1-\delta)q_{i,t+1}\right]\right]. \tag{20}$$

These first order conditions are the same for each two consecutive decision periods. Equating the two equations, we obtain

$$E_t\left[M_{t+1}\left[\frac{f(x_{t+1},z_{i,t+1})-W_{i,t+1}+\frac{c_{i,t+1}}{2}\left(\frac{H_{i,t+1}}{N_{i,t+1}}\right)^2+(1-\delta)q_{i,t+1}}{c_{i,t}\frac{H_{i,t}}{N_{i,t}}}\right]\right]=1, \tag{21}$$

which is the Euler equation for the hiring return:

$$R^H_{i,t+1}\equiv\frac{f(x_{t+1},z_{i,t+1})-W_{i,t+1}+\frac{c_{i,t+1}}{2}\left(\frac{H_{i,t+1}}{N_{i,t+1}}\right)^2+(1-\delta)q_{i,t+1}}{c_{i,t}\frac{H_{i,t}}{N_{i,t}}}. \tag{22}$$

As pointed out by Cochrane (1991), in a complete market, a firm can construct a mimicking portfolio for its hiring return. As a result, it will adjust its workforce until the mimicking portfolio return equals the hiring return. In other words, the equity return for industry i, $R_{i,t+1}$ equals its hiring return, $R^H_{i,t+1}$. Hence, by plugging equations (15), (16) and (19) into (22), we have the following expression for equity returns in industry i:

$$R_{i,t+1}=\frac{1}{S_t}\left[\frac{f(x_{t+1},z_{i,t+1})-W_{i,t+1}+\frac{c_{i,t+1}}{2}\left(\frac{N_{i,t+2}}{N_{i,t+1}}\right)^2-\frac{c_{i,t+1}}{2}(1-\delta)^2}{\kappa_h Q(\tilde{z}_{i,t})\left(\frac{N_{i,t+1}}{N_{i,t}}-(1-\delta)\right)}\right], \tag{23}$$

which links the cross-sectional dispersion in productivity shocks, S_t, to future stock returns in industry i. Averaging over all industries using their corresponding market capitalization as weights $(w_{i,t})$ we obtain

$$R_{m,t+1}=\frac{1}{S_t}\left[\sum_{i=1}^{N}w_{i,t}\frac{f(x_{t+1},z_{i,t+1})-W_{i,t+1}+\frac{c_{i,t+1}}{2}\left(\frac{N_{i,t+2}}{N_{i,t+1}}\right)^2-\frac{c_{i,t+1}}{2}(1-\delta)^2}{\kappa_h Q(\tilde{z}_{i,t})\left(\frac{N_{i,t+1}}{N_{i,t}}-(1-\delta)\right)}\right]. \tag{24}$$

It is important to note that $R_{i,t+1}$ ($R_{m,t+1}$) is the gross stock return and therefore is always positive. As a result, considering that S_t is always positive, the second ratio on the right hand side is also always positive. This implies that, ceteris paribus, there is an inverse relation between the reallocation shocks and future stock market returns as well as future returns of each industry. Our simulation results in the next section show that the cross-sectional dispersion in industry-specific productivity shocks S_t and the cross-sectional dispersion in industry-specific stock returns CSV are highly correlated. Given these findings, (24) suggests that CSV can predict future market excess returns with negative sign, which is consistent with our empirical results in Section III. To gain more economic insights into why the relationship between CSV and future returns is negative, note that the denominator in (23) is the marginal cost of hiring an additional worker. With high adjustment costs, a dollar invested in hiring a worker obtains less workforce, and hence lower return to investment compared to a case where adjustment costs are low. In other words, higher adjustment costs lead to lower future returns.

A potential concern about the analytical solution above is that the current and future employment growth, which are included in the second ratio on the right hand side, are endogenous. We address this concern in two ways. First, we use CSV to predict industry returns for individual industry portfolios. Hence, we perform 49 predictive regressions, one for each of the 49 industry portfolios. The results for $k = 12$ are reported in the Internet Appendix. We find that CSV predicts industry returns with a negative sign for 48 out of 49 industries.[17] These empirical findings suggest that regardless of the changes in current and future workforce, the relationship between CSV and future stock returns is always negative. Hence, it logically extends to future market returns as well, which is what we show in our main empirical analysis. Second, we solve the model numerically.

B. The General Model Setup

In the general model that we solve numerically, α can differ from one and wages are determined endogenously. First, we parametrize the production function as

$$f(x_t, z_{i,t}) = e^{x_t + \kappa_p z_{i,t}}, \tag{25}$$

where κ_p is a constant. We assume that the aggregate and standardized idiosyncratic productivity shocks follow AR(1) processes as

$$x_t = \rho_x x_{t-1} + \sigma_x \epsilon_t^x, \tag{26}$$

$$\tilde{z}_{i,t} = \rho_{\tilde{z}} \tilde{z}_{i,t-1} + \sigma_{\tilde{z}} \epsilon_{i,t}^{\tilde{z}}, \tag{27}$$

where ϵ_t^x and $\epsilon_{i,t}^{\tilde{z}}$ are i.i.d. standard normal random variables, and they are independent of each other. We impose $\sigma_{\tilde{z}} = \sqrt{1 - \rho_{\tilde{z}}^2}$ to guarantee that the unconditional variance of $\tilde{z}_{i,t}$ is equal to one. We also define $s_t = \log(S_t)$ as the log of the reallocation shocks, evolving as an AR(1) process:

$$s_t = \rho_s s_{t-1} + \sigma_s \epsilon_t^s, \tag{28}$$

[17] The coefficient estimate is statistically significant for 33 or 38 industries (based on either the Hodrick (1992) or Newey-West (1987) adjusted t-ratios) and most industries have substantial in-sample and out-of-sample R^2s.

where ϵ_t^s is i.i.d. standard normal, and independent of ϵ_t^x and $\epsilon_{i,t}^{\tilde{z}}$. This in turn implies that the cross-sectional standard deviation of the idiosyncratic productivity shocks $z_{i,t}$ is proportional to S_t.

Similar to Kuehn et al. (2017), we assume that wages are the outcome of a Nash bargain between workers and the firm in an individual bargaining process. Workers are assumed to receive unemployment benefit b in case of a breakdown in their negotiations with the firm. The rent associated with the employment of a worker is split between the worker and the firm in accordance with the worker's bargaining weight $\eta \in (0, 1)$, which is assumed to be exogenous. This rent consists of the savings in hiring costs when the position is filled, and the surplus associated with the marginal product of labor. In addition, hiring the marginal worker lowers wage payments to the firm's current labor force. Hence, wages can be written as

$$W_{i,t} = \eta \left[\frac{\alpha}{1 - \eta(1 - \alpha)} \frac{Y_{i,t}}{N_{i,t}} + c_{i,t} \frac{H_{i,t}}{N_{i,t}} \right] + (1 - \eta)b. \tag{29}$$

The effect of the decreasing return to scale of the production function is reflected in the term $\alpha/(1 - \eta(1 - \alpha))$, while the second term in bracket represents marginal adjustment costs.

Following Berk et al. (1999) and Zhang (2005), we specify an exogenous pricing kernel without explicitly modeling the consumer's problem. The log of the stochastic discount factor is specified as

$$\log M_{t+1} = \log\beta + \lambda(x_t - x_{t+1}), \tag{30}$$

in which $\lambda > 0$ is the constant price of risk, and $0 < \beta < 1$ is the time discount factor. Our choice of the pricing kernel, which implies a constant price of risk, is intended to maintain consistency with our focus on the production side of the economy. Effectively, it serves our purpose to pin down the effect of labor adjustment costs as the channel that induces return predictability while shutting down alternative consumption-based channels that potentially generate return predictability through time series variations in the price of risk. We acknowledge, however, that the assumption of a constant price of risk can potentially lead to the failure of our model to precisely capture a number of the empirical regularities that are documented in the literature.

C. Model Calibration and Simulation Results

We solve the firm's optimization problem numerically by applying the value function iteration procedure on a discritized state space. A detailed explanation of this procedure is provided in

Appendix B. We simulate 500 panels each containing 50 industries and 900 months. Using the firm's optimal hiring and dividend policies, we extract firm (i.e., industry) returns which aggregate to market returns. We can now construct CSV using the procedure described in Section II.A. We test the predictive relationship between CSV and excess market returns using simulated data.

C.1. Calibration of Parameters

Our benchmark model is calibrated according to the values reported in Table IX. We adhere to the literature in our calibration of the parameters that have been given a value in existing studies. For additional free parameters we make our choices in such a way to achieve an approximate match between the aggregate moments generated by our model and those implied by historical data.

[Insert Table IX about here]

The first set of parameters concerns those related to the technology of the firms. We set the total separation rate to $\delta = 3\%$, roughly the same as the separation rate identified by Davis et al. (2006). This includes the rates associated with both layoffs and voluntary quits by workers. We also set $\eta = 0.72$ as the workers' bargaining power, following among others Shimer (2005).

The model assumes that workers obtain a utility of b from being unemployed. This would be the outside option from rejecting a job offer, which would in turn affect wages. Shimer (2005) argues that this option is limited to the unemployment benefits received by the worker, implying $b = 0.4$. On the other hand, Hagedorn and Manovskii (2008) suggest that home production and leisure obtained by the individual as a result of being unemployed should be included in the unemployment activities, and hence $b = 1$. We take a relatively conservative stance by setting $b = 0.51$. This could be interpreted as the higher value of a riskless flow of monetary compensation in the form of unemployment benefits compared to other forms of utility obtained by the unemployed individual. Finally, the output elasticity of labor, α, is set to 0.3 following Morales-Jimenez (2017).

Next, we calibrate parameters related to pricing kernel as well as the stochastic processes that drive the price. We set $\rho_x = 0.95$ and $\sigma_x = 0.0137$ following among others Bai (2016). The parameters $\beta = 0.997$ and $\lambda = 1$ are set to match the first two moments of the risk-free rate implied by the data. The values of $\rho_{\tilde{z}} = 0.965$ is chosen close to Zhang (2005), while $\sigma_{\tilde{z}} = 0.262$ is set to make the standard deviation of $\tilde{z}_{i,t}$ equal to one.

Since the parameters associated with the reallocation shocks are not directly observable, we set them is such a way to provide us with a mean and standard deviation for CSV as close to those implied by the data as possible. To this end we set $\rho_s = 0.96$ and $\sigma_s = 0.2$. Finally, we set $\kappa_p = 0.443$ in order to make the standard deviation of the idiosyncratic component in the production function close to 0.58, which is the value used in Belo et al. (2017) and is based on their parameter values for the stochastic process associated with the idiosyncratic productivity.

We verify our calibration by comparing a set of moments based on artificial data generated by the calibrated model with those implied by the data. The results are reported in Table X. Overall, the model does a reasonable job at generating aggregate moments close to the values suggested by the data, implying that the calibrated model captures important dynamics of the stock market.[18] In addition, our simulations show a correlation of 76% between the reallocation shocks S_t and CSV. This supports our argument in Section IV.A where the implicit assumption was that the predictive relationship between S_t and excess market returns implies a predictive relationship between CSV and excess market returns.

[Insert Table X about here]

C.2. CSV and Future Stock Returns

In order to assess the extent to which we capture the effect of labor adjustment costs in our model, we run a predictive regression of market excess return on CSV based on artificial data. To this end, we follow the procedure outlined in Section III.A for $k = 1$, 3, 12, 24 and 36. We define the market return as the value-weighted average of the industry stock returns. The results are provided in Table XI.

[Insert Table XI about here]

The simulation-based results are in line with our empirical findings. The coefficient associated with CSV is negative over all horizons. The coefficient estimate is statistically significant for 12, 24 and 36-month horizons. The magnitudes of the coefficients and the R^2s in our simulated tests are smaller than those of the empirical analysis. A potential explanation is that CSV is more volatile in our calibrated model compared to the data.

[18] Note that the observed mismatch between some of the moments generated by our model and those derived from the actual data can be due to our simplifying assumption that the price of risk is constant.

Our model provides us with a new testable implication as well. The adjustment cost is specified as an increasing function of the idiosyncratic productivity shock. This implies that labor adjustment costs are higher for the top performing industries (with high $\tilde{z}_{i,t}$) than for the bottom performing industries (with lower $\tilde{z}_{i,t}$). As a result, in addition to predicting market returns, we expect that CSV also predicts industry returns, where the predictive power for each industry is positively related to its past performance. This implies a higher coefficient associated with CSV for the outperforming industries compared to other industries.

We test this model prediction by first creating quintile portfolios from our set of industries based on their past performance. Specifically, each month, we sort industries into quintiles based on their past 12-month idiosyncratic returns (estimated using (4)). Then, for each quintile, we calculate the equally weighted excess return over the following k months. This gives us a time-series of returns for each of the quintile portfolios (i.e., past losers, 2, 3, 4 and past winners). Next, we use CSV to predict the continuously compounded excess returns on each of these five quintiles. This is similar to (7), except that we use the quintile industry excess returns as the dependent variable instead of market excess return.

The results in Table XII support this hypothesis. Panel A shows the test results based on simulated data. While industries in all performance quintiles have a negative CSV coefficient, the coefficient associated with the winner industry is significantly more negative compared to those in the loser industry. In other words, higher industry-level labor adjustment costs are associated with higher CSV coefficient in this predictive regression. This model prediction is validated when we perform this test on actual data as well. The results presented in Panel B, which are based on actual data, show that the coefficient estimate of CSV for past winners is larger in magnitude than that of past losers in all cases. Further, past winners typically have higher R^2s than past losers.

[Insert Table XII about here]

Summarizing, we incorporate sectoral shifts in a production-based asset pricing model by allowing for time-varying and asymmetric labor adjustment costs. We show that this generates a link between CSV and future stock returns in the model. Additional empirical tests show that the predictive relation is negative and significant for individual industries, regardless of whether they expand or reduce their workforce. Further, the model implies a stronger relation for top performing industries, which is confirmed in the data.

V. Further Discussion

A. Alternative Explanations

We examine three alternative explanations for the negative predictive power of CSV for future market returns. First, we test whether capital adjustment costs rather than labor adjustment costs are the first-order impediment for firms to respond to sectoral reallocation shocks. Firms can face adjustment costs when trying to increase their capital stock, and this can reduce their investment returns and hence stock returns. We investigate this alternative explanation by applying a similar approach as in Subsection III.C using a measure for the difficulty at which an industry can make adjustments to its capital stock. Industries that rely on highly industry-specific physical assets are expected to find it more costly to adjust their capital, especially when the need for the reallocation of capital is high due to sectoral reallocation shocks. On the other hand, industries that mostly rely on assets that can be redeployed across a wide range of industries should experience a less severe increase in their capital adjustment costs during these periods. Therefore, under this alternative explanation, we expect that the predictive power of CSV would diminish by the extent to which industries rely on redeployable assets. We test this hypothesis using the industry-level asset redeployability index constructed by Kim and Kung (2017) as a measure of the industries' ability to obtain physical capital from other industries. Hence, this measure is negatively related to the industries' intrinsic capital adjustment cost. In each month, we construct CSV_{HR} as the CSV that is based on industries that belong to the highest tercile of industries in terms of their redeployability index. Similarly, we construct CSV_{LR} using industries in the lowest tercile in terms of redeployability index.[19]

[Insert Table XIII about here]

Table XIII compares the predictive power of the CSV constructed using high-redeployability industries with that of the low-redeployability industries. Due to the shorter time period over which asset redeployability data is available, we focus on in-sample predictive regressions. Panel A reports

[19] We define industries using the same BEA-based industry classification used Kim and Kung (2017). We exclude industries that consist of fewer than five firms at each time to alleviate the effect of firm-specific shocks on CSV, and to maintain a high correlation between the CSV based on the BEA industry classification and the one based on 49 Fama-French industry classification.

the results when the asset-level redeployability score is constructed using industry market capitalization as weights, while Panel B shows the results when the score is simply based on the number of industries that use the asset. The results do not show a notable difference between the predictive power of the CSVs based on the two types of industries. This contradicts the alternative hypothesis that capital adjustment costs generate the predictive power of CSV for future market returns. The results in Gavazza (2011) provide a possible rationale for this finding. If firms optimally hold on to their most industry-specific assets in response to adverse profitability shocks, they can simply adjust the capacity utilization of their assets in place in response to sectoral reallocation shocks. This makes their capital adjustment costs less sensitive to sectoral reallocation shocks.

Second, we investigate the role of investor over- or underreaction to industry-specific shocks. If investors overreact to good news in expanding industries, or they underreact to bad news in declining industries at time t, the returns presumably will be corrected one period later. Hence, according to this hypothesis, at time $t+1$ both expanding and contracting industries experience lower returns, which may drive the negative response of the market at $t+1$. An important implication of the over- or underreaction hypothesis is that only extreme industries will respond at time $t+1$, as opposed to all industries. However, our results in Table XII Panel B show that CSV negatively predicts future equity returns for virtually all industries, not just a few extreme past winners or past loser industries. As a result, the over- or underreaction does not seem to be a plausible explanation for the observed return predictability.

A third alternative explanation is based on slow information diffusion. According to the sectoral shifts hypothesis, sectoral shocks signal lower future economic activity. In order to assess the full extent of the sectoral reallocation that these shocks bring about and the effect on future aggregate unemployment, investors need to consider all industries. However, investors often specialize in one or a few industries only (Hong et al., 2007). If information diffuses slowly across industries, the full effect of the sectoral reallocation shocks will be incorporated in other industry returns and in market returns with a delay. However, our finding that the predictability is significant even for long horizons (up to three years) is challenging for a slow information diffusion explanation.[20] Further, when we interact CSV with an aggregate measure of analyst coverage,[21] we do not find

[20]The cross-asset predictability identified by among others, Hong et al. (2007), Cohen and Frazzini (2008) and Menzley and Ozbas (2010) is typically confined to the one-year horizon.

[21]We construct a monthly measure of average analyst coverage for the firms in each industry. We then take the

that the predictive ability of CSV increases during times when analyst coverage is lower. These results contrast a slow information diffusion explanation.

B. Robustness Tests

Below we discuss various robustness tests, which are reported in the Internet Appendix.

CSV based on different horizons

First, our results are robust when using past three-month and past 24-month (rather than past 12-month) industry idiosyncratic returns to calculate CSV. Similar to our main measure, these two alternative CSV measures negatively predict lower future excess market returns. At the same time, significance of the coefficients and the in-sample and out-of-sample R^2s are somewhat lower than for the 12-month CSV. Further, they also positively predict unemployment growth up to one year ahead.

CSV based on different weighting schemes

Second, we use different weighting schemes to calculate CSV. Our main measure equally weights all industries. As a first alternative, we use employment-based weights. This is in line with our economic interpretation of CSV as a proxy for sectoral shifts. When an industry is hit by a shock, the subsequent need for labor reallocation is expected to be higher when that industry has a large share of the labor market. Unfortunately, detailed industry-level employment data at a monthly frequency for a long sample period are not available. We use employment data for 35 industries (see Appendix A), ending in April 2003. We construct value-weighted industry equity portfolios for matched industries using individual stock returns from the CRSP. Consistent with our base results, the employment weighted CSV predicts future equity market returns with a negative sign. The coefficient estimate is significant for all $k > 1$. The equally weighted CSV has stronger predictive power, which may stem from the fact that these 35 industries are less well balanced. Unreported results show that the average employment share varies from 0.12% (tobacco) to 23.89% (services). Also, we should keep in mind that other factors may play a role as well, such as the presence of unions and the specificity of human capital in the industry. Further, when the employment weighted CSV measure is based on past 3-month returns, it significantly predicts future unemployment growth.

equally weighted average of the analyst coverage of the top and bottom 20 percentile of industries ranked in terms of past 12 months idiosyncratic return. Results are reported in the Internet Appendix.

Our second alternative weighting scheme uses market capitalization-based weights. Note that there is less economic reason to weight sectoral reallocation shocks by the equity market capitalization of the industry. In fact, employment-based weights can be quite different from value weights. For instance, unreported results show that the average employment-based weight of retail trade is 17.56%, while the average value weight is only 4.96%. We use the original set of 49 industry equity portfolios over the full sample period to construct CSV with value weights. Similar to the equally weighted CSV, the value-weighted CSV significantly predicts future stock market returns with a negative sign. However, in-sample and out-of-sample R^2s are lower. Further, the value-weighted CSV has predictive ability for future unemployment growth, but here again, the results are weaker than for the equally weighted CSV. In sum, while our results on stock market predictability are robust to using market capitalization based weights, the equally weighted CSV is a better proxy for sectoral shifts and has stronger predictive ability for market returns.

CSV based on total industry returns

Several existing papers use cross-sectional return dispersion measures to forecast market returns (e.g., Goyal and Santa-Clara, 2003, Garcia et al., 2014). They typically find that their measures, which are interpreted as proxies for aggregate idiosyncratic risk, positively predict future market returns. There are two main differences with our empirical measure. First, we use idiosyncratic returns rather than total returns. Second, we use industry returns rather than individual stock returns. To further compare our results to these papers, we first construct CSV based on total industry returns. A CSV measure based on total returns is a less suitable proxy for sectoral reallocation shocks, because total returns are affected by aggregate demand shocks as well. Using idiosyncratic returns allows us to control for these aggregate demand shocks. Indeed, CSV based on total industry returns does not significantly predict future unemployment growth. It still has predictive power for future stock market returns, but in contrast to Goyal and Santa-Clara (2003) and Garcia et al. (2014), the sign is always negative. This is consistent with our main analysis.

CSV based on individual stock returns

Next, we use individual stocks to construct CSV. To be more comparable with existing papers, we again use total returns in this analysis. We use both equal weights and market capitalization based weights. The equally weighted individual stock return CSV measure has no predictive ability for future equity market returns, while the market capitalization based measure has some significant coefficients, mainly for larger k. Importantly, the coefficient estimate is always negative, similar to our base results. The lack of significant positive coefficients is in line with Bali et al. (2005) and Wei and Zhang (2005). Further, we find that these alternative CSV measures predict lower rather than higher unemployment growth.

CSV based on other portfolio returns

In our next robustness test, we use idiosyncratic returns on 100 size and book-to-market ranked portfolios to construct CSV, similar to Maio (2015). We find that it predicts future market returns with a negative sign, but the industry-based CSV measure outperforms, especially for higher k. Moreover, the resulting CSV measure fails to predict unemployment growth.

The above analyses highlight that the industry dimension of return dispersion measures is a key driver of their predictive power for future stock market returns. The robustness tests also show that using a better proxy for sectoral shifts (with more economic motivation) leads to better predictability of equity returns and aggregate unemployment growth. This is in line with our proposed economic channel through which CSV predicts equity market returns, namely sectoral labor reallocation.

Sub sample periods

Finally, we perform our analysis for three separate subsample periods. First, we start in January 1973 (rather than 1952), which is comparable to many existing studies on return predictability. We find an even stronger link between CSV and future stock market returns with higher in-sample and out-of-sample R^2s than for the full sample period. Also, the predictive power of CSV for future unemployment growth is substantially stronger over this period. Next, we split our sample in two halves: from January 1952 to December 1982 and from January 1983 to December 2013. The corresponding results confirm that the strongest predictive power for both stock market returns and unemployment growth occurs in the later part of the sample period. This part of our sample period includes the 1991–1992 and 2001 crises, during which sectoral shifts played an important role according to Groshen and Potter (2003).

To put these findings in perspective, it is important to note that the relative importance of sectoral reallocation shocks compared to aggregate demand shocks as a driver for unemployment likely changes over time. During times when labor is more mobile or when the economy is doing well, the effect of reallocation shocks is expected to be weaker (e.g., Davis, 1987).

VI. Conclusion

This paper proposes a new variable that helps predict future stock market returns: the cross-sectional volatility of industry-specific stock returns (CSV). We find that increases in CSV strongly and robustly predict lower future market returns. Importantly, the predictive ability translates into significant utility improvements when we use CSV in a trading strategy. In addition, CSV substantially outperforms a large number of well-known alternative predictors of stock market returns.

CSV has a clear economic interpretation. Following, among others, Loungani et al. (1990) and Brainard and Cutler (1993), we show that CSV serves as a proxy for sectoral labor reallocation shocks. As such, it can be linked to time-varying labor adjustment costs. When the cross-sectional dispersion in industry returns is large, there is a greater need for reallocation of workers from low performing to high performing industries. Due to limited labor mobility (for instance, as a result of industry-specific human capital), reallocation takes time and resources. In other words, labor adjustment costs increase. As a result, the return to a dollar invested in a firm's workforce is lower. In sum, sectoral labor reallocation shocks proxied by CSV induce higher labor adjustment costs, which lead to lower future returns.

Consistent with this economic mechanism, we find that the predictive ability of CSV depends on those industries that rely the most on high skilled labor, which is costlier to adjust. Furthermore, we show that our empirical results are in line with a production-based asset pricing model with time-varying and asymmetric labor adjustment costs.

Appendix

A. Labor market data

Our list of alternative predictive variables includes, among others, three labor market variables from Chen and Zhang (2011): payroll growth (PYRL), the Net hiring rate (NetHR) and the Net job creation rate in manufacturing (NetJC). PYRL is defined as the log growth rate of the monthly seasonally adjusted total nonfarm payrolls of all employees, from the BLS. NetHR is defined as the gross hiring rate minus separation rate. We obtain the data from 1977Q1 to 2002Q4 from Merz and Yashiv (2007). Starting 2001Q1 we use data from the Jobs Opening and Labor Turnover Survey (JOLTS). We follow Chen and Zhang (2011) and scale the second series using the ratio between the average of the series during the overlapping period, i.e., 2001−2002. NetJC is defined as the difference between the job creation rates and job destruction rates from John Haltiwanger's website. We convert the two quarterly variables into monthly values by setting the values each month equal to the most recent value available in that month.

Part of the analysis uses monthly industry-level employment data for a set of 35 industries from the Current Employment Statistics survey. We use total nonfarm payroll employees per industry. In 2003, the CES industry classification changes. As there is no one-to-one match between the old and the new industry classification, we end our sample for this part of our analysis in May 2003. All employment data are seasonally adjusted.

For the sectoral mismatch index and mismatch unemployment rate, we follow the procedure proposed by Şahin et al. (2014) and construct the time series of the two variables for the period between January 2001 and December 2013. The codes for these two variables and the series themselves are available on the AEA website up to 2011. We update the data until the end of our sample period in December 2013.

In our test for the link between market return predictability and the level of labor skill, we combine three datasets. First, the BLS Occupational Employment Statistics (OES) tracks employment across industries for each occupation in the economy from 1988 onwards. Following Donangelo (2014), for each year prior to 1996, we combine data from the previous three years to ensure continuity in industry coverage in our data. This restricts our sample period for this test to 1990 and after. The dataset also contains median hourly wage estimates for each industry-occupation starting in 1997. Prior to 1997, we obtain average hourly wage estimates from the Census Current Population Survey – Merged Outgoing Rotation Group (CPS-MORG).[22] Next, the resulting dataset is matched with the Dictionary of Occupational Titles (DOT), which contains score for attributes of a wide range of occupations in the economy. To this end, we link each nine-digit DOT occupation code with the occupation codes used in the OES using the 2000 Standard Occupation Code (SOC) as the reference occupation classification. We use the crosswalk tables provided by the National Crosswalk Service Center and the mapping tables used by Belo et al. (2017).

B. Computational Algorithm

We start by discretizing the state space (x, $\tilde{z}$, s) into a grid consisting of nine grid points for x, seven grid points for z, and seven grid points for s. To this end, we utilize Rouwenhorst's (1995) method by which the states follow a Markov chain with finite states. We use this method in our simulation because each of the above processes have a persistence level greater than 0.9. We also specify a grid consisting of 50 points for the labor stock N with lower bound $\underline{N} = 0.01$ and upper bound $\bar{N}$=25. The distance between grid points is determined using the recursive procedure suggested by McGrattan (1999). Similarly, for the choice variable N', which represents the optimal employment $N_{i,t+1}$ in Equation (15), we specify a log-linear grid of 5000 over the same interval.

We solve the firm's maximization problem (18) on each grid point using the value function iteration procedure. Having obtained the optimal policies over each grid point, we construct a simulated path for each state variable for a panel of 50 firms, each representing an industry. We find the value functions and the corresponding hiring decisions on the simulation paths that are off the grid points using linear interpolation. We first neutralize the effect of the initial conditions defined for the state variables by running a simulation over 10,000 months, by which we obtain the stationary cross-sectional distribution of the idiosyncratic productivity shocks and that of the optimal values of the choice variables. The end values generated by this procedure are then fed as the initial values to our main simulation procedure, in which firm (industry) stock returns are calculated based on the optimal dividend and the firm value over 900 months.

[22] For each industry-occupation in the OES we use the corresponding earnings-weighted average hourly wage of individuals aged 18 to 65 in its matching broad industry-occupation observations in CPS-MORG. For the industry-occupations with no CPS-MORG match, we use the nationwide hourly wage for that occupation. We follow the BLS regarding industry classifications and use two-digit SIC codes until 2001 and three-digit NAICS codes afterwards.

References

[1] Aaronson, Daniel, Ellen R. Rissman, and Daniel G. Sullivan, 2004, Can sectoral reallocation explain the jobless recovery? *FRBC Economic Perspectives* 2Q, 36–49.

[2] Abraham, Katherine G. and Lawrence F. Katz, 1986, Cyclical unemployment: sectoral shifts or aggregate disturbances? *Journal of Political Economy* 94, Part 1, 507–522.

[3] Bai, Hang, 2016, Unemployment and credit risk, Working Paper, University of Connecticut.

[4] Bali, Turan, Nusret Cakici, Xuemin Yan, and Zhe Zhang, 2005, Does idiosyncratic risk really matter? *Journal of Finance* 60, 905–929.

[5] Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch, 2014, Labor hiring, investment, and stock return predictability, *Journal of Political Economy* 122, 129–177.

[6] Belo, Frederico, Xiaoji Lin, Jun Li, and Xiaofei Zhao, 2017, Labor-force heterogeneity and asset prices: the importance of skilled labor, *Review of Financial Studies* 30, 3669–3709.

[7] Berk, Jonathan, Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

[8] Black, Fisher S., 1995, *Exploring General Equilibrium*, MIT Press, Massachusetts.

[9] Blanchard, Olivier and Peter Diamond, 1989, The Beveridge curve, *Brookings Papers on Economic Activity* 1989–1, 1–76, reprinted in *The Economics of Unemployment*, P. N. Junankar editor, Edward Elgar Publisher.

[10] Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael Roberts, 2007, On the importance of measuring payout yield: implications for empirical asset pricing, *Journal of Finance* 62, 877–915.

[11] Brainard, S. Lael and David M. Cutler, 1993, Sectoral shifts and cyclical unemployment reconsidered, *The Quarterly Journal of Economics* 108, 219–244.

[12] Campbell, John Y. and Samuel B. Thompson, 2008, Predicting excess stock returns out of sample: can anything beat the historical average? *Review of Financial Studies* 21, 1509–1531.

[13] Chen, Long and Lu Zhang, 2011, Do time-varying risk premiums explain labor market performance? *Journal of Finance Economics* 99, 385–399.

[14] Cochrane, John, 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–237.

[15] Cohen, Lauren and Andrea Frazzini, 2008, Economic links and predictable returns, *Journal of Finance* 63, 1977–2011.

[16] Davis, Steven J., 1987, Allocative disturbances and specific capital in real business cycle theories, *American Economic Review* 77, 326–332.

[17] Davis, Steven J., R. Jason Faberman, and John Haltiwanger, 2006, The flow approach to labor markets: New data sources and micro-macro links, *Journal of Economic Perspectives* 20, 3–26.

[18] Dixit, Avinash and Rafael Rob, 1994, Switching costs and sectoral adjustments in general equilibrium with uninsured risk, *Journal of Economic Theory* 62, 48–69.

[19] Donangelo, Andres, 2014, Labor mobility: Implications for asset pricing, *Journal of Finance* 69, 1321–1346.

[20] Fama, Eugene F. and Kenneth French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

[21] Gallipoli, Giovanni and Guanluigi Pelloni, 2013, Macroeconomic effects of job reallocations: a survey, *Review of Economic Analysis* 5, 127–176.

[22] Garcia, Rene, Daniel Mantilla-Garcia, and Lionel Martellini, 2014, A model-free measure of aggregate idiosyncratic volatility and the prediction of market returns, *Journal of Financial and Quantitative Analysis* 49, 1133–1165.

[23] Gavazza, Alessandro, 2011, The role of trading frictions in real asset markets. *American Economic Review* 101, 1106–1143.

[24] Goyal, Amit and Pedro Santa-Clara, 2003, Idiosyncratic risk matters! *Journal of Finance* 58, 975–1008.

[25] Groshen, Erica L. and Simon Potter, 2003, Has structural change contributed to a jobless recovery, *Current Issues in Economics and Finance* 9, 1–7.

[26] Hagedorn, Marcus and Iourii Manovskii, 2008, The cyclical behavior of equilibrium unemployment and vacancies revisited, *American Economic Review* 98, 1692–1706.

[27] Hodrick, Robert J., 1992, Dividend yields and expected stock returns: alternative procedures for inference and measurement, *Review of Financial Studies* 5, 257–286.

[28] Hong, Harrison, Walter Torous, and Rossen Valkanov, 2007, Do industries lead stock markets? *Journal of Financial Economics* 83, 367–396.

[29] Jorgensen, Bjorn, Jing Li, and Gil Sadka, 2012, Earnings dispersion and aggregate stock returns *Journal of Accounting and Economics* 53, 1–20.

[30] Kalay, Alon, Suresh Nallareddy, and Gil Sadka, 2014, Conditional earnings dispersion, the macroeconomy and aggregate stock returns, Working Paper, Columbia University and the University of Texas at Dallas.

[31] Katz, Lawrence F. and Lawrence H. Summers, 1989, Industry rents: evidence and implications, *Brookings Papers on Economic Activity. Microeconomics* 1989, 209–290.

[32] Kim, Hyunseob and Howard Kung, 2017, The asset redeployability channel: how uncertainty affects corporate investment. *Review of Financial Studies* 30, 245–280.

[33] Kuehn, Lars-Alexander, Mikhail Simutin, and Jessie Jiaxu Wang, 2017, A labor capital asset pricing model, *Journal of Finance* 72, 2131–2178.

[34] Lettau, Martin and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.

[35] Lilien, David M., 1982, Sectoral shifts and cyclical unemployment, *Journal of Political Economy* 90, 777–793.

[36] Loungani, Prakash and Bharat Trehan, 1997, Explaining unemployment: sectoral vs. aggregate shocks, *FRBSF Economic Review* 1, 3–15.

[37] Loungani, Prakash, Mark Rush, and William Tave, 1990, Stock market dispersion and unemployment, *Journal of Monetary Economics* (June), 367–388.

[38] Loungani, Prakash, Mark Rush, and William Tave, 1991, Stock market dispersion and business cycles, *FRBC Economic Perspectives* (January), 2–8.

[39] Maio, Paulo, 2015, Cross-sectional return dispersion and the equity premium, *Journal of Financial Markets* 29, 87–109.

[40] McGrattan, Ellen R., 1999, Application of weighted residual methods to dynamic economic models, in Ramon Marimon and Andrew Scott, eds. *Computational Methods for the Study of Dynamic Economies* (Oxford University Press, Oxford).

[41] Menzley, Lior and Oguzhan and Ozbas, 2010, Market segmentation and cross-predictability of returns, *Journal of Finance* 65, 1555–1580.

[42] Merz, Monika, and Eran Yashiv, 2007, Labor and the market value of the firm, *American Economic Review* 97, 1419–1431.

[43] Morales-Jimenez, Camilo, 2017, The cyclical behavior of unemployment and wages under information frictions, Working Paper, University of Maryland.

[44] Neal, Derek, 1995, Industry-specific human capital: evidence from displaced workers, *Journal of Labor Economics* 13, 653–677.

[45] Newey, Whitney and Kenneth West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 59, 817–858.

[46] Rapach, David E., Matthew C. Ringgenberg, and Guofu Zhou, 2016, Short interest and aggregate stock returns, *Journal of Financial Economics* 121, 46–65.

[47] Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: combination forecasts and links to the real economy, *Review of Financial Studies* 23, 821–862.

[48] Rissman, Ellen, 1993, Wage growth and sectoral shifts: Phillips curve redux, *Journal of Monetary Economics* 31, 395–416.

[49] Rouwenhorst, K. Geert, 1995, Asset pricing implications of equilibrium business cycle models, in Thomas F. Cooley, ed. *Frontiers of Business Cycle Research*, 294–330.

[50] Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante, 2014, Mismatch unemployment, *American Economic Review* 104, 3529–3564.

[51] Shimer, Robert, 2005, The cyclical behavior of equilibrium unemployment and vacancies, *American Economic Review* 95, 25–49.

[52] Shin, Kwanho, 1997, Inter- and intrasectoral shocks: effects on the unemployment rate, *Journal of Labor Economics* 15, 376–401.

[53] Stivers, Chris and Licheng Sun, 2010, Cross-sectional return dispersion and time variation in value and momentum premiums, *Journal of Financial and Quantitative Analysis* 45, 987–1014.

[54] Vissing-Jørgensen, Annette, and Orazio P. Attanasio, 2003, Stock-market participation, intertemporal substitution, and risk-aversion, *American Economic Review Papers and Proceedings* 93, 383–391.

[55] Wei, Steven and Chu Zhang, 2005, Idiosyncratic risk does not matter: a re-examination of the relationship between verage returns and average volatilities, *Journal of Banking and Finance* 29, 603–621.

[56] Weiss, Laurence, 1986, Asymmetric adjustment costs and sectoral shifts, *Essays in Honor of Kenneth J. Arrow*: Volume 2, Equilibrium Analysis, Chapter 9, Cambridge University Press.

[57] Welch, Ivo and Amit Goyal, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies*, 21, 1455–1508.

Table I. Summary Statistics

The table reports summary statistics for 13 potential predictors of market returns. CSV is the equally weighted cross-sectional volatility of 49 industry returns, based on the continuously compounded past 12-month industry idiosyncratic returns, where industry idiosyncratic returns are calculated as the sum of the intercept and residuals of the CAPM, estimated over the past 36 months. The other twelve predictors are: the log dividend-price ratio (logDP), the one-month T-bill rate, which is demeaned by its three-month moving average (RF), the log price earnings ratio (logPE), the log net payout yield (logNPY), the default premium (DEF), calculated as the difference between the yields on Baa and Aaa rated corporate bonds, the term spread (TERM), calculated as the difference between the yields of 10-year government bonds and three-month Treasury bills, the inflation rate (INFL), based on the change in the consumer price index (all items, urban), the consumption-wealth ratio (CAY), payroll growth (PYRL), net hiring rate (NetHR), net job creation in manufacturing (NetJC), and short interest (SII). All data are at the monthly frequency. The sample period covers January 1952 to December 2013 but six predictors are available for a shorter sample period: logNPY ends in December 2010, TERM starts in April 1953, CAY starts in April 1952 and ends in September 2013, SII starts in January 1973, NetHR starts in March 1977, and NetJC ends in May 2005. The table reports the mean, median, standard deviation, first and second order autocorrelation coefficients (AR(1) and AR(2)), the minimum and maximum values of each predictor and its correlation with CSV (corr. CSV).

	Mean	Median	Std. Dev.	AR(1)	AR(2)	min	max	corr. CSV
CSV	0.155	0.145	0.044	0.909	0.821	0.077	0.389	1.000
logDP	−3.544	−3.498	0.389	0.991	0.981	−4.547	−2.843	−0.317
RF	0.00%	0.00%	0.07%	0.406	0.028	−0.48%	0.42%	−0.039
logPE	2.868	2.910	0.398	0.995	0.988	1.890	3.790	0.239
logNPY	−2.196	−2.144	0.211	0.980	0.963	−3.235	−1.700	−0.423
DEF	0.97%	0.85%	0.45%	0.970	0.922	0.32%	3.38%	0.100
TERM	1.47%	1.43%	1.20%	0.956	0.893	−2.65%	4.42%	−0.136
INFL	0.29%	0.25%	0.31%	0.615	0.482	−1.79%	1.79%	0.144
CAY	0.03%	0.06%	1.77%	0.991	0.973	−3.88%	3.44%	−0.123
PYRL	0.14%	0.18%	0.20%	0.926	0.849	−0.59%	0.72%	−0.138
NetHR	0.22%	0.18%	0.33%	0.816	0.632	−0.50%	1.50%	−0.101
NetJC	−0.22%	−0.07%	1.22%	0.899	0.798	−4.22%	4.60%	−0.228
SII	0.000	−0.110	1.000	0.942	0.901	−2.514	2.907	0.112

Table II. Industry Stock Returns and Future Industry-level Employment Changes

This table shows a link between industry-level equity returns and industry-level employment changes. Each month, we sort 35 portfolios into quintiles based on their past 12-month industry-specific returns. Then, we record the industry-level employment growth for these industries for the following k months, calculated as the continuously compounded average employment growth over all industries within the quintile. The sample period is from January 1952 to April 2003. Panel A reports an unconditional analysis and Panel B conditions on months in which CSV is in the top 10% (taken over the full sample period). Below the average industry-level employment changes, we report the corresponding Newey-West (1987) adjusted t-ratios based on $k-1$ lags. The final column WML reports the difference between the top quintile (past winners) and bottom quintile (past losers). ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Unconditional Analysis

		Loser	2	3	4	Winner	WML
$k=1$	ΔEmpl. (%)	−0.05	−0.04	0.06	0.05	0.13	0.17
	t-ratio	(−0.73)	(−0.92)	(1.78)*	(1.76)*	(3.68)***	(2.56)**
$k=3$	ΔEmpl. (%)	−0.20	−0.09	0.21	0.14	0.30	0.50
	t-ratio	(−1.16)	(−0.79)	(2.38)**	(1.84)*	(3.94)***	(3.15)***
$k=12$	ΔEmpl. (%)	−0.66	−0.02	0.36	0.50	0.86	1.52
	t-ratio	(−1.26)	(−0.04)	(0.90)	(1.45)	(2.50)**	(4.15)***
$k=24$	ΔEmpl. (%)	−0.91	0.11	0.53	0.97	1.30	2.21
	t-ratio	(−0.99)	(0.16)	(0.78)	(1.57)	(1.96)**	(3.78)***
$k=36$	ΔEmpl. (%)	−0.68	0.61	1.25	1.53	1.40	2.08
	t-ratio	(−0.56)	(0.66)	(1.40)	(1.68)*	(1.51)	(2.52)**

Panel B: Conditioning on CSV Being in the Top 10%

		Loser	2	3	4	Winner	WML
$k=1$	ΔEmpl. (%)	0.09	−0.17	−0.04	−0.12	0.02	−0.06
	t-ratio	(0.25)	(−2.44)**	(−0.74)	(−1.21)	(0.19)	(−0.17)
$k=3$	ΔEmpl. (%)	−0.29	−0.57	−0.24	−0.32	0.50	0.78
	t-ratio	(−0.37)	(−2.18)**	(−0.97)	(−1.17)	(2.16)**	(0.99)
$k=12$	ΔEmpl. (%)	−1.95	−1.68	−1.25	−0.79	0.47	2.43
	t-ratio	(−1.69)*	(−1.61)	(−1.04)	(−0.74)	(0.47)	(2.45)**
$k=24$	ΔEmpl. (%)	−3.47	−2.49	−1.85	−1.03	−0.34	3.13
	t-ratio	(−2.08)**	(−1.34)	(−0.88)	(−0.57)	(−0.22)	(3.44)***
$k=36$	ΔEmpl. (%)	−3.52	−1.59	−0.07	0.10	0.17	3.69
	t-ratio	(−1.32)	(−0.70)	(−0.03)	(0.04)	(0.09)	(1.36)

Table III. Predicting Aggregate Unemployment Growth Using CSV

This table reports results of the following predictive regression of k-month ahead aggregate unemployment growth on CSV,

$$\Delta un_{t:t+k} = b_0 + b_1 CSV_t + \varepsilon_{t:t+k},$$

where $\Delta un_{t:t+k} = un_{t+k} - un_t$, and un_t is based on a log transformation of the unemployment rate: $un_t = \log(UN_t/(1 - UN_t))$. The second column ($\hat{b}_1$) reports estimates of b_1 and the next column (t-ratio) gives the Newey-West (1987) adjusted t-ratios of $\hat{b}_1$, based on $k - 1$ lags. The fourth column presents the regression's in-sample R^2 and the final column reports the correlation between the log excess stock market returns and unemployment growth, where both are taken over the same k-month horizon. Panel A is based on total unemployment, while Panels B and C are based on long term (27+ weeks) and short term (0–5 weeks) unemployment respectively. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth				
k	$\hat{b}_1$	t-ratio	R^2_{IS}	$\text{corr}(r_{M,t:t+k}, \Delta un_{t:t+k})$
1	0.080	(2.46)**	0.90%	−0.018
3	0.229	(2.20)**	1.75%	−0.071
12	0.889	(1.88)*	3.48%	−0.237**
24	1.529	(2.43)**	5.35%	−0.370***
36	1.423	(1.94)*	3.68%	−0.460***
Panel B: Predicting Long Term Unemployment Growth (27+ weeks)				
k	$\hat{b}_1$	t-ratio	R^2_{IS}	$\text{corr}(r_{M,t:t+k}, \Delta un_{t:t+k})$
1	0.235	(2.93)***	1.24%	−0.001
3	0.612	(2.97)***	2.66%	0.062
12	2.169	(2.18)**	4.35%	0.021
24	3.791	(2.56)**	6.04%	−0.153
36	4.112	(2.35)**	5.54%	−0.289**
Panel C: Predicting Short Term Unemployment Growth (0–5 weeks)				
k	$\hat{b}_1$	t-ratio	R^2_{IS}	$\text{corr}(r_{M,t:t+k}, \Delta un_{t:t+k})$
1	0.011	(0.19)	0.01%	0.007
3	0.051	(0.66)	0.10%	−0.108**
12	0.173	(0.82)	0.37%	−0.341***
24	0.258	(0.89)	0.65%	−0.400***
36	0.156	(0.50)	0.20%	−0.397***

Table IV. Predicting Sectoral Mismatch Index and Mismatch Unemployment Growth Using CSV

This table reports results of predictive regression for labor market-based proxies of the sectoral mismatch between unemployed workers and job vacancies. Panel A reports the result of the following predictive regression

$$\mathcal{M}_{t+k} = b_0 + b_1 CSV_t + \varepsilon_t,$$

where $\mathcal{M}_{t+k}$ is the sectoral mismatch index based industries with heterogenous productivity. Panel B reports the results of the following predictive regression of k-month ahead mismatch unemployment growth on CSV,

$$\Delta un_{\mathcal{M},t:t+k} = b_0 + b_1 CSV_t + \varepsilon_{t:t+k},$$

where $\Delta un_{\mathcal{M},t:t+k} = un_{\mathcal{M},t+k} - un_{\mathcal{M},t}$, and $un_{\mathcal{M},t}$ is based on a log transformation of the mismatch unemployment rate: $un_{\mathcal{M},t} = \log(UN_{\mathcal{M},t}/(1-UN_{\mathcal{M},t}))$. The second column ($\hat{b}_1$) reports estimates of b_1 and the next column (t-ratio) gives the Newey-West (1987) adjusted t-ratios of $\hat{b}_1$, based on 12 lags for Panel A, and $k-1$ lags for Panel B. The fourth column presents the regression's in-sample R^2. The sample period covers January 2001 to December 2013. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Sectoral Mismatch Index			
k	$\hat{b}_1$	t-ratio	R^2_{IS}
1	0.029	(0.91)	1.22%
3	0.046	(1.50)	3.04%
12	0.121	(3.44)***	20.58%
24	0.072	(2.10)**	9.15%
36	0.056	(1.29)	5.02%
Panel B: Predicting Sectoral Mismatch Unemployment			
k	$\hat{b}_1$	t-ratio	R^2_{IS}
1	0.743	(4.16)***	31.04%
3	1.886	(2.92)***	34.12%
12	3.995	(4.72)***	46.28%
24	4.569	(4.86)***	36.69%
36	3.860	(2.38)**	20.50%

Table V. Predicting Stock Market Returns Using CSV and Alternative Predictors

The table reports results of the following predictive regression:

$$r_{t:t+k} = \alpha + \beta z_t + \varepsilon_{t:t+k},$$

where $r_{t:t+k}$ is the continuously compounded k-month excess return on the market from month t to month $t+k$. We estimate predictive regressions for the following predictive variables (z_t): the proxy for sectoral shifts (CSV), the log dividend price ratio (logDP), the de-meaned risk-free rate (RF), the log price earnings ratio (logPE), the log net payout yield (logNPY), the default spread (DEF), term spread (TERM), inflation rate (INFL), the consumption-wealth ratio (CAY), payroll growth (PYRL), net hiring rate (NetHR), net job creation in manufacturing (NetJC), and short interest index (SII). The sample period covers January 1952 to December 2013 but five predictors are available for a shorter sample period: logNPY ends in December 2010, TERM starts in April 1953, CAY starts in April 1952 and ends in September 2013, SII starts in January 1973, NetHR starts in March 1977, and NetJC ends in May 2005. The table reports the regression coefficient estimate ($\hat{\beta}$), the corresponding t-ratios of Hodrick (1992) and Newey-West (1987) with $k-1$ lags, as well as the in-sample and out-of-sample R^2s. The five panels show results for $k = 1, 3, 12, 24$ and 36 months. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: $k = 1$

	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
CSV	-0.089	$(-2.23)^{**}$	$(-2.26)^{**}$	0.80%	0.47%
logDP	0.008	$(1.70)^{*}$	$(1.72)^{*}$	0.46%	-0.55%
RF	-7.280	$(-2.23)^{**}$	$(-2.40)^{**}$	1.23%	0.65%
logPE	-0.006	(-1.25)	(-1.25)	0.26%	-1.05%
logNPY	0.011	(1.18)	(1.21)	0.30%	-1.47%
DEF	0.340	(0.77)	(0.77)	0.12%	-0.60%
TERM	0.303	$(2.09)^{**}$	$(2.10)^{**}$	0.68%	-0.56%
INFL	-1.215	$(-2.01)^{**}$	$(-1.98)^{**}$	0.77%	-0.67%
CAY	0.152	$(1.77)^{*}$	$(1.78)^{*}$	0.38%	-0.08%
PYRL	-2.177	$(-2.77)^{***}$	$(-2.92)^{***}$	1.05%	-0.23%
NetHR	0.353	(0.63)	(0.63)	0.07%	-0.15%
NetJC	-0.377	$(-2.98)^{***}$	$(-3.14)^{***}$	1.14%	0.05%
SII	-0.006	$(-2.73)^{***}$	$(-2.86)^{***}$	1.56%	2.17%

Table V - continued

Panel B: $k = 3$					
	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
CSV	−0.378	(−3.06)***	(−3.95)***	4.36%	3.87%
logDP	0.025	(1.87)*	(2.29)**	1.49%	−1.52%
RF	−11.786	(−1.93)*	(−1.98)**	0.97%	0.40%
logPE	−0.018	(−1.37)	(−1.64)	0.82%	−2.92%
logNPY	0.032	(1.13)	(1.50)	0.77%	−4.10%
DEF	1.119	(0.87)	(0.89)	0.40%	−2.20%
TERM	0.810	(1.88)*	(2.12)**	1.47%	−1.62%
INFL	−2.715	(−1.62)	(−1.65)*	1.15%	−1.99%
CAY	0.519	(2.00)**	(2.19)**	1.33%	−0.24%
PYRL	−6.191	(−2.70)***	(−3.06)***	2.55%	−0.82%
NetHR	−2.053	(−1.15)	(−1.37)	0.68%	−4.23%
NetJC	−1.005	(−2.73)***	(−3.03)***	2.50%	−0.32%
SII	−0.019	(−2.84)***	(−3.22)***	4.98%	7.23%

Panel C: $k = 12$					
	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
CSV	−1.420	(−3.17)***	(−4.77)***	14.24%	14.88%
logDP	0.105	(2.01)**	(2.33)**	6.13%	−10.57%
RF	−27.192	(−2.37)**	(−2.13)**	1.20%	−1.43%
logPE	−0.084	(−1.61)	(−1.93)*	4.12%	−11.33%
logNPY	0.178	(1.68)*	(2.40)**	5.76%	−17.97%
DEF	4.055	(1.00)	(1.21)	1.22%	−3.54%
TERM	3.354	(2.15)**	(2.60)***	5.76%	1.00%
INFL	−11.002	(−2.04)**	(−2.52)***	4.35%	−0.42%
CAY	2.132	(1.98)**	(2.22)**	5.08%	0.33%
PYRL	−20.274	(−2.58)***	(−3.11)***	6.35%	−1.31%
NetHR	−4.202	(−0.79)	(−0.92)	0.71%	−7.57%
NetJC	−3.418	(−2.66)***	(−2.43)**	6.86%	0.30%
SII	−0.068	(−2.44)**	(−2.68)***	13.50%	13.03%

Table V - continued

Panel D: $k = 24$					
	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
CSV	-2.296	$(-3.07)^{***}$	$(-4.92)^{***}$	20.45%	19.90%
logDP	0.198	$(1.93)^{*}$	$(2.30)^{**}$	11.82%	-27.39%
RF	-24.740	(-1.55)	$(-1.75)^{*}$	0.55%	-1.51%
logPE	-0.159	(-1.54)	$(-1.95)^{*}$	8.00%	-26.34%
logNPY	0.365	$(1.83)^{*}$	$(2.91)^{***}$	12.80%	-25.64%
DEF	2.792	(0.39)	(0.55)	0.31%	-4.94%
TERM	5.006	$(1.93)^{*}$	$(3.00)^{***}$	7.18%	12.88%
INFL	-12.281	(-1.28)	$(-2.64)^{***}$	2.94%	-0.16%
CAY	4.170	$(1.86)^{*}$	$(2.46)^{**}$	10.34%	4.37%
PYRL	-21.257	$(-1.69)^{*}$	$(-2.17)^{**}$	3.84%	0.10%
NetHR	2.789	(0.35)	(0.43)	0.18%	-4.05%
NetJC	-1.467	(-0.77)	(-0.75)	0.70%	-0.82%
SII	-0.060	(-1.15)	(-1.29)	6.14%	1.54%

Panel E: $k = 36$					
	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
CSV	-2.423	$(-2.53)^{**}$	$(-3.76)^{***}$	17.24%	15.05%
DP	0.259	$(1.71)^{*}$	$(2.51)^{**}$	15.24%	-36.06%
RF	-20.589	(-1.13)	$(-1.87)^{*}$	0.29%	1.68%
logPE	-0.209	(-1.37)	$(-2.06)^{**}$	10.59%	-33.46%
logNPY	0.406	(1.45)	$(3.00)^{***}$	13.00%	-29.24%
DEF	2.668	(0.28)	(0.40)	0.22%	-19.75%
TERM	6.867	$(2.07)^{**}$	$(3.58)^{***}$	10.56%	14.64%
INFL	-12.940	(-1.04)	$(-2.71)^{***}$	2.50%	-0.19%
CAY	5.775	$(1.72)^{*}$	$(2.65)^{***}$	14.72%	5.65%
PYRL	-24.341	(-1.62)	$(-2.82)^{***}$	3.86%	4.06%
NetHR	5.440	(0.48)	(0.60)	0.48%	-1.67%
NetJC	-0.411	(-0.19)	(-0.20)	0.04%	1.17%
SII	-0.040	(-0.55)	(-0.64)	2.06%	-2.93%

Table VI. Predicting Stock Market Returns Using Multiple Regressions

The table reports results of the following predictive regression:

$$r_{t:t+k} = \alpha + \sum_{i \in S} \beta_i z_{i,t} + \varepsilon_{t:t+k},$$

where $r_{t:t+k}$ is the continuously compounded k-month excess return on the market from month t to month $t+k$ and S represents the index of the subset of variables used as forecasting variables in the predictive regression. For each specification, we select regressors from the following set of predictive variables ($z_{i,t}$): the proxy for sectoral shifts (CSV), the log dividend price ratio (logDP), the demeaned risk-free rate (RF), the log price earnings ratio (logPE), the log net payout yield (logNPY), the default spread (DEF), term spread (TERM), inflation rate (INFL), and the consumption-wealth ratio (CAY), payroll growth (PYRL), net hiring rate (NetHR), net job creation in manufacturing (NetJC), and short interest index (SII). The sample period covers January 1952 to December 2013 but three predictors are available for a shorter sample period: logNPY ends in December 2010, TERM starts in April 1953, CAY starts in April 1952 and ends in September 2013, SII starts in January 1973, NetHR starts in March 1977, and NetJC ends in May 2005. For each specification, we restrict the sample period to the longest period that is available for all independent variables used in that specification. Panel A reports the in-sample R^2 for each specification and each forecasting horizon. Panel B presents out-of-sample R^2 for each specification and each forecasting horizon. Panel C reports the regression coefficient estimate ($\hat{\beta}$), the corresponding t-ratios of Hodrick (1992) and Newey-West (1987) with $k-1$ lags, as well as the in-sample and out-of-sample R^2s for two recommended specifications, in which the forecasting variables are CSV and TERM, or CSV and PYRL. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table VI - continued

Panel A: In-Sample R^2

	$k = 1$	$k = 3$	$k = 12$	$k = 24$	$k = 36$
All ex. CSV	5.83%	9.76%	26.84%	42.00%	52.73%
All with CSV	5.83%	10.80%	32.10%	58.23%	62.66%
[CSV, SII]	2.14%	8.58%	27.83%	26.86%	19.16%
[CSV, PYRL]	2.14%	7.99%	23.67%	27.25%	23.88%
[CSV, PYRL, SII]	2.78%	10.56%	33.22%	31.19%	25.41%

Panel B: Out-of-Sample R^2

	$k = 1$	$k = 3$	$k = 12$	$k = 24$	$k = 36$
All ex. CSV	−8.12%	−24.48%	−65.49%	−66.19%	−91.16%
All with CSV	−10.09%	−22.95%	−46.70%	0.93%	−48.56%
[CSV, SII]	2.74%	13.35%	34.63%	24.38%	12.01%
[CSV, PYRL]	0.65%	4.66%	19.28%	24.39%	21.87%
[CSV, PYRL, SII]	1.78%	11.49%	36.26%	28.47%	17.51%

Panel C: Recommended Specifications

		Specification: [CSV, PYRL, SII]				
k		CSV	PYRL	SI	R^2_{IS}	R^2_{OOS}
1	$\hat{\beta}$	−0.093	−2.086	−0.006	2.78%	1.78%
	t-ratio$_{Hodr}$	(−1.89)*	(−1.60)	(−2.62)***		
	t-ratio$_{NW}$	(−1.91)*	(−1.65)*	(−2.75)***		
3	$\hat{\beta}$	−0.397	−6.619	−0.018	10.56%	11.49%
	t-ratio$_{Hodr}$	(−2.63)***	(−1.76)*	(−2.65)***		
	t-ratio$_{NW}$	(−3.70)***	(−2.15)**	(−3.19)***		
12	$\hat{\beta}$	−1.568	−22.106	−0.061	33.22%	36.26%
	t-ratio$_{Hodr}$	(−2.88)***	(−1.81)*	(−2.23)**		
	t-ratio$_{NW}$	(−5.39)***	(−3.62)***	(−3.23)***		
24	$\hat{\beta}$	−2.321	−25.129	−0.050	31.19%	28.47%
	t-ratio$_{Hodr}$	(−2.58)***	(−1.27)	(−0.94)		
	t-ratio$_{NW}$	(−3.64)***	(−2.71)***	(−1.35)		
36	$\hat{\beta}$	−2.465	−33.993	−0.033	25.41%	17.51%
	t-ratio$_{Hodr}$	(−2.15)**	(−1.46)	(0.45)		
	t-ratio$_{NW}$	(−2.92)***	(−3.20)***	(−0.67)		

Table VII. Trading Strategies based on CSV and Alternative Predictors

The table reports the economic value of predictive power of CSV when used as a predictive variable in a trading strategy, following Rapach et al. (2010). The realized gain is calculated from the perspective of a mean-variance investor who allocates between the market portfolio and the risk-free asset. At each time t, the weight allocated to the market portfolio is

$$\hat{w}_t = \frac{1}{\gamma}\frac{\hat{r}_{t:t+k}}{\hat{\sigma}^2_{t:t+k}}.$$

and $1-\hat{w}_t$ is allocated to the risk-free asset, where γ is the coefficient of risk aversion, for which we use a value of three. $\hat{r}_{t:t+k}$ is the k-month continuously compounded predicted excess market return based on the predictive regression (7)

$$\hat{r}_{t:t+k} = \hat{\alpha}_t + \hat{\beta}_t z_t,$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are estimated using data up to time t. The forecast of the market excess return's variance is denoted by $\hat{\sigma}^2_{t:t+k}$, which is estimated using the sample variance of the excess returns over a rolling window of the past five years multiplied by k. The 13 predictive variables (z_t) that we use are: the proxy for sectoral shifts (CSV), the log dividend price ratio (logDP), the de-meaned risk-free rate (RF), the log price earnings ratio (logPE), the log net payout yield (logNPY), the default spread (DEF), term spread (TERM), inflation rate (INFL), and the consumption-wealth ratio (CAY), payroll growth (PYRL), net hiring rate (NetHR), net job creation in manufacturing (NetJC), and short interest index (SII). Panel A shows the annaulized sample mean and standard deviation of the excess returns for portfolios where CSV is used as the predictive variable for different investment horizons, and compares them to the benchmark portfolio where the market excess return is predicted using the historical mean excess return. Panel B illustrates the difference between the realized certainty equivalence of the trading strategy for each predictive variable amd that of the benchmark. The certainty equivalent (CE) for each strategy is defined as

$$CE = \bar{r}_p - \frac{\gamma}{2}\hat{\sigma}^2_p,$$

where $\bar{r}_p$ and $\hat{\sigma}^2_p$ are the annualized sample mean and variance respectively of the excess portfolio return associated with each trading strategy, calculated over all out-of-sample returns. The t-ratios associated with the CE differences are derived in the Internet Appendix. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively, for a one-sided test.

Table VII - continued

Panel A: Portfolios Constructed Using CSV vs. Historical Mean Benchmark			
	Mean	Std. Dev.	Certainty Equivalent
CSV ($k = 1$)	7.64%	16.56%	3.53%
CSV ($k = 3$)	10.85%	18.61%	5.65%
CSV ($k = 12$)	10.43%	17.01%	6.09%
CSV ($k = 24$)	8.72%	16.26%	4.75%
CSV ($k = 36$)	6.76%	15.30%	3.25%
Benchmark	3.99%	15.79%	0.25%

Panel B: Certainty Equivalent Differences for Different Predictive Variables

k		CSV	logDP	RF	logPE	logNPY	DEF	TERM	INFL	CAY	PYRL	NetHR	NetJC	SII
1	ΔCE (%)	3.28*	−4.24	−5.14	−2.15	−5.75	−3.25	−1.96	−7.40	0.36	−4.95	−0.93	−0.07	7.03*
	t-ratio	(1.37)	(−1.26)	(−1.35)	(−1.24)	(−2.20)	(−1.53)	(−0.41)	(−1.12)	(0.11)	(−1.37)	(−0.61)	(−0.03)	(1.55)
3	ΔCE (%)	5.40**	−5.22	−0.60	−2.84	−6.16	−3.61	−1.68	−4.32	0.46	−5.26	−8.68	−0.04	8.85**
	t-ratio	(1.79)	(−1.39)	(−0.42)	(−1.40)	(−2.41)	(−1.74)	(−0.41)	(−0.94)	(0.12)	(−1.40)	(−2.01)	(−0.02)	(1.90)
12	ΔCE (%)	5.83***	−7.81	−0.06	−2.90	−7.00	−1.68	1.65	0.05	1.12	−2.78	−7.35	1.09	6.10*
	t-ratio	(2.55)	(−1.72)	(−0.05)	(−1.14)	(−1.65)	(−1.26)	(0.68)	(0.02)	(0.36)	(−1.00)	(−1.64)	(0.68)	(1.32)
24	ΔCE (%)	4.50***	−7.07	0.58	−1.12	−2.59	−0.14	2.61***	0.66	0.99	0.13	−1.66	−0.05	0.33
	t-ratio	(2.63)	(−1.64)	(0.58)	(−0.48)	(−0.71)	(−0.13)	(2.44)	(0.47)	(0.29)	(0.10)	(−0.56)	(−0.06)	(0.15)
36	ΔCE (%)	2.99***	−4.10	0.94	0.09	−0.35	−0.13	2.83***	0.98	0.86	1.04	−0.14	−0.37	−0.92
	t-ratio	(2.49)	(−1.08)	(0.91)	(0.04)	(−0.12)	(−0.10)	(2.41)	(0.78)	(0.27)	(0.89)	(−0.05)	(−0.40)	(−0.38)

Table VIII. Predicting Stock Market Returns Using Labor Skill-Based CSV

The table reports results of a predictive regression where the dependent variable is the continuously compounded k-month excess return on the market from month t to month $t+k$. We use as the independent variables the high skill CSV (CSV_{HS}) and low skill CSV (CSV_{LS}) at time t, constructed respectively using industries that belong to the upper tercile and lower tercile in terms of industry skill. Panel A reports the results when industry skill is defined as the percentage of workers in the industry that are high skill workers. Panel B shows the results when industry skill is defined as the percentage of total wages in the industry that is associated with high skill workers. Industries are defined at the two-digit SIC level prior to 2002, and at three-digit NAICS level afterwards. The sample is from 1991 to 2013. The table reports the regression coefficient estimate, the corresponding t-ratios of Hodrick (1992) and Newey-West (1987) with $k-1$ lags, as well as the in-sample R^2s. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

	$k=1$			$k=3$			$k=12$			$k=24$			$k=36$		
Panel A: Industry Skill Level Weighted by Employment															
CSV_{HS}	−0.094		−0.128	−0.244		−0.316	−0.922		−1.147	−1.438		−1.612	−2.184		−2.447
$t\text{-ratio}_{Hodr}$	(−2.24)**		(−2.43)**	(−1.99)**		(−2.13)**	(−2.05)**		(−2.16)**	(−1.74)*		(−2.17)**	(−1.90)*		(−2.74)***
$t\text{-ratio}_{NW}$	(−2.33)**		(−2.55)**	(−2.77)***		(−2.61)***	(−3.85)***		(−3.71)***	(−2.19)**		(−3.14)***	(−2.83)***		(−4.63)***
CSV_{LS}		−0.029	0.078		−0.106	0.163		−0.479	0.503		−0.989	0.388		−1.500	0.578
$t\text{-ratio}_{Hodr}$		(−0.50)	(1.06)		(−0.65)	(0.82)		(−0.88)	(0.81)		(−0.95)	(0.46)		(−1.00)	(0.55)
$t\text{-ratio}_{NW}$		(−0.50)	(1.09)		(−0.79)	(0.91)		(−0.90)	(0.75)		(−1.14)	(0.75)		(−1.29)	(0.69)
R^2_{IS}	2.43%	0.12%	2.99%	4.90%	0.49%	5.61%	14.81%	2.09%	16.22%	16.86%	4.20%	17.25%	26.92%	6.79%	27.54%
Panel B: Industry Skill Level Weighted by Wages															
	$k=1$			$k=3$			$k=12$			$k=24$			$k=36$		
CSV_{HS}	−0.095		−0.121	−0.246		−0.303	−0.982		−1.177	−1.517		−1.796	−2.133		−2.331
$t\text{-ratio}_{Hodr}$	(−2.25)**		(−2.35)**	(−2.02)**		(−2.13)**	(−2.19)**		(−2.35)**	(−1.80)*		(−2.32)**	(−1.83)*		(−2.43)**
$t\text{-ratio}_{NW}$	(−2.38)**		(−2.52)**	(−2.76)***		(−2.58)***	(−4.44)***		(−4.47)***	(−2.47)**		(−4.17)***	(−2.73)***		(−4.83)***
CSV_{LS}		−0.028	0.073		−0.096	0.160		−0.452	0.550		−0.735	0.788		−1.405	0.558
$t\text{-ratio}_{Hodr}$		(−0.47)	(0.98)		(−0.55)	(0.79)		(−0.76)	(0.85)		(−0.67)	(0.90)		(−0.86)	(0.47)
$t\text{-ratio}_{NW}$		(−0.47)	(1.01)		(−0.60)	(0.80)		(−0.70)	(0.73)		(−0.72)	(1.21)		(−0.95)	(0.48)
R^2_{IS}	2.53%	0.10%	2.97%	5.05%	0.32%	5.68%	17.17%	1.51%	18.73%	19.12%	1.88%	20.62%	26.21%	4.79%	26.74%

Table IX. Model Parameters

The table reports the values of the calibrated parameters used in our benchmark model at monthly frequency

Description	Symbol	Value
Technology		
Output elasticity of labor	α	0.30
Total separation rate	δ	0.03
Adjustment cost coefficient	κ_h	1
Industry level productivity coefficient	κ_p	0.443
Workers' bargaining power	η	0.72
Workers' unemployment benefit	b	0.51
Preferences and shocks		
Time discount factor	β	0.997
Price of risk	λ	1
Persistence coefficient of aggregate productivity shock	ρ_x	0.95
Conditional volatility of aggregate productivity shock	σ_x	0.0137
Persistence coefficient of sectoral reallocation shock	ρ_s	0.96
Conditional volatility of sectoral reallocation shock	σ_s	0.2
Persistence coefficient of industry level productivity shock	$\rho_{\tilde{z}}$	0.965
Conditional volatility of industry level productivity shock	$\sigma_{\tilde{z}}$	0.262

Table X. Aggregate Moments

The table reports the target aggregate moments for the calibrated model. We present the median along with 40 percentile and 60 percentile of the moments from our cross-simulation distribution based on 500 simulations, and compare it with moments derived from historical data. Moments associated with returns are reported in annual frequency. The sample period for the moments associated with historical data is from January 1952 to December 2013.

		Simulation		
Moments	Data	40%	Median	60%
$\mathbb{E}[R_m - R_f](\%)$	7.01	1.29	1.71	2.12
$\sigma[R_m - R_f](\%)$	15.02	15.28	17.03	19.14
$\mathbb{E}[R_f](\%)$	3.45	3.38	3.52	3.68
$\sigma[R_f](\%)$	0.88	0.71	0.75	0.78
$\mathbb{E}[CSV]$	0.155	0.134	0.145	0.159
$\sigma[CSV]$	0.044	0.119	0.132	0.144
$\rho[CSV]$	0.909	0.901	0.915	0.926

Table XI. Simulation Results: Market Return Predictability

The table reports results of the following predictive regression based on simulated data:

$$r_{t:t+k} = \alpha + \beta CSV_t + \varepsilon_{t:t+k},$$

where $r_{t:t+k}$ is the continuously compounded k-month excess return on the market from month t to month $t+k$, and CSV_t is the time t value of the CSV, both obtained through simulation. The table reports the cross-simulation median of the regression coefficient estimate ($\hat{\beta}$) and Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios, as well as the in-sample and out-of-sample R^2s for 500 simulations. The five columns show results for $k = 1, 3, 12, 24$ and 36 months. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

	Excess market return predictability				
	$k=1$	$k=3$	$k=12$	$k=24$	$k=36$
$\hat{\beta}$	−0.039	−0.107	−0.306	−0.439	−0.498
t-ratio$_{NW}$	(−1.45)	(−2.04)**	(−3.11)***	(−3.67)***	(−3.72)***
$R^2_{IS}(\%)$	1.01	2.84	7.98	10.65	11.59
$R^2_{OOS}(\%)$	−0.39	0.14	3.76	5.85	5.08

Table XII. Predicting Industry Equity Returns Using CSV

The table reports results of predictive regressions of CSV for future industry returns. Each month, we sort returns on industry portfolios into quintiles, based on their past 12-month compounded idiosyncratic returns. Next, we use CSV to predict the time series of the continuously compounded k-month excess returns on each of the industry quintile portfolios (equally weighted) using a similar predictive regression as (7). Panel A shows the results when time series are generated using simulations, where the reported values are the cross-simulation median of the regression coefficient estimate ($\hat{\beta}$) and Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios, as well as the in-sample R^2s for 500 simulations. Panel B reports the results based on actual data from January 1952 to December 2013. The panel reports the in-sample regression coefficient estimates ($\hat{\beta}$), the corresponding t-ratios of Hodrick (1992) and Newey-West (1987) with $k-1$ lags, as well as the in-sample and out-of-sample R^2s. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Industry Excess Return Predictability based on Simulated Data						
		Loser	2	3	4	Winner
$k=1$	$\hat{b}_1$	−0.014	−0.008	−0.007	−0.012	−0.062
	t-ratio$_{NW}$	(−0.90)	(−0.73)	(−0.71)	(−0.80)	(−1.44)
	$R^2_{IS}(\%)$	0.39	0.19	0.17	0.24	1.05
$k=3$	$\hat{b}_1$	−0.042	−0.023	−0.021	−0.034	−0.169
	t-ratio$_{NW}$	(−1.24)	(−1.08)	(−1.00)	(−1.18)	(−2.02)**
	$R^2_{IS}(\%)$	1.05	0.54	0.40	0.62	2.79
$k=12$	$\hat{b}_1$	−0.116	−0.074	−0.073	−0.116	−0.449
	t-ratio$_{NW}$	(−1.70)*	(−1.44)	(−1.28)	(−1.67)*	(−3.00)***
	$R^2_{IS}(\%)$	2.58	1.33	1.13	1.71	7.49
$k=24$	$\hat{b}_1$	−0.174	−0.121	−0.121	−0.195	−0.596
	t-ratio$_{NW}$	(−1.88)*	(−0.55)	(−1.46)	(−1.85)*	(−3.50)***
	$R^2_{IS}(\%)$	3.15	1.74	1.76	2.49	10.24
$k=36$	$\hat{b}_1$	−0.213	−0.151	−0.154	−0.224	−0.671
	t-ratio$_{NW}$	(−1.78)*	(−1.45)	(−1.40)	(−1.83)*	(−3.49)***
	$R^2_{IS}(\%)$	3.27	2.03	1.98	2.80	10.79

Table XII - continued

Panel B: Industry Excess Return Predictability based on Actual Data						
		Loser	2	3	4	Winner
$k = 1$	$\hat{\beta}$	-0.059	-0.030	-0.076	-0.066	-0.099
	t-ratio$_{Hodr}$	(-1.15)	(-0.76)	$(-1.96)^{*}$	$(-1.68)^{*}$	$(-2.08)^{**}$
	t-ratio$_{NW}$	(-1.14)	(-0.76)	$(-1.95)^{*}$	$(-1.68)^{*}$	$(-2.12)^{**}$
	R^2_{IS}	0.21%	0.07%	0.49%	0.39%	0.69%
	R^2_{OOS}	-0.21%	-0.27%	0.24%	0.11%	0.31%
$k = 3$	$\hat{\beta}$	-0.276	-0.242	-0.291	-0.357	-0.378
	t-ratio$_{Hodr}$	$(-1.78)^{*}$	$(-1.94)^{*}$	$(-2.40)^{**}$	$(-2.96)^{***}$	$(-2.48)^{**}$
	t-ratio$_{NW}$	$(-2.25)^{**}$	$(-2.27)^{**}$	(-3.11)	$(-3.35)^{***}$	$(-2.56)^{**}$
	R^2_{IS}	1.38%	1.31%	2.14%	3.18%	2.86%
	R^2_{OOS}	0.48%	0.24%	1.12%	2.85%	2.06%
$k = 12$	$\hat{\beta}$	-1.059	-0.878	-0.977	-1.143	-1.427
	t-ratio$_{Hodr}$	$(-1.82)^{*}$	$(-1.96)^{**}$	$(-2.24)^{**}$	$(-2.69)^{***}$	$(-2.82)^{***}$
	t-ratio$_{NW}$	$(-2.47)^{**}$	$(-2.48)^{**}$	$(-3.12)^{***}$	$(-3.89)^{***}$	$(-3.21)^{***}$
	R^2_{IS}	5.33%	4.68%	6.18%	8.68%	10.66%
	R^2_{OOS}	3.46%	1.63%	3.92%	11.10%	14.70%
$k = 24$	$\hat{\beta}$	-1.504	-1.295	-1.476	-1.744	-2.623
	t-ratio$_{Hodr}$	(-1.62)	$(-1.78)^{*}$	$(-1.99)^{**}$	$(-2.51)^{**}$	$(-3.36)^{***}$
	t-ratio$_{NW}$	$(-3.07)^{***}$	$(-3.37)^{***}$	$(-3.82)^{***}$	$(-4.28)^{***}$	$(-4.98)^{***}$
	R^2_{IS}	7.08%	6.53%	8.65%	11.52%	18.51%
	R^2_{OOS}	7.36%	-2.18%	3.32%	16.39%	27.35%
$k = 36$	$\hat{\beta}$	-1.334	-1.197	-1.352	-1.709	-2.593
	t-ratio$_{Hodr}$	(-1.15)	(-1.29)	(-1.40)	$(-1.90)^{*}$	$(-2.63)^{***}$
	t-ratio$_{NW}$	$(-2.17)^{**}$	$(-2.22)^{**}$	$(-2.77)^{***}$	$(-3.29)^{***}$	$(-3.71)^{***}$
	R^2_{IS}	4.83%	4.53%	5.62%	8.44%	13.75%
	R^2_{OOS}	11.07%	-13.14%	-4.11%	10.01%	24.71%

Table XIII. Predicting Stock Market Returns Using CSV Based on Industries With High– and Low– Asset Redeployability

The table reports results of a predictive regression where the dependent variable is the continuously compounded k-month excess return on the market from month t to month $t+k$. We use as the independent variables the high-redeployability CSV (CSV_{HR}) and low-redeployability CSV (CSV_{LR}) at time t, constructed respectively using industries that belong to the highest tercile and lowest tercile in terms of industry-level asset redeployability index, defined by Kim and Kung (2017). Panel A reports the results when asset-level redeployability score is defined based on the market capitalization of the industries that use the asset. Panel B shows the results when asset-level redeployability score is defined based on the number of industries that utilize the asset. Industries are defined based on BEA industry classification, where we consider industries that consist of at least 5 firms. The sample period is from 1985 to 2013. The table reports the regression coefficient estimate, the corresponding t-ratios of Hodrick (1992) and Newey-West (1987) with $k-1$ lags, as well as the in-sample R^2s. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table XIII - continued

		Panel A: Value-Weighted Asset Redeployability Score			Panel B: Equally-Weighted Asset Redeployability Score		
$k=1$	CSV_{HR}	−0.115		−0.073	−0.142		−0.108
	$t\text{-ratio}_{Hodr}$	(−2.35)**		(−1.27)	(−2.69)***		(−1.68)*
	$t\text{-ratio}_{NW}$	(−2.47)**		(−1.32)	(−2.91)***		(−1.80)*
	CSV_{LR}		−0.137	−0.083		−0.140	−0.064
	$t\text{-ratio}_{Hodr}$		(−2.55)**	(−1.34)		(−2.56)**	(−0.97)
	$t\text{-ratio}_{NW}$		(−2.79)***	(−1.46)		(−2.81)***	(−1.06)
	R^2_{IS}	2.59%	2.53%	3.19%	3.44%	2.54%	3.76%
$k=3$	CSV_{HR}	−0.412		−0.277	−0.482		−0.365
	$t\text{-ratio}_{Hodr}$	(−2.68)***		(−1.70)*	(−2.95)***		(−2.04)**
	$t\text{-ratio}_{NW}$	(−4.65)***		(−2.80)***	(−4.99)***		(−3.08)***
	CSV_{LR}		−0.472	−0.270		−0.475	−0.217
	$t\text{-ratio}_{Hodr}$		(−2.54)**	(−1.37)		(−2.52)**	(−1.05)
	$t\text{-ratio}_{NW}$		(−3.69)***	(−1.73)*		(−3.67)***	(−1.29)
	R^2_{IS}	10.00%	8.99%	11.86%	11.82%	8.76%	12.95%
$k=12$	CSV_{HR}	−1.699		−1.500	−1.937		−1.862
	$t\text{-ratio}_{Hodr}$	(−2.82)***		(−2.79)***	(−2.96)***		(−3.06)***
	$t\text{-ratio}_{NW}$	(−5.45)***		(−3.48)***	(−6.37)***		(−4.12)***
	CSV_{LR}		−1.496	−0.387		−1.467	−0.135
	$t\text{-ratio}_{Hodr}$		(−2.35)**	(−0.81)		(−2.31)**	(−0.28)
	$t\text{-ratio}_{NW}$		(−4.41)***	(−0.99)		(−4.11)***	(−0.32)
	R^2_{IS}	37.7%	20.27%	38.55%	42.35%	18.76%	42.44%
$k=24$	CSV_{HR}	−2.253		−2.246	−2.316		−2.389
	$t\text{-ratio}_{Hodr}$	(−2.52)**		(−2.65)***	(−2.48)**		(−2.63)***
	$t\text{-ratio}_{NW}$	(−4.26)***		(−6.10)***	(−3.36)***		(−4.99)***
	CSV_{LR}		−1.693	−0.014		−1.604	0.130
	$t\text{-ratio}_{Hodr}$		(−1.85)*	(−0.02)		(−1.78)*	(0.18)
	$t\text{-ratio}_{NW}$		(−1.75)*	(−0.02)		(−1.60)	(0.15)
	R^2_{IS}	32.24%	12.62%	32.24%	29.27%	10.87%	29.31%
$k=36$	CSV_{HR}	−1.989		−1.628	−2.239		−2.070
	$t\text{-ratio}_{Hodr}$	(−1.84)*		(−1.62)	(−1.98)**		(−1.89)*
	$t\text{-ratio}_{NW}$	(−2.28)**		(−1.97)**	(−2.40)**		(−2.52)**
	CSV_{LR}		−1.957	−0.716		−1.836	−0.310
	$t\text{-ratio}_{Hodr}$		(−1.80)*	(−0.95)		(−1.74)*	(−0.4)
	$t\text{-ratio}_{NW}$		(−1.78)*	(−0.75)		(−1.62)	(−0.31)
	R^2_{IS}	17.52%	11.22%	18.44%	19.04%	9.48%	19.2%

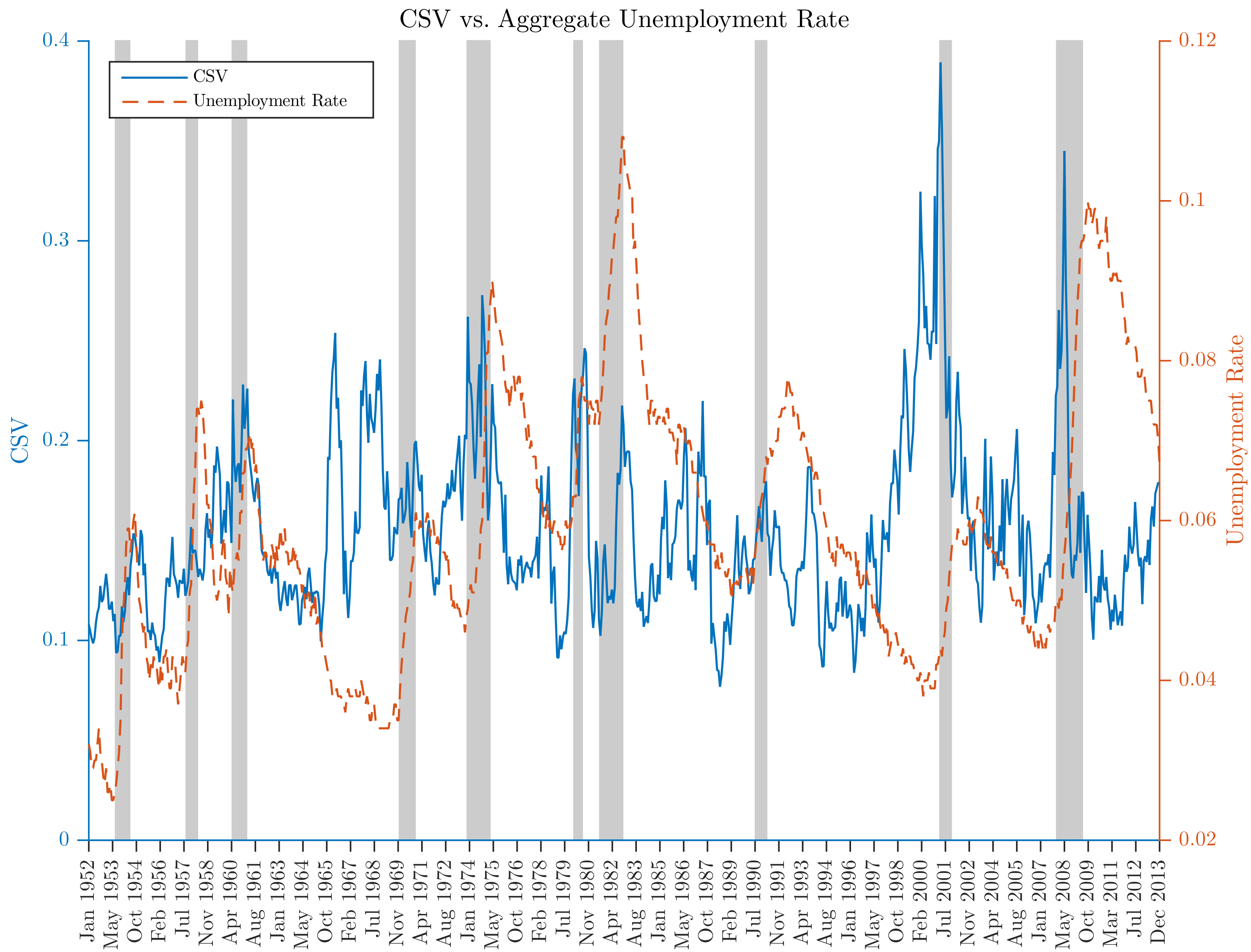

Figure 1. Time series of CSV and aggregate unemployment. The solid line shows the time series of our monthly CSV measure, which is computed based on the past 12 months of industry-specific returns (i.e., from month $t-11$ to t). The dashed line shows the aggregate unemployment rates (in levels) at time t. The shaded areas correspond to NBER recession dates.

Internet Appendix for

Sectoral Labor Reallocation and Return Predictability

Esther Eiling, Raymond Kan and Ali Sharifkhani

February 2, 2018

This document presents the derivation of a statistical test for the equality of certainty equivalents of two different trading strategies as well as the results of several robustness tests for the paper. First, we present our derivation of the certainty equivalent statistical test. For our robustness tests, we begin by showing the results when we use our proxy for sectoral reallocation shocks (CSV) to predict individual industry rather than market equity returns. Then, we use different horizons of past industry returns (3 months and 24 months) and different weighting schemes (employment-based weights and market capitalization-based weights) to construct CSV. Further, we use total rather than idiosyncratic industry returns in constructing our CSV measure, as well as use CSVs based on individual stocks or based on 100 size and book-to-market ranked portfolios rather than industry portfolios. Finally, we perform the analysis for three different subsample periods.

A. Test of Certainty Equivalent Differences

Suppose we have returns on two portfolios, $r_{1,t}$ and $r_{2,t}$ for $t = 1, \ldots, T$. Let $\mu_i = E[r_{i,t}]$ and $\sigma_i^2 = \mathrm{Var}[r_{i,t}]$. The certainty equivalents of the two portfolios are given by

$$U_1 = \mu_1 - \frac{\gamma}{2}\sigma_1^2, \qquad U_2 = \mu_2 - \frac{\gamma}{2}\sigma_2^2. \tag{IA.1}$$

The sample estimates of U_1 and U_2 are given by

$$\hat{U}_1 = \hat{\mu}_1 - \frac{\gamma}{2}\hat{\sigma}_1^2, \qquad \hat{U}_2 = \hat{\mu}_2 - \frac{\gamma}{2}\hat{\sigma}_2^2, \tag{IA.2}$$

where $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ are the sample mean and variance of portfolio i, for $i = 1, 2$. Under joint stationarity and ergodicity assumption on $\{r_{1t}, r_{2t}\}$, the joint asymptotic distribution of $(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2)$ is given by

$$\sqrt{T}\begin{bmatrix} \hat{\mu}_1 - \mu_1 \\ \hat{\mu}_2 - \mu_2 \\ \hat{\sigma}_1^2 - \sigma_1^2 \\ \hat{\sigma}_2^2 - \sigma_2^2 \end{bmatrix} \overset{A}{\sim} N\left(\mathbf{0}, \sum_{j=-\infty}^{\infty} E[h_t h_{t+j}']\right), \tag{IA.3}$$

where

$$h_t = \begin{bmatrix} r_{1,t} - \mu_1 \\ r_{2,t} - \mu_2 \\ (r_{1,t} - \mu_1)^2 - \sigma_1^2 \\ (r_{2,t} - \mu_2)^2 - \sigma_2^2 \end{bmatrix}. \tag{IA.4}$$

Using the delta method, we have

$$\sqrt{T}(\hat{U}_1 - \hat{U}_2 - (U_1 - U_2)) \overset{A}{\sim} N\left(0, \sum_{j=-\infty}^{\infty} E[g_t g_{t+j}]\right), \tag{IA.5}$$

where

$$g_t = r_{1,t} - \frac{\gamma}{2}(r_{1,t} - \mu_1)^2 - r_{2,t} + \frac{\gamma}{2}(r_{2,t} - \mu_2)^2 - (U_1 - U_2). \tag{IA.6}$$

For constructing the estimated standard error of $\hat{U}_1 - \hat{U}_2$, we replace g_t with its sample counterpart

$$\hat{g}_t = r_{1,t} - \frac{\gamma}{2}(r_{1,t} - \hat{\mu}_1)^2 - r_{2,t} + \frac{\gamma}{2}(r_{2,t} - \hat{\mu}_2)^2 - (\hat{U}_1 - \hat{U}_2), \tag{IA.7}$$

and use zero lags as returns are almost serially uncorrelated.

Table AI. Predicting Individual Industry Returns Using CSV

This table reports results of predictive regressions for individual industry returns of 49 industries. We regress our proxy for sectoral shifts (CSV) on log industry excess returns over the next 12 months:

$$r_{i,t:t+12} = \alpha_i + \beta_i CSV_t + \varepsilon_{i,t:t+12},$$

where $r_{i,t:t+12}$ is the continuously compounded 12-month excess return for industry i from month t to month $t+12$. The table reports the regression coefficient estimate ($\hat{\beta}$), the corresponding t-ratios of Hodrick (1992) and Newey-West (1987) with $k-1$ lags, as well as the in-sample and out-of-sample R^2s. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Industry	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
Agriculture	−1.45	(−2.84)***	(−2.96)***	9.19%	10.13%
Food Products	−0.63	(−1.69)*	(−1.90)*	3.06%	0.93%
Candy and Soda	−1.08	(−1.59)	(−2.14)**	4.61%	4.09%
Beer and Liquor	−1.25	(−2.65)***	(−3.56)***	9.65%	8.36%
Tobacco Products	−0.29	(−0.53)	(−0.49)	0.35%	−4.14%
Recreation	−1.37	(−2.09)**	(−2.43)**	4.09%	−3.35%
Entertainment	−1.28	(−1.61)	(−1.86)*	3.65%	1.62%
Printing and Publishing	−1.41	(−2.69)***	(−2.67)***	6.48%	1.17%
Consumer Goods	−0.99	(−2.33)**	(−2.55)**	6.06%	3.00%
Apparel	−0.63	(−1.03)	(−1.42)	1.30%	−1.82%
Healthcare	−0.49	(−0.63)	(−0.64)	0.41%	−10.73%
Medical Equipment	−0.88	(−1.88)*	(−2.51)**	3.92%	1.20%
Pharmaceutical Products	−1.36	(−3.04)***	(−5.06)***	10.67%	10.82%
Chemicals	−1.19	(−2.30)**	(−3.04)***	7.47%	5.44%
Rubber and Plastic Products	−1.10	(−1.95)*	(−2.33)**	5.00%	−0.68%
Textiles	−0.92	(−1.27)	(−1.41)	2.45%	−1.42%
Construction Materials	−1.12	(−1.95)*	(−2.35)**	5.64%	3.49%
Construction	−0.93	(−1.40)	(−1.76)	2.54%	0.58%
Steel Works Etc	−1.62	(−1.88)*	(−2.40)**	7.17%	1.33%
Fabricated Products	−1.57	(−2.22)**	(−3.33)***	8.86%	14.36%
Machinery	−1.25	(−1.83)*	(−2.65)***	6.68%	2.58%

Table AI - continued

Industry	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
Electrical Equipment	−1.96	(−3.03)***	(−4.55)***	16.38%	14.53%
Automobiles and Trucks	−1.47	(−2.04)**	(−2.97)***	6.48%	4.32%
Aircraft	−1.95	(−2.66)***	(−3.77)***	10.25%	1.31%
Shipbuilding, Railroad Equipment	−0.99	(−1.53)	(−1.65)*	3.66%	−1.60%
Defense	−0.23	(−0.35)	(−0.33)	0.21%	−8.98%
Precious Metals	−0.09	(−0.10)	(−0.12)	0.02%	−13.61%
Non-Metallic and Industrial Metal Mining	−0.91	(−1.32)	(−1.68)*	2.40%	−1.32%
Coal	0.20	(0.18)	(0.20)	0.07%	−3.99%
Petroleum and Natural Gas	−1.07	(−2.21)**	(−2.89)***	6.73%	5.87%
Utilities	−0.97	(−2.35)**	(−2.73)***	8.58%	6.34%
Communication	−1.76	(−3.49)***	(−4.07)***	17.69%	15.68%
Personal Services	−0.76	(−1.49)	(−1.45)	1.50%	−8.49%
Business Services	−1.47	(−2.62)***	(−4.45)***	10.02%	9.95%
Computers	−2.37	(−2.36)**	(−3.05)***	13.14%	7.67%
Computer Software	−2.75	(−2.66)***	(−4.50)***	8.23%	20.70%
Electronic Equipment	−2.43	(−2.52)**	(−3.78)***	14.51%	7.78%
Measuring and Control Equipment	−1.72	(−2.14)**	(−3.32)***	9.35%	6.45%
Business Supplies	−0.40	(−0.80)	(−0.99)	0.74%	−0.39%
Shipping Containers	−0.97	(−1.60)	(−1.74)*	5.00%	0.37%
Transportation	−0.99	(−1.85)*	(−2.63)***	4.42%	3.02%
Wholesale	−1.04	(−2.08)**	(−2.88)***	4.93%	4.76%
Retail	−0.80	(−1.65)*	(−2.57)**	3.48%	2.33%
Restaurants, Hotels, Motels	−0.97	(−1.82)*	(−2.32)**	2.92%	1.76%
Banking	−0.95	(−1.70)*	(−1.91)*	3.62%	2.30%
Insurance	−0.99	(−1.80)*	(−2.03)**	4.75%	2.71%
Real Estate	−0.97	(−1.37)	(−1.34)	1.79%	−1.40%
Trading	−1.28	(−1.81)*	(−2.98)***	5.90%	5.31%
Other	−1.78	(−2.95)***	(−3.98)***	8.51%	10.13%

Table AII. Predictive Regressions Using Alternative Horizons for CSV

This table reports results of predictive regressions on future unemployment growth (Panel A) and future excess market returns (Panel B) using CSV based on past three-month ($m = 3$) or past 24 month ($m = 24$) industry idiosyncratic returns, rather than using $m = 12$ as in our primary measure. We run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta CSV_t^m + \varepsilon_{t:t+k}.$$

For Panel A, $y_{t:t+k} = un_{t+k} - un_t$ where un_t is based on a log transformation of the unemployment rate: $un_t = \log(UN_t/(1 - UN_t))$. For Panel B, $y_{t:t+k} = r_{t:t+k}$, i.e., the continuously compounded k-month excess return on the market from month t to month $t + k$. CSV_t^m is the cross-sectional volatility of past m-month industry idiosyncratic returns. Panel A reports the regression coefficient estimate ($\hat{\beta}$), the corresponding Newey-West (1987) adjusted (based on $k - 1$ lags) t-ratios and the in-sample R^2s. In addition, Panel B reports Hodrick (1992) t-ratios and the out-of-sample R^2. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth

		$m = 3$			$m = 24$	
k	$\hat{\beta}$	t-ratio	R^2_{IS}	$\hat{\beta}$	t-ratio	R^2_{IS}
1	0.22	(3.27)***	1.59%	0.08	(3.16)***	1.78%
3	0.51	(2.46)**	1.96%	0.19	(2.60)***	2.71%
12	1.49	(1.53)	2.23%	0.51	(1.71)*	2.61%
24	2.08	(1.57)	2.27%	0.61	(1.13)	1.89%
36	2.21	(1.57)	2.04%	0.56	(0.92)	1.28%

Panel B: Predicting Excess Stock Market Returns

			$m = 3$		
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.12	(−1.27)	(−1.26)	0.32%	−0.12%
3	−0.36	(−1.49)	(−1.75)*	0.92%	−0.36%
12	−1.99	(−2.36)**	(−2.80)***	6.40%	3.09%
24	−3.81	(−2.50)**	(−3.36)***	12.88%	10.95%
36	−4.44	(−2.18)**	(−3.09)***	13.32%	12.20%
			$m = 24$		
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.06	(−2.25)**	(−2.24)**	0.84%	0.54%
3	−0.20	(−2.50)**	(−2.79)***	2.65%	1.61%
12	−0.66	(−2.39)**	(−2.39)**	6.84%	4.81%
24	−1.13	(−2.41)**	(−2.46)**	11.02%	4.42%
36	−1.55	(−2.53)**	(−2.87)***	15.80%	11.12%

Table AIII. Predictive Regressions where CSV has Employment-Based Weights

This table reports results of predictive regressions on future unemployment growth (Panel A) and future excess market returns (Panel B) where we use as a predictor CSV with employment-based weights, rather than equal weights as in our primary measure. Industry-level employment data are available for a set of 35 industries for the period from January 1952 to April 2003. We use the CRSP data on individual stocks to create value-weighted returns on 35 matched industry equity portfolios. Similar to our main measure, we estimate industry idiosyncratic returns with respect to the CAPM, estimated over the past 36 months. We calculate CSV based on past $m = 3$, 12 and 24-month industry-specific returns. We run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta CSV_t^m + \varepsilon_{t:t+k},$$

where $y_{t:t+k}$ is either the k-month ahead unemployment growth or the continuously compounded k-month excess return on the market from month t to month $t+k$. CSV_t^m is the cross-sectional volatility of past m-month industry idiosyncratic returns. Panel A reports the regression coefficient estimate, the corresponding Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios and the in-sample R^2s. In addition, Panel B reports Hodrick (1992) t-ratios and the out-of-sample R^2. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth

	$m = 3$			$m = 12$		
k	$\hat{\beta}$	t-ratio	R^2_{IS}	$\hat{\beta}$	t-ratio	R^2_{IS}
1	0.15	(1.80)*	0.47%	0.03	(0.63)	0.05%
3	0.29	(1.21)	0.43%	0.01	(0.12)	0.00%
12	0.98	(1.04)	0.68%	0.22	(0.35)	0.13%
24	2.61	(2.03)**	2.68%	1.15	(1.26)	1.93%
36	2.75	(1.81)*	2.30%	1.68	(1.36)	3.16%

	$m = 24$		
k	$\hat{\beta}$	t-ratio	R^2_{IS}
1	0.01	(0.26)	0.01%
3	−0.01	(−0.12)	0.00%
12	0.05	(0.14)	0.02%
24	0.45	(0.56)	0.74%
36	0.97	(0.98)	2.97%

Table AIII - continued

Panel B: Predicting Excess Stock Market Returns					
		$m = 3$			
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.04	(−0.36)	(−0.35)	0.02%	−0.45%
3	−0.48	(−1.63)	(−2.13)*	1.24%	0.39%
12	−2.43	(−2.39)**	(−3.99)***	7.35%	6.99%
24	−4.72	(−3.01)***	(−5.12)***	14.80%	17.63%
36	−5.63	(−3.92)***	(−4.02)***	14.65%	17.95%
		$m = 12$			
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.08	(−1.49)	(−1.52)	0.41%	0.20%
3	−0.34	(−2.13)**	(−2.79)***	2.19%	1.98%
12	−1.15	(−2.16)**	(−3.09)***	6.03%	6.66%
24	−2.72	(−3.10)***	(−4.34)***	18.24%	16.49%
36	−3.03	(−2.48)**	(−3.35)***	15.61%	12.07%
		$m = 24$			
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.04	(−1.38)	(−1.38)	0.32%	0.03%
3	−0.18	(−1.99)**	(−2.51)**	1.74%	0.91%
12	−0.72	(−2.18)**	(−2.61)***	6.18%	2.87%
24	−1.40	(−2.07)**	(−3.50)***	12.14%	1.22%
36	−2.04	(−2.18)**	(−3.89)***	20.10%	4.91%

Table AIV. Predictive Regressions where CSV has Market Capitalization-Based Weights

This table reports results of predictive regressions on future unemployment growth (Panel A) or future excess market returns (Panel B) where our predictor is CSV (for 49 industries) with market capitalization-based weights, rather than equal weights as in our primary measure. We run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta CSV_t^m + \varepsilon_{t:t+k},$$

where $y_{t:t+k}$ is either the k-month ahead unemployment growth or the continuously compounded k-month excess return on the market from month t to month $t+k$. CSV_t^m is the market capitalization weighted cross-sectional volatility of past m-month industry idiosyncratic returns. Panel A reports the regression coefficient estimate, the corresponding Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios and the in-sample R^2s. In addition, Panel B reports Hodrick (1992) t-ratios and the out-of-sample R^2,. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth

	$m=3$			$m=12$		
k	$\hat{\beta}$	t-ratio	R^2_{IS}	$\hat{\beta}$	t-ratio	R^2_{IS}
1	0.24	$(3.27)^{***}$	1.69%	0.06	$(1.70)^{*}$	0.46%
3	0.59	$(2.75)^{***}$	2.46%	0.16	(1.48)	0.81%
12	1.45	(1.47)	1.93%	0.48	(0.99)	0.92%
24	1.76	(1.31)	1.47%	1.06	$(1.72)^{*}$	2.36%
36	1.38	(0.89)	0.72%	0.97	(1.23)	1.60%

	$m=24$		
k	$\hat{\beta}$	t-ratio	R^2_{IS}
1	0.02	(0.86)	0.12%
3	0.04	(0.68)	0.13%
12	0.06	(0.24)	0.04%
24	0.37	(0.82)	0.63%
36	0.36	(0.68)	0.49%

Table AIV - continued

Panel B: Predicting Excess Stock Market Returns					
			$m = 3$		
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	-0.16	(-1.64)	(-1.61)	0.57%	0.04%
3	-0.39	(-1.55)	$(-1.86)^{*}$	0.98%	-0.20%
12	-1.64	$(-1.90)^{*}$	$(-2.44)^{**}$	3.96%	1.27%
24	-3.01	$(-2.03)^{**}$	$(-2.38)^{**}$	7.32%	4.16%
36	-3.40	$(-1.72)^{*}$	$(-1.97)^{*}$	7.14%	3.64%
			$m = 12$		
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	-0.09	$(-2.08)^{**}$	$(-2.14)^{**}$	0.79%	0.41%
3	-0.29	$(-2.36)^{**}$	$(-3.15)^{***}$	2.43%	1.39%
12	-1.10	$(-2.43)^{**}$	$(-3.87)^{***}$	7.90%	7.95%
24	-2.02	$(-2.60)^{***}$	$(-4.33)^{***}$	14.59%	15.17%
36	-2.02	$(-1.97)^{**}$	$(-2.67)^{***}$	11.17%	8.68%
			$m = 24$		
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	-0.05	$(-1.88)^{*}$	$(-1.93)^{*}$	0.58%	0.23%
3	-0.16	$(-1.91)^{*}$	$(-2.65)^{***}$	1.50%	0.66%
12	-0.65	$(-2.11)^{**}$	$(-3.72)^{***}$	6.08%	6.11%
24	-1.12	$(-2.27)^{**}$	$(-3.22)^{***}$	9.87%	8.58%
36	-1.40	$(-2.18)^{**}$	$(-2.91)^{***}$	11.76%	9.19%

Table AV. Predictive Regressions where CSV is based on Total Industry Returns

This table reports results of predictive regressions on future unemployment growth (Panel A) or future excess market returns (Panel B) where our predictor variable CSV is based on total rather than idiosyncratic industry returns (for 49 industries) over the past 12 months ($m = 12$). We run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta CSV_t + \varepsilon_{t:t+k},$$

where $y_{t:t+k}$ is either the k-month ahead unemployment growth or the continuously compounded k-month excess return on the market from month t to month $t+k$. CSV_t is the equally weighted cross-sectional volatility of past 12-month compounded total industry returns. Panel A reports the regression coefficient estimate, the corresponding Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios and the in-sample R^2s. In addition, Panel B reports Hodrick (1992) t-ratios and the out-of-sample R^2. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth

k	$\hat{\beta}$	t-ratio	R^2_{IS}
1	−0.01	(−0.66)	0.05%
3	−0.04	(−0.51)	0.08%
12	0.04	(0.11)	0.01%
24	0.52	(1.07)	1.10%
36	0.67	(1.22)	1.44%

Panel B: Predicting Excess Stock Market Returns

k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.06	(−2.17)**	(−2.13)**	0.69%	0.51%
3	−0.28	(−3.32)***	(−4.47)***	4.20%	4.22%
12	−0.97	(−3.54)***	(−4.14)***	11.70%	13.32%
24	−1.72	(−3.58)***	(−5.26)***	20.17%	19.17%
36	−1.93	(−3.01)***	(−4.08)***	19.37%	14.20%

Table AVI. Predictive Regressions where CSV is based on Individual Stock Returns

This table reports results of predictive regressions on future unemployment growth (Panel A) or future excess market returns (Panel B) where our predictor variable CSV is based on total returns on individual stocks rather than idiosyncratic returns on 49 industry portfolios. We use stock returns from CRSP for all stocks traded on the NYSE, AMEX and Nasdaq. We run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta CSV_t + \varepsilon_{t:t+k},$$

where $y_{t:t+k}$ is either the k-month ahead unemployment growth or the continuously compounded k-month excess return on the market from month t to month $t+k$. CSV_t is the equally or market capitalization weighted cross-sectional volatility of past 12-month compounded total returns on individual stocks. Panel A reports the regression coefficient estimate, the corresponding Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios and the in-sample R^2s. In addition, Panel B reports Hodrick (1992) t-ratios and the out-of-sample R^2. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth

	Equally weighted			Value weighted		
k	$\hat{\beta}$	t-ratio	R^2_{IS}	$\hat{\beta}$	t-ratio	R^2_{IS}
1	−0.01	(−3.56)***	1.60%	−0.02	(−3.16)***	1.24%
3	−0.04	(−3.25)***	3.10%	−0.05	(−2.75)***	2.32%
12	−0.13	(−2.18)**	4.00%	−0.13	(−1.29)	2.04%
24	−0.16	(−1.40)	2.91%	−0.05	(−0.29)	0.16%
36	−0.14	(−1.08)	1.95%	0.01	(0.07)	0.01%

Panel B: Predicting Excess Stock Market Returns

			Equally weighted		
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	0.00	(−0.84)	(−0.83)	0.09%	−0.48%
3	−0.02	(−1.24)	(−1.65)	0.52%	−0.63%
12	−0.07	(−1.28)	(−1.23)	1.54%	−3.33%
24	−0.13	(−1.60)	(−1.41)	3.58%	−11.66%
36	−0.16	(−1.51)	(−1.17)	3.86%	−22.26%
			Value weighted		
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.01	(−1.10)	(−1.11)	0.21%	−0.40%
3	−0.03	(−1.35)	(−2.07)**	0.81%	0.01%
12	−0.14	(−1.63)	(−2.49)**	3.91%	1.10%
24	−0.29	(−1.96)*	(−3.41)***	9.23%	2.31%
36	−0.39	(−2.03)**	(−3.30)***	12.90%	0.36%

Table AVII. Predictive Regressions where CSV is based on 100 Size and BM Ranked Portfolios

This table reports results of predictive regressions on future unemployment growth (Panel A) or future excess market returns (Panel B) where our predictor variable CSV is based on idiosyncratic returns on 100 size and book-to-market ranked portfolios rather than 49 industry portfolios. We run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta CSV_t + \varepsilon_{t:t+k},$$

where $y_{t:t+k}$ is either the k-month ahead unemployment growth or the continuously compounded k-month excess return on the market from month t to month $t+k$. CSV_t is the equally weighted cross-sectional volatility of past 12-month compounded idiosyncratic returns on 100 size and book-to-market ranked portfolios. Panel A reports the regression coefficient estimate, the corresponding Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios and the in-sample R^2s. In addition, Panel B reports Hodrick (1992) t-ratios and the out-of-sample R^2. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth

k	$\hat{\beta}$	t-ratio	R^2_{IS}
1	0.02	(0.44)	0.02%
3	0.07	(0.58)	0.09%
12	0.54	(1.02)	0.75%
24	0.96	(1.19)	1.22%
36	0.70	(0.62)	0.52%

Panel B: Predicting Excess Stock Market Returns

k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.14	(−2.48)**	(−2.37)**	1.12%	0.97%
3	−0.45	(−2.88)***	(−3.78)***	3.73%	3.89%
12	−1.20	(−2.21)**	(−3.12)***	5.99%	5.62%
24	−1.81	(−1.93)*	(−2.29)**	7.33%	2.38%
36	−1.67	(−1.33)	(−1.57)	4.77%	−1.34%

Table AVIII. The Role of Informed Investors

This table reports the effect of analyst coverage on the predictive power of CSV. At each month, we construct a measure of analyst coverage for each industry based on the average number of analyst announcement per firm over the past 6 months in the corresponding industry (Panel A) or the percentage of firms in the industry with at least one analyst announcement over the past 6 months (Panel B). We next construct $coverage_t$ at each month t by calculating the equally weighted average of the analyst coverage of the top and bottom 20 percentile of industries ranked in terms of past 12 months idiosyncratic return. We then run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta_1 CSV_t + \beta_2 CSV_t \times coverage_t + \varepsilon_{t:t+k},$$

where $y_{t:t+k}$ is the continuously compounded k-month excess return on the market from month t to month $t+k$. CSV_t is the equally weighted cross-sectional volatility of past 12-month compounded idiosyncratic returns on 49 industry portfolios. Each panel reports the regression coefficient estimate, the corresponding t-ratios of Hodrick (1992) and Newey-West (1987) with $k-1$ lags, and the in-sample R^2s. The sample period covers January 1984 to December 2013 ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Average Number of Analysts per Firm in the Industry

		$k=1$	$k=3$	$k=12$	$k=24$	$k=36$
CSV	$\hat{b}_1$	−0.16	−0.56	−1.77	−2.24	−2.61
	t-ratio$_{Hodr}$	(−2.19)**	(−2.60)***	(−2.15)**	(−1.20)	(−0.95)
	t-ratio$_{NW}$	(−2.22)**	(−3.40)***	(−3.59)***	(−1.80)**	(−1.39)
CSV× coverage	$\hat{b}_1$	0.03	0.04	−0.05	−0.18	−0.20
	t-ratio$_{Hodr}$	(1.05)	(0.49)	(−0.13)	(−0.19)	(−0.13)
	t-ratio$_{NW}$	(1.08)	(0.58)	(−0.16)	(−0.31)	(−0.21)
	R^2_{IS}	1.76	8.89	30.47	28.92	27.62

Panel B: Percentage of Firms in the Industry with at Least One Analyst

		$k=1$	$k=3$	$k=12$	$k=24$	$k=36$
CSV	$\hat{b}_1$	−0.16	−0.63	−1.89	−2.03	−1.91
	t-ratio$_{Hodr}$	(−1.44)	(−1.84)*	(−1.45)	(−0.74)	(−0.50)
	t-ratio$_{NW}$	(−1.44)	(−2.27)**	(−2.37)**	(−1.24)	(−0.76)
CSV× coverage	$\hat{b}_1$	0.20	0.52	0.18	−2.06	−4.12
	t-ratio$_{Hodr}$	(0.51)	(0.43)	(0.04)	(−0.19)	(−0.27)
	t-ratio$_{NW}$	(0.52)	(0.51)	(0.05)	(−0.35)	(−0.45)
	R^2_{IS}	1.61	8.91	30.43	28.96	28.12

Table AIX. Predictive Regressions for Different Subsample Periods

This table reports results of predictive regressions on future unemployment growth (Panel A) or future excess market returns (Panel B) for three different subsample periods: from January 1973 to December 2013, from January 1952 to December 1982 and from January 1983 to December 2013. We run the following predictive regressions:

$$y_{t:t+k} = \alpha + \beta CSV_t + \varepsilon_{t:t+k},$$

where $y_{t:t+k}$ is either the k-month ahead unemployment growth or the continuously compounded k-month excess return on the market from month t to month $t+k$. CSV_t is the equally weighted cross-sectional volatility of past 12-month compounded idiosyncratic returns on 49 industry portfolios. Panel A reports the regression coefficient estimate, the corresponding Newey-West (1987) adjusted (based on $k-1$ lags) t-ratios and the in-sample R^2s. The final column reports the correlation between the log excess stock market returns and unemployment growth, where both are taken over the same k-month horizon. In addition, Panel B reports Hodrick (1992) t-ratios and the out-of-sample R^2. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

Panel A: Predicting Aggregate Unemployment Growth

		1973:01–2013:12		
k	$\hat{\beta}$	t-ratio	R^2_{IS}	$\text{corr}(r_{M,t:t+k}, \Delta un_{t:t+k})$
1	0.12	(3.70)***	3.85%	−0.014
3	0.38	(4.02)***	9.32%	−0.111
12	1.34	(3.03)***	13.13%	−0.379***
24	2.12	(4.18)***	13.13%	−0.569***
36	1.85	(2.60)***	7.08%	−0.643***
		1952:01–1982:12		
k	$\hat{\beta}$	t-ratio	R^2_{IS}	$\text{corr}(r_{M,t:t+k}, \Delta un_{t:t+k})$
1	0.04	(0.67)	0.14%	−0.003
3	0.06	(0.26)	0.06%	−0.002
12	0.12	(0.14)	0.04%	−0.115
24	0.88	(0.69)	1.20%	−0.133
36	0.22	(0.15)	0.07%	−0.174
		1983:01–2013:12		
k	$\hat{\beta}$	t-ratio	R^2_{IS}	$\text{corr}(r_{M,t:t+k}, \Delta un_{t:t+k})$
1	0.11	(3.15)***	3.60%	−0.038
3	0.35	(4.38)***	10.59%	−0.186*
12	1.46	(3.34)***	20.09%	−0.444***
24	2.13	(3.97)***	15.79%	−0.631***
36	2.11	(2.70)***	9.97%	−0.703***

Table AIX - continued

Panel B: Predicting Excess Stock Market Returns					
		1973:01–2013:12			
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.09	(−1.87)*	(−1.89)*	0.80%	0.47%
3	−0.39	(−2.60)***	(−3.39)***	4.63%	3.76%
12	−1.52	(−2.82)***	(−4.57)***	17.66%	15.10%
24	−2.21	(−2.55)**	(−4.21)***	23.06%	22.50%
36	−2.22	(−2.01)**	(−2.87)***	18.05%	16.97%
		1952:01–1982:12			
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.06	(−1.00)	(−1.01)	0.34%	−0.62%
3	−0.25	(−1.32)	(−1.66)*	1.55%	−2.49%
12	−0.78	(−1.27)	(−1.51)	3.40%	−2.44%
24	−2.04	(−2.01)**	(−2.36)**	12.62%	4.77%
36	−2.07	(−1.49)	(−1.93)*	9.80%	−14.67%
		1983:01–2013:12			
k	$\hat{\beta}$	t-ratio$_{Hodr}$	t-ratio$_{NW}$	R^2_{IS}	R^2_{OOS}
1	−0.11	(−2.09)**	(−2.11)**	1.32%	0.93%
3	−0.47	(−2.88)***	(−3.83)***	7.80%	6.32%
12	−1.82	(−2.92)***	(−5.65)***	29.04%	23.85%
24	−2.49	(−2.38)**	(−4.60)***	28.27%	26.85%
36	−2.89	(−2.11)**	(−3.61)***	26.93%	24.33%