Abstract

I focus my research on asset pricing with a focus on portfolio allocation and general equilibrium models. Portfolio allocation is of utmost importance for the industry. We show the importance of optimal portfolio allocation for large unbalanced equity data sets compared to naive diversification.
ESSAYS ON ASSET PRICING AND PORTFOLIO ALLOCATION

by

Sébastien COUPY

A thesis submitted to the
Geneva School of Economics and Management,
University of Geneva, Switzerland,
in fulfillment of the requirements for the degree of
PhD in Finance

Members of the thesis committee:
Prof. Tony BERRADA, Adviser, University of Geneva
Prof. Inès CHAIEB, Chair, University of Geneva
Prof. Olivier SCAILLET, University of Geneva
Prof. Marcel RINDISBACHER, Boston University

Thesis No. 25
July 2016
Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Prof. Tony Berrada for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Inès Chaieb, Prof. Olivier Scaillet, and Pr. Marcel Rindisbacher, for their insightful comments and encouragement.

I thank my colleagues for the stimulating discussions and for all the fun we have had in these past years. I am especially grateful to Christopher, Adrien, Benjamin, Paola and Tamara.

Last but not the least, I would like to thank my parents for supporting me throughout writing this thesis and my life in general.
Abstract

In my dissertation, I develop both empirical and theoretical results. In Chapter 1, I study the impact of modern portfolio theory on very large datasets and show that using individual stocks instead of portfolios of stocks increases significantly the out-of-sample performance of different investment strategies. In Chapter 2, I develop a general equilibrium model that can generate the cross-sectional variation as well as the time-varying variation of the pairwise correlation between financial stocks observed in the data. My findings have important implications for portfolio allocation.
Dans ma thèse, je développe des résultats théoriques ainsi qu’empiriques. Dans le chapitre 1, j’étudie l’impact de la théorie moderne de portefeuille sur de très grandes bases de données et démontre que l’utilisation d’actions individuelles à la place de portefeuilles d’actions augmente significativement la performance hors échantillon de différentes stratégies d’investissement. Dans le chapitre 2, je développe un modèle d’équilibre général qui est capable de générer la variation transversale ainsi que la variation temporelle de la corrélation entre les actifs financiers observée dans les données. Mes résultats ont d’importantes implications en terme d’allocation de portefeuille.
# Contents

Acknowledgments ................................................. i
Abstract ................................................................iii
Résumé ................................................................v
Introduction .......................................................... 1

1 Outperforming Naive Diversification Using Stock Level Information 5
  1.1 Introduction ....................................................... 5
  1.2 Mean-Variance Framework .................................. 8
     1.2.1 Setup ....................................................... 8
     1.2.2 Issues ..................................................... 9
  1.3 Moments estimation ............................................. 10
     1.3.1 Factor model estimation ................................. 10
  1.4 Alternative approaches ........................................ 13
  1.5 Empirical results ................................................ 14
     1.5.1 Strategies .................................................. 14
     1.5.2 Data and estimation methodology ...................... 15
     1.5.3 Out-of-sample performance evaluation ................. 16
     1.5.4 Empirical results ........................................ 18
  1.6 Extension and Robustness ..................................... 23
     1.6.1 Extension ................................................... 23
     1.6.2 Robustness ............................................... 31
  1.7 Conclusion ....................................................... 33

2 Pairwise Correlation Dynamics and Incomplete Information 35
  2.1 Introduction ....................................................... 35
  2.2 The model ......................................................... 37
     2.2.1 Uncertainty ............................................... 38
## CONTENTS

2.2.2 Firms ........................................... 38  
2.2.3 Representative agent .......................... 39  
2.2.4 Financial markets ............................. 40  

2.3 Equilibrium ....................................... 40  
2.3.1 State price density ............................. 41  
2.3.2 Stock price .................................... 42  
2.3.3 Stock returns and volatility .................... 42  
2.3.4 Market beta ................................... 44  
2.3.5 Cross-sectional correlation ..................... 44  

2.4 Theoretical implications of the model ...................... 45  
2.4.1 Calibration .................................... 45  
2.4.2 Theoretical implications ......................... 46  

2.5 Empirical results .................................. 50  
2.5.1 Data ........................................... 50  
2.5.2 Empirical patterns observed in the data ........ 52  
2.5.3 Model implied cross-correlation ................. 52  

2.6 Implication for portfolio allocation ..................... 62  
2.7 Conclusion ....................................... 70  

Conclusion ............................................. 71  

Appendix to Pairwise Correlation Dynamics and Incomplete Information 73
Introduction

My thesis consists of two individual essays. The first chapter of this dissertation is about portfolio allocation. Since the work of Markowitz (1952), the standard paradigm of portfolio allocation has revolved around the trade-off between risk and return. Modern portfolio theory establishes the set of optimal portfolios, which exploits the benefits of diversification. The efficient frontier simply results from the minimization of the portfolio variance given an expected portfolio return. This approach, while remarkably simple, requires reasonable estimates of the first and second moments of the excess returns’ distribution. A direct implementation using sample estimates of these moments provides an unstable and often extreme portfolio allocation (see Merton (1980)). These shortcomings have initiated various strands of literature, which address issues related to excess returns’ moments estimation and portfolio weights’ stability.

One direction that this literature\(^1\) has taken is to impose additional structure to the data generating process. Several models have been proposed in the literature, such as the CAPM of Sharpe (1964) and Lintner (1965), the 3-factor model introduced by Fama and French (1992; 1993) and its extension proposed by Carhart (1997).

The first chapter of this thesis contributes to the literature on portfolio optimization along several dimensions. We perform portfolio allocation on very large unbalanced panel data, namely the whole CRSP database. We show that Markowitz portfolio can offer significant benefits over naive diversification.

To be more precise, using a factor model structure and a novel estimation methodology introduced in Gagliardini, Ossola, and Scaillet (forthcoming), we implement the portfolio allocation using the entire CRSP database over the sample period (1967-2012) and show that it dominates naive diversification (1/N) in terms of out-of-sample Sharpe ratio (gross and net of transaction costs). This is remarkable as DeMiguel, Garlappi, and Uppal (2009b) demonstrate the surprising performance of the 1/N strategy.

The success of this approach is due to its ability to address three important problems that arise in large-scale portfolio allocation. First, estimation issues are mitigated by the two-pass approach à la Fama-MacBeth developed in Gagliardini et al. (forthcoming).

\(^1\)Brandt (2010) reviews the literature.
Second, stability issues are overcome by imposing additional structure to the data generating process. The introduction of a factor model simplifies the inversion of the large dimensional covariance matrix, which is instrumental in defining the optimal portfolio. Third, the methodology allows for unbalanced panel data, and therefore avoids discarding any information that may prove useful in estimating means and covariances of stock returns.

The results of this first chapter shed light on the importance of correlation for portfolio allocation. It is then natural to model correlation in order to improve portfolio performance. It is the goal of the second chapter of this thesis. Correlation is the key to portfolio diversification and therefore portfolio risk management. However, as the market goes into recession, when the need to control risk is presumably the highest, correlation goes up resulting in the least effect of portfolio risk reduction technique. This observation is true for correlation between individual stocks (Chordia, Goyal, and Tong 2011), between individual stocks and the market (Ang and Chen 2002, Hong, Tu, and Zhou 2007), and between international stocks (Erb, Harvey, and Viskanta 1994, Longin and Solnik 2001).

This chapter aims to provide a rational explanation to the phenomenon. We want to pinpoint the underlying structure that creates the important but intriguing patterns of volatility and correlation, namely time varying, clustering, and asymmetric to market conditions. We want to know whether correlation risk is priced. The answer will help the quest of understanding equity risk premia. Finally, we want to build efficient portfolios for investors.

We propose a rational expectation intertemporal equilibrium model of asset prices with aggregate and idiosyncratic risks, where information about the aggregate component is incomplete. Learning about the unobservable component creates business cycles that are conditioned by the time varying probability of a given state. Shocks to the idiosyncratic component affect stock characteristics conditional on the aggregate state of the economy, creating a dynamic pattern of cross sectional heterogeneity. In this model, a conditional version of the CAPM holds and only aggregate shocks carry a risk premium. However, idiosyncratic volatility is affected by the aggregate state of the economy, and therefore the cross section of correlations varies over time.

Our model generates results that are in line with empirical findings. On the time series dimension, both volatility of and correlations between stocks are higher during extreme market conditions\(^2\). On the cross sectional dimension, cross correlations within

\(^2\)Extensive research has shown the asymmetric pattern of volatility that arises from bad economic condition (see French, Schwert, and Stambaugh (1987), Schwert (1989), Campbell and Hentschel (1992), Bekaert and Wu (2000))
low beta stocks are lower than those within high beta stocks\textsuperscript{3}. Our efficient frontier, built for each investment horizon by proper conditioning, reduces significantly the risk of the global minimum variance portfolio.

Our study focuses on a simple cash flow risk model and is highly tractable. In addition, our model allows for relatively easy simulations, which in turns ease portfolio allocation.

\textsuperscript{3}See Ang and Chen (2002) and Chordia et al. (2011) for empirical results of the difference in correlations between low and high beta stocks
Chapter 1

Outperforming Naive Diversification
Using Stock Level Information

Tony Berrada$^1$ & Sébastien Coupé

Abstract. We construct mean-variance portfolios using a factor model approach. We show the importance of portfolio allocation for large unbalanced equity data sets using the full CRSP database. We compare the performance of our portfolio construction methodology to the 1/N naive diversification strategy, standard shrinkage procedures, and alternative factor model estimation. We document significant out-of-sample performance improvement in terms of Sharpe ratios, turnover and certainty equivalent. We show that it is due to improved expected returns estimation coming from the 2-pass regression approach.

JEL Classification: G11

Keywords: Mean-Variance, factor model, optimization, 2 stage least square

1.1 Introduction

Since the work of Markowitz (1952), the standard paradigm of portfolio allocation has revolved around the trade-off between risk and return. Modern portfolio theory establishes the set of optimal portfolios, which exploits the benefits of diversification. The efficient frontier simply results from the minimization of the portfolio variance given an
expected portfolio return. This approach, while remarkably simple, requires reasonable estimates of the first and second moments of the excess returns’ distribution. A direct implementation using sample estimates of these moments provides an unstable and often extreme portfolio allocation (see Merton (1980)). These shortcomings have initiated various strands of literature, which address issues related to excess returns’ moments estimation and portfolio weights’ stability.

There are two main directions that this literature has followed. In the first the focus is on correcting the sample covariance matrix and the mean vector. A prominent example of this approach is known as shrinkage, where the sample mean/covariance estimates are shrunk towards a target value. The latter is biased but has no estimation error. The methodologies vary in the way they trade off estimation error against bias. Important contributions in this area are Jorion (1985; 1986) who considers Bayesian approaches to correct both estimates and proposes the Bayes-Stein estimator and Ledoit and Wolf (2004a; b) who focus on the covariance matrix and propose several corrections. This paper is more closely related to Jorion (1985; 1986) as the method we use aims at correcting the sample mean of the excess returns distribution.

The second direction focuses on imposing additional structure to the data generating process. Sharpe (1964) and Lintner (1965) initiate this approach and develop the CAPM. In this model, a stock’s excess return is entirely defined by a coefficient $\beta$, representing its sensitivity to the market, and the excess market return. Ross (1976) generalizes this idea through the arbitrage pricing theory, where an arbitrary number of unspecified factors drive a security’s excess return. Fama and French (1992; 1993) introduce a 3-factor model, which adds the size (SMB) and book-to-market factors (HML) to the CAPM. Carhart (1997) extends the model by taking into account momentum as a fourth factor.

This paper contributes to the literature on portfolio optimization along several dimensions. We perform portfolio allocation on very large unbalanced panel data, namely the whole CRSP database. We show that Markowitz portfolio can offer significant benefits over naive diversification. In addition, we show that one can improve portfolio performance for small sample using information from full sample. In particular, computing risk premia from the complete information available leads to more stable results than restricting the computation to a subsample, even if an investor cannot invest in every stock available on the market. To be more precise, using a factor model structure and a novel estimation methodology introduced in Gagliardini et al. (forthcoming), we implement the portfolio allocation using the entire CRSP database over the sample period (1967-2012) and show that it dominates naive diversification (1/N) in terms of out-of-sample Sharpe ratio (gross and net of transaction costs). This is remarkable as DeMiguel et al.

\footnote{Brandt (2010) reviews the literature.}
(2009b) demonstrate the surprising performance of the 1/N strategy. To the best of our knowledge, very few papers use a large unbalanced panel of individual stocks for mean-variance portfolio allocation, a notable exception is Cosemans, Frehen, Schotman, and Bauer (2014) who analyze the global minimum variance portfolio in a one factor setting. Relying on a simple 3-factor (Fama French) or 4-factor (Carhart) model, out-of-sample Sharpe ratios range from 0.60 to 0.91 compared to 0.50 for naive diversification. Turnover is lower than (1/N) for global minimum variance portfolios and at most 1.7 times larger for tangency portfolios. Including transaction costs, the performance of the factor model methodology is still far better than the benchmark.

The success of this approach is due to its ability to address three important problems that arise in large-scale portfolio allocation. First, estimation issues are mitigated by the two-pass approach à la Fama-MacBeth developed in Gagliardini et al. (forthcoming). Second, stability issues are overcome by imposing additional structure to the data generating process. The introduction of a factor model simplifies the inversion of the large dimensional covariance matrix, which is instrumental in defining the optimal portfolio. Third, the methodology allows for unbalanced panel data, and therefore avoids discarding any information that may prove useful in estimating means and covariances of stock returns.

We find that out-of-sample Sharpe ratios are higher and turnover is lower for tangency portfolios when using the two-pass estimation methodology rather than the simple one-pass approach. We show that using cross-sectional information through a second pass estimation procedure, instead of individual regressions for each stock, mitigates noise and significantly improves the factor model’s ability to estimate expected returns.

When considering minimum variance portfolios with different specific target returns, we find that the approach developed in this paper leads to significantly improved results compared to the simple 1-pass regression. In particular, the 2-pass methodology is able to produce a broad spectrum of results - from 9% to 29% for excess returns and 9% to 67% for volatility - while the 1-pass model produces very similar results for portfolio excess returns and volatility, whichever the target excess return.

We show that using a suitable econometric model that allows for large panel of asset returns significantly improves the estimation of expected returns. We document an important increase in portfolio performance.

In order to compare the performance of the methodology to other strategies based on sample mean estimates or Bayesian techniques, we must restrict the sample to a balanced structure. We show that the factor model structure we propose allows extracting significant information from the individual stock return series and dominates the alternative strategies in term of Sharpe ratio (gross and net of transaction costs), certainty
equivalent, and turnover. We hope this will help restoring faith in the mean-variance framework.

The paper is organized as follows. In Section 1.2, we describe the Markowitz framework and its issues. In Section 1.3, we present the methodology behind the proposed portfolios. Section 1.4 contains a presentation of related literature. Section 1.5 presents the results. Section 1.6 proposes extensions and robustness checks. Section 1.7 concludes.

1.2 Mean-Variance Framework

1.2.1 Setup

In this section we summarize the main elements of the Markowitz (1952) approach to portfolio choice. This is a 1-period model with two dates 0 and $T$ in which investors care only about the mean and variance of their end of period wealth. There are $N$ financial assets in the economy. It is well known that the optimal portfolio weights are entirely defined by the first two moments of the excess return distribution, which we denote $\mu$ and $\Sigma$. Following the vast literature on this topic, we focus on three specific portfolios, namely the global minimum variance portfolio, $GMVP$, the tangency portfolio, $TP$, which produces the largest Sharpe ratio, and the minimum variance portfolio with a target excess return, $MVPT$. Formally, their respective relative portfolio weights are given by the following expressions

\begin{align}
  w_{TP} &= \frac{\Sigma^{-1}\mu}{\mathbf{1}^T\Sigma^{-1}\mu} \\
  w_{GMVP} &= \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} \\
  w_{MVPT} &= \frac{1}{D}\Sigma^{-1}(\mathbf{1}A - \mu B + (\mu C - \mathbf{1}B)E)
\end{align}

where $\mathbf{1}$ is the unit vector of dimension equal to the number of securities in the portfolio, $A = \mu'\Sigma^{-1}\mu$, $B = \mu'\Sigma^{-1}\mathbf{1}$, $C = \mathbf{1}'\Sigma^{-1}\mathbf{1}$, $D = BC - A^2$, and $E$ is the target excess return.

We assume that the excess returns take the following form

\[ R_{i,t} = a_i + b_i f_t + \epsilon_{i,t} \]  

$f_t$ can be any factors that characterize the returns. The no arbitrage asset pricing restriction implies that $a_i = b_i'\nu$ and Equation 1.3 can be rewritten

\[ R_{i,t} = b_i(\nu + f_t) + \epsilon_{i,t}. \]
The first two moments of the distribution are

\[ \mu = b'\lambda \]  \hspace{1cm} (1.4) \\
\[ \Sigma = b'\Sigma_f b + \Sigma_e \] \hspace{1cm} (1.5)

where \( \lambda = \nu + E(f_t) \) defines the risk premia, and \( \Sigma_f \) the factor variance-covariance matrix. Factor loadings \( b \) and risk premia \( \lambda \) are estimated following the procedure developed by Gagliardini et al. (forthcoming). It is a 2-pass estimation procedure that offers several advantages over existing alternatives such as naive diversification, standard one-pass estimation or shrinkage procedures (alternative approaches are described in Section 1.4). In particular, the 2-pass methodology produces accurate estimates of equity risk premia, which in turns allows for a better and more stable estimation of expected returns compared to a standard 1-pass approach. Additionally, it features a missing-at-random design that allows for datasets with missing values. This is of utmost importance when one looks at individual stocks instead of portfolio of stocks. The estimation procedure is described in details in Section 1.3.1.

The portfolio average excess return and variance are

\[ \mu_{pf} = w'\mu \] \\
\[ \sigma_{pf}^2 = w'\Sigma w \]

### 1.2.2 Issues

There are three types of issues that must be addressed in order to implement the methodology developed in the previous section: (i) Moment estimation errors (ii) Stability and (iii) Panel dimension.

Time series dimension is in most cases short, making the estimates noisy. Sample mean estimates are not admissible estimators (see Stein (1956), DeMiguel, Garlappi, Nogales, and Uppal (2009a)) and, although better, the sample covariances are not satisfactory either. As equations (1.1) and (1.2) show, the inverse of the covariance matrix is required to compute the optimal weights. It follows that errors are magnified through the optimization algorithm. As pointed out by Michaud (1989), one needs the input ‘to be more precise than the inaccurate information available’.

In addition, the inverse covariance matrix is subject to stability issue. The estimate becomes ill-conditioned as soon as the number of stocks increases in the neighborhood of the time dimension. A slight variation in the input might have a large effect on the optimal weight distribution. Most authors test strategies on data sets where they group individual stocks into portfolios of stocks beforehand to reduce the cross-section
dimension in order to overcome this issue.

Lastly, available data at the individual stock level are not exempt of errors and omissions. One will most likely face a database containing missing values. When considering portfolios of stocks, the problem is alleviated. In this paper, we are interested in portfolios based on individual stocks. Hence, we need a model allowing for unbalanced panel data.

The use of a factor model gives a structure to the returns and stabilizes the problem. The methodology proposed in this paper based on Gagliardini et al. (forthcoming) fixes estimation issues and allows for unbalanced data sets. In particular, this method improves significantly the estimation of the risk premia, hence the first moment of the excess returns distribution.

1.3 Moments estimation

1.3.1 Factor model estimation

We follow a three-step approach. In a first step, we estimate the factor loadings $b$ and risk premia $\lambda$ following Gagliardini et al. (forthcoming). Secondly, we reconstruct the average excess return $\mu$ and the excess return covariance matrix $\Sigma$ according to equation (1.4) and (1.5). In a third step, we compute the inverse covariance matrix using the inversion lemma, also called the Sherman-Morrison-Woodbury identity. We assume a strict factor model. We took into consideration a non-strict factor model and applied soft-thresholding technique following Fan, Liao, and Mincheva (2011) but the results are at best similar and most of the time worse.

The procedure to compute the estimates of $a_i$, $b_i$ and $\lambda$ consists of a 2-pass estimation à la Fama and MacBeth (1973). Gagliardini et al. (forthcoming) derive large sample properties for large equity data sets. The model allows both for a conditional and an unconditional setting. We choose to use the latter with a rolling window.

The methodology in Gagliardini et al. (forthcoming) produces accurate estimates of equity risk premia even if time-series estimators of the factor loadings are noisy. In addition, it is simple to implement and is numerically tractable. Most importantly, it supports unbalanced data sets through a missing-at-random design.

We now formally introduce the model. We can rewrite equation (1.3) as $R_t = \beta'x_t + \epsilon_t$ where $\beta = [a; b]$ and $x = [1; f]$. An indicator function $I_{i,t}$ accounts for the unbalanced nature of the panel. It takes a value of 1 if the return of stock $i = 1, ..., N$ exists at time $t$ and zero otherwise. The no-arbitrage condition implies $a = b'\nu$, which is equivalent to $E(R_t) = b'\lambda$ where $\lambda = \nu + E(f_i)$ is a vector of risk premia. It is worth noting that in
case of tradable factor, the theory suggests that risk premia should be equal to the factor means, \( E(f_t) \). In practice, however, for standard factors such as the Fama-French factors and momentum, the implementation of such portfolios faces high transaction costs due to frequent rebalancing and short selling constraints. A non-zero \( \nu \) might capture these market imperfections (see Cremers, Petajisto, and Zitzewitz (2012) and Frazzini, Israel, and Moskowitz (2012)). Eventually, we have \( V(R_t) = V(b'f_t) + V(\epsilon_t) = b'\Sigma_f b + \Sigma_\epsilon \) where \( \Sigma_f \) is the factor covariance matrix, and \( \Sigma_\epsilon \) the covariance matrix of the residuals.

In the first pass, we compute:

\[
\hat{\beta}_i = \hat{Q}_{x,i}^{-1} \frac{1}{T_i} \sum_i I_{i,t} x_t R_{i,t}
\]

where \( \hat{Q}_{x,i} = \frac{1}{T} \sum_{t} I_{i,t} x_t x'_t \), and \( T_i = \sum_t I_{i,t} \) is the total number of observation for stock \( i \). In order to take into account misspecifications, a trimming device \( 1^n_i \) that only keeps stocks with a) enough observations and b) a time-series regression reasonably scaled is introduced. We measure the latter by computing the condition number of matrix \( \hat{Q}_{x,i} \), which is defined as \( \text{CN}(\hat{Q}_{x,i}) = \sqrt{\text{eig}_{\text{max}}(\hat{Q}_{x,i})/\text{eig}_{\text{min}}(\hat{Q}_{x,i})} \), where \( \text{eig}_{\text{max}} \) and \( \text{eig}_{\text{min}} \) are the maximum and minimum eigenvalue respectively. This number indicates multicollinearity problems. The trimming device takes a value of one when both the condition number is lower than 15 (see Greene 2008) and the ratio \( \tau_{i,T} = T/T_i \) is lower than 10.\(^3\) It is worth noting that \( \tau \) is a lower bound to reduce noise and over-fitting. Increasing this number improves results as a lower number of stocks are included in Markowitz portfolio, leading to a lower turnover, which will come as an advantage compared to naive diversification where all stocks are included.

The second pass focuses on the estimation of \( \nu \). The model is based on a weighted least square (WLS) approach and is performed by regressing the \( \hat{\alpha}_i \)'s on the \( \hat{\beta}_i \)'s while keeping the non-trimmed assets only. The weights are given by \( w_i = v_i^{-1} \) with the estimates of \( v_i \) being

\[
\hat{v}_i = \tau_{i,T} c'_{i1} \hat{Q}_{x,i}^{-1} \hat{S}_{ii} \hat{Q}_{x,i}^{-1} c_{i1}
\]

where \( \hat{S}_{ii} = \frac{1}{T} \sum_{t} I_{i,t} \hat{\epsilon}_{i,t} x_t x'_t \), \( \hat{\epsilon}_{i,t} = R_{i,t} - \hat{\beta}_i x_t \) and \( c_{i1} = (1, -\nu'_i)' \).

From the OLS estimator \( \hat{\nu}_i = (\sum_i 1^n_i \hat{b}_i \hat{b}_i')^{-1} \sum_i 1^n_i \hat{b}_i \hat{\alpha}_i \), we are able to estimate \( c_{i1} \), which is a first-step estimator with unit weights.

\(^3\)120/12 which is the length of the rolling window divided by the number of month in a year, arbitrarily chosen.
The WLS estimator is thus given by

$$\hat{\nu} = \hat{Q}_b^{-1} \frac{1}{N} \sum_i \hat{w}_i \hat{b}_i \hat{a}_i$$

where $\hat{Q}_b = \frac{1}{N} \sum_i \hat{w}_i \hat{b}_i \hat{b}_i'$ and $\hat{w}_i = 1_N \hat{\nu}_i^{-1}$. The advantage of using a WLS approach is that the statistical precision of the first pass is taken into account by the weights attributed to each asset.

The estimator of risk premia is then

$$\hat{\lambda} = \hat{\nu} + \frac{1}{T} \sum_t f_t$$

The fast convergence rate of $\hat{\nu}$ is $\sqrt{nT}$. Thus, a second-order asymptotic bias shows up when $N$ gets large. Gagliardini et al. (forthcoming) propose the correction

$$\hat{\lambda}_B = \hat{\nu} - \frac{1}{T} \hat{B}_\nu + \frac{1}{T} \sum_t f_t$$

where $\hat{B}_\nu = \hat{Q}_b^{-1}(\frac{1}{N} \sum_i \hat{w}_i \hat{r}_t, T E^t \hat{Q}_x^{-1} \hat{S}_t \hat{Q}_x^{-1} c_{\nu})$ and $E_2 = (0 : I_{d_K})'$.

Equipped with risk premia estimates ($\lambda$, common to all stocks) and factor loadings ($\beta_t$, specific to each stock), we are able to reconstruct the mean excess return vector as well as the covariance matrix of returns $\hat{\Sigma}$ for each point in time $t$. To obtain the vector of mean, we use the asset pricing restriction $E(R_t) = b'\lambda$, which gives $\hat{R} = \hat{b}'\hat{\lambda}$. Similarly, we obtain the relationship $V(R_t) = \hat{\Sigma} = b'\Sigma f' + \Sigma_e$ for the covariance matrix where $\Sigma_f$ is the factor covariance matrix and $\Sigma_e$ is the residual variance.

In a portfolio allocation context, one is interested in the precision matrix more than in the covariance matrix itself. We construct it without directly inverting the covariance matrix by using the Sherman-Morrison-Woodbury identity ($SMW$). The main feature of the formula is that it only requires an inversion of a $K$ by $K$ matrix and a sparse matrix, improving both precision and computational time. In our case, $\Sigma_f$ is a squared full rank matrix as long as $K < T$, which is a special case of the SMW formula. We obtain the following decomposition

$$\hat{\Sigma}^{-1} = (\Sigma_e + b'\Sigma_f b)^{-1} = \Sigma_e^{-1} - \Sigma_e^{-1} b' (\Sigma_f^{-1} + b \Sigma_e^{-1} b')^{-1} b \Sigma_e^{-1}.$$

To summarize, the procedure described in this section addresses stability issues by imposing a structure to the excess return. It handles the unbalanced nature of panel data through the procedure detailed in Gagliardini et al. (forthcoming). The methodology
addresses the main three issues described earlier, with a focus on the estimation of the first moment of the excess return distribution, and we are confident that by choosing adequate factors, one can mitigate estimation issues as well.

1.4 Alternative approaches

An important body of literature addresses estimation issues and factor models are one among many ways of dealing with the implementation issues of the Markowitz approach. In order to assess the benefits of our methodology, we compare the performance of our approach to well-known strategies.

A first natural alternative to consider is sample estimation of the two moments. While sample estimates are unbiased they are subject to estimation error as discussed in the previous section. The covariance matrix being singular as soon as \( N > T \), the scope of such a method is restrained in empirical applications given the limited available data.

A prominent approach to address estimation and stability issues are the so-called ‘shrinkage’ methods. The idea is to shrink the estimated moments toward a specific target, which might be far from the true distribution, but does not contain errors. A combination of the sample and the target produces a trade-off between error and information.

One of the first to investigate the field is Jorion (1985; 1986) who, based on the work of Stein (1956) and James and Stein (1961), proposes to correct the sample mean by averaging it with the minimum variance portfolio mean. The author also suggests one shrink the sample covariance matrix in a Bayesian fashion. Ledoit and Wolf (2004a; b) focus on the minimum variance portfolio and hence on the regularization of the covariance matrix only. They advocate a scaled identity matrix or a constant correlation matrix as suitable choices.\(^4\)

Several authors propose alternative strategies by adding constraints to the portfolio optimization problem or by neglecting some estimates. As an example, Jagannathan and Ma (2003) add a no short selling constraint, which significantly reduces turnover. The strategy relies however on sample estimation and is subject to the standard shortcomings. Kirby and Ostdiek (2009) obtain good results by neglecting covariances. This approach ignores important information, but its performance is better than previous portfolio methods, both in term of Sharpe ratio, certainty equivalent and turnover.

DeMiguel et al. (2009b) show that naive diversification is a very successful strategy. Surprisingly, investing equally in every stock generates a superior performance. We consider this portfolio as our benchmark throughout the paper. The main advantages are

\(^4\)We choose the latter in the empirical part since it performs better. Matlab routine freely available from Wolf’s website.
low turnover and the absence of short selling. On the other hand, it never benefits of obvious correlation between stocks. In particular, one can expect it to perform well when stocks are highly correlated but not otherwise.

The main advantage of those algorithms over factor models is that they do not impose any ex-ante assumptions on the returns’ structure. By assuming a factor model for the returns, one makes the hypothesis that the chosen factors can explain the returns’ dynamics. A natural comparison with our proposed technique is a simple one-pass linear regression. In the standard single pass approach, for each stock available at the date of the portfolio setup, we simply regress the excess returns on the factor(s) and we use

\[ E(r) = a_i + b_i E(f) \]

as an estimate of the mean excess return. The only information used in this case is at an individual level and not from the full panel, whereas in our case \( a_i = b_i \nu \) with \( \nu \) common to all stocks. In the single pass approach, the noise from the estimate \( a_i \) is not mitigated by a second pass. Since the estimates are obtained by individual regression on each stock, there is no benefit from the information available at the cross-sectional level of the panel data.

### 1.5 Empirical results

#### 1.5.1 Strategies

Our focus in this paper is on the ability of a factor model specification of the excess returns to improve portfolio allocation performance through increased parameters’ stability and better estimation.

We consider three factor models that are all standard in the literature. The first model has a single factor, the excess market return (ERM). The second model is the Fama-French three-factor model (Fama and French (1993)) which includes the excess market return (ERM), the Small-minus-Big (SMB) size factor and High-minus-Low (HML) growth factor. Finally we follow Carhart (1997) and add a momentum (MOM) factor to the excess return specification. The factor model specifications are denoted 1-factor, 3-factor and 4-factor, respectively.

For each model we implement 3 strategies: (i) the Global Minimum Variance Portfolio (GMVP), (ii) the Tangency Portfolio (TP) and (iii) the Minimum Variance Portfolio with Target (MVPT).

---

\(^5\) Zhang (2012) studies the properties of the tangency portfolio with the addition of a no short-sell constraint in a one factor model.
Table 1.1: List of models

<table>
<thead>
<tr>
<th>Model</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benchmark</strong></td>
<td></td>
</tr>
<tr>
<td>1. Naive diversification</td>
<td>1/N</td>
</tr>
<tr>
<td>2. Tangency portfolio</td>
<td>TP</td>
</tr>
<tr>
<td>3. Minimum variance portfolio</td>
<td>GMVP</td>
</tr>
<tr>
<td>4. Minimum variance portfolio with target return</td>
<td>MVPT</td>
</tr>
<tr>
<td><strong>Shrinkage methods</strong></td>
<td></td>
</tr>
<tr>
<td>5. Minimum variance with short sell constraint</td>
<td>MVc</td>
</tr>
<tr>
<td>6. Bayes-Stein portfolio</td>
<td>BS</td>
</tr>
<tr>
<td>7. Ledoit-Wolf portfolio</td>
<td>LW</td>
</tr>
<tr>
<td><strong>Alternatives options</strong></td>
<td></td>
</tr>
<tr>
<td>8. Volatility timing strategy</td>
<td>VT</td>
</tr>
<tr>
<td>9. Risk-and-reward timing strategy</td>
<td>RRT</td>
</tr>
</tbody>
</table>

This table represents all strategies used throughout the paper.

We choose different targets for $MVPT$ to analyze its behavior. We set the target equal to 5%, 15%, and 9%, which corresponds to the observed out-of-sample average excess return of the $1/N$ strategy.

An important benchmark to consider is the Naive diversification ($1/N$) portfolio developed in DeMiguel et al. (2009b). While remarkably simple, it generates an excellent performance both in term of Sharpe ratio and turnover.

We list all strategies used throughout the paper in Table 1.1.

1.5.2 Data and estimation methodology

As soon as the number of stocks $N$ approaches the number of observations $T$, the covariance matrix becomes singular. Most studies in the literature consider portfolios of stocks instead of individual stocks for portfolio allocation problems. Ang, Liu, and Schwarz (2008) show that pre-grouping stocks is not optimal. Information is lost and a bias is introduced. Hence, we choose to consider individual stocks in this paper. In particular, we take the entire Centre of Research in Security Price (CRSP) database from 1967 to 2012 into account. We obtain returns for over 20,000 securities spread over 552 months ($T$), totalizing two million observations. We clean the database by deleting securities having an entry corresponding to an error in CRSP (i.e. 44.0/55.0/.../99.0) and having

---

6a notable exception is Zhang (2012) who uses a single pass estimation approach and a one factor model on a subsample of CRSP.
returns in the 0.05% highest return region. Finally, we keep only securities with a share code equal to 10 or 11, which represents stocks. Our sample consists of 18,356 stocks and contains 4089 stocks on average. We split the sample into subsamples to check for discrepancies between time periods.

On the estimation side, we subtract the 30 days T-Bill risk-free rate to get excess returns. We then estimate the two moments of interest, i.e. the mean and the covariance matrix, every month based on a rolling window of $M$ months. There is a trade-off between choosing large estimation window for accuracy and short estimation for relevance as market conditions are likely to change, the smallest window is advisable for relevance but not for accuracy. We consider a window length of 120 months. We discuss the results for a 60 months and 180 months windows in Section 1.6.

We obtain 432 ($T - M$) out-of-sample estimators of $\mu_{pf}$ and $\Sigma_{pf}$. Taking the sample mean and the square root of the variance leads to the average excess return and the volatility of a given strategy. We compute turnover using the 432 weight vectors.

We consider the HAC estimator developed by Newey and West (1994) with ten lags to compute the factor covariance matrix $\Sigma_f$ due to the presence of autocorrelation in the factors (see Gagliardini et al. (forthcoming)).

In line with Cremers et al. (2012) findings, the coefficient $\nu$ capturing market imperfections is different from zero and hence, improves expected return estimation.

\subsection{Out-of-sample performance evaluation}

We compute several criteria to assess portfolio performance. Out-of-sample Sharpe ratio (SR) measures the excess average return over the volatility of a given strategy. We have:

$$SR_{pf} = \frac{\mu_{pf}}{\sigma_{pf}}$$

We also compute the ‘Certainty Equivalent ($CE$\textsuperscript{10}$’), also known as the mean-variance criterion, under the assumption of quadratic utility. The $CE$ is defined as:

$$CE = \mu_{pf} - 0.5\gamma \sigma_{pf}^2$$

where $\gamma$ is the risk aversion parameter. We set $\gamma$ to one and five throughout the rest

---

\textsuperscript{7}Removing the highest returns prevents large entry errors to affect results. The final monthly return range is -100% to +100% approximately. A change in this criterion doesn’t affect the results.

\textsuperscript{8}We show the periods 1967-2012 and 1988-2012 in this paper, other periods lead to similar results.

\textsuperscript{9}Computation of this variable using the sample covariance operator leads to similar results. These results are available upon request.

\textsuperscript{10}We test the statistical significance for $H_0 : CE_x = CE_y$ using an adjustment to the technique developed by Ledoit and Wolf (2008)
of the paper. A gamma of one \((CE_1)\) corresponds to log-utility and penalizes less the benchmark strategy, which is not optimized to have a low volatility. A gamma of five \((CE_5)\), as in Kirby and Ostdiek (2009), allows comparing strategies from a risk averse investor point of view. The \(CE\) represents the minimum sure return at which an investor would be willing to abandon a risky opportunity. The certainty equivalent differs from the Sharpe ratio by explicitly taking into account the investors attitude towards risk.

We include average monthly turnover and use it as a proxy for transaction costs. It is computed as follows:

\[
\tau_{pf} = \frac{1}{T - M} \sum_{t=M}^{T-1} \sum_{i=1}^{N} \left( |\hat{w}_{i,t+1} - \hat{w}_{i,t+1}| \right)
\]

where \(w_{i,t+1}\) is the weight of stock \(i\) in the portfolio just before rebalancing. The total cost of one transaction can vary greatly for different market capitalizations. There is also significant variation through time. Furthermore, trading stocks on NYSE or NASDAQ entails different costs. In order to account for all these factors, we compute Sharpe ratios after transaction costs using a range of possible candidates, from 50 basis points to 120 basis points. This range reflects transactions costs level reported in different studies. Keim and Madhavan (1997) find that buy orders incur trading costs of 50 basis points while sell orders incur costs of 55 basis points for the period 1991-1993. They note, however, that trading on NASDAQ costs between 123 and 143 basis points. Scaillet and Bajgrowicz (2012) show that on shore accounts face fees between 31 and 126 basis points. Using mutual funds transactions data, Edelen, Evans, and Kadlec (2013) compute an average trading cost of 1.44% per year, and find that, for most trading strategies, funds trading small caps faces transaction costs 32 basis point higher than those trading large caps. Using a sample of approximately 150 NASDAQ firms and 150 NYSE/Amex firms selected in a capitalization-stratified random draw in each of the years from 1993 to 2005, Hasbrouck (2009) finds an average trading cost ranging from 106 bps to 112 bps depending on the estimation methodology.

While technological progress have made explicit transaction costs lower over time, the total cost, which includes the implementation shortfall (market impact and timing delay cost), varies significantly across periods. ITG, an independent broker and financial technology provider, publishes regular reports (ITG (2009; 2013)) on trading costs and found that trading costs for small caps was around 80 bps for the period 2005-2009, but increased to 120 bps (even 160 bps for micro caps) for the period 2009-2013.

We subtract \(\tau \times c\) to the sample mean portfolio excess return in order to get the out-of-sample Sharpe ratio net of transaction costs. A large turnover negatively impact portfolio performance due to market trading fees. This measure is of utmost importance
for traders. It might not be viable to implement a strategy with a large Sharpe ratio or CE if the turnover is high.

We measure portfolio weight concentration using the Herfindahl index (hereafter HI), 
\[ HI = \sum_{i=1}^{N} w_i^2. \] This measure is important as it can help understand large swing in portfolio performance. Tangency portfolio without constraint is notoriously assigning enormous weights on some stocks.

In addition, we use a measure called equivalent transaction cost (ETC later on) which represents the transaction cost level in basis point such that the strategy’s SR is equal to the SR obtained through naive diversification (1/N). For example, an ETC of 0.005 means that an investor facing transaction costs of 50 basis points will be indifferent between naive diversification and the model under scrutiny on the basis of Sharpe ratio.

### 1.5.4 Empirical results

In this section we report the performance of the individual stocks sample. Table 1.2 and 1.3 depict the results from the largest and longest database under consideration, which represents 18,356 stocks and spans the time from January 1967 to December 2012.\(^{11}\) Naive diversification over the period gives an average excess return of 8.6% for a volatility of 17.2%. The Sharpe ratio is 49.8% and the average monthly turnover 10.5%, leading to out-of-sample Sharpe ratios net of transaction costs ranging from 46.2% to 41%.

The 2-pass approach uses the entire panel data to identify the factor structure, including time series of stocks that are not necessarily present at the time of portfolio construction. It allows incorporating more information than the standard approach resulting in a significant performance gain.

Table 1.2 presents results for tangency portfolios constructed with two different approaches, the 1-pass and 2-pass methodologies, and within each approach using three factor models, namely a 1-factor, 3-factor and 4-factor model. Results are similar between factor models. Sharpe ratios are much higher and turnover much lower for the 2-pass approach than fro the 1-pass approach. The impact of turnover on Sharpe ratios after transaction costs is striking, leading to an even stronger gain using the 2-pass approach. We test the statistical significance of the difference in Sharpe ratios by testing \( H_0 : SR_x = SR_y \) using the test derived by Ledoit and Wolf (2008). We perform the test for three portfolio pairs, i.e. between \( 1/N \) & 1pass, \( 1/N \) & 2pass and 1pass & 2pass. Every factor model portfolio outperforms the benchmark in term of Sharpe ratio before transaction costs, regardless of the methodology. The 2-pass methodology provides similar risk-reward characteristics across factor models, with an average excess return of 10%

\(^{11}\) Additional subsamples results are available upon request.
Table 1.2: All CRSP between 1967 and 2012, tangency portfolio

<table>
<thead>
<tr>
<th></th>
<th>( \mu )</th>
<th>( \sigma )</th>
<th>( SR )</th>
<th>( \tau )</th>
<th>( SR_{at,50} )</th>
<th>( SR_{at,80} )</th>
<th>( SR_{at,100} )</th>
<th>( SR_{at,120} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benchmark</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.09</td>
<td>0.17</td>
<td>0.50</td>
<td>0.10</td>
<td>0.46</td>
<td>0.44</td>
<td>0.43</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>Factor models, 1-pass</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(_1)-factor</td>
<td>0.10</td>
<td>0.18</td>
<td>0.54</td>
<td>0.49</td>
<td>0.38</td>
<td>0.28</td>
<td>0.22</td>
<td>0.15</td>
</tr>
<tr>
<td>p-value(_1N-1p)</td>
<td>(0.80)</td>
<td></td>
<td>(0.65)</td>
<td>(0.39)</td>
<td>(0.23)</td>
<td>(0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(_3)-factor</td>
<td>0.09</td>
<td>0.17</td>
<td>0.52</td>
<td>0.54</td>
<td>0.33</td>
<td>0.22</td>
<td>0.15</td>
<td>0.07</td>
</tr>
<tr>
<td>p-value(_1N-1p)</td>
<td>(0.92)</td>
<td></td>
<td>(0.40)</td>
<td>(0.19)</td>
<td>(0.09)</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(_4)-factor</td>
<td>0.09</td>
<td>0.17</td>
<td>0.54</td>
<td>0.52</td>
<td>0.36</td>
<td>0.24</td>
<td>0.17</td>
<td>0.10</td>
</tr>
<tr>
<td>p-value(_1N-1p)</td>
<td>(0.80)</td>
<td></td>
<td>(0.51)</td>
<td>(0.23)</td>
<td>(0.12)</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Factor models, 2-pass</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(_1)-factor</td>
<td>0.10</td>
<td>0.17</td>
<td>0.63</td>
<td>0.09</td>
<td>0.49</td>
<td>0.58</td>
<td>0.56</td>
<td>0.55</td>
</tr>
<tr>
<td>p-value(_1N-2p)</td>
<td>(0.01)</td>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value(_1p-2p)</td>
<td>(0.60)</td>
<td></td>
<td>(0.18)</td>
<td>(0.07)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(_3)-factor</td>
<td>0.10</td>
<td>0.17</td>
<td>0.62</td>
<td>0.13</td>
<td>0.58</td>
<td>0.55</td>
<td>0.53</td>
<td>0.51</td>
</tr>
<tr>
<td>p-value(_1N-2p)</td>
<td>(0.01)</td>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value(_1p-2p)</td>
<td>(0.44)</td>
<td></td>
<td>(0.09)</td>
<td>(0.03)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(_4)-factor</td>
<td>0.10</td>
<td>0.17</td>
<td>0.61</td>
<td>0.17</td>
<td>0.55</td>
<td>0.51</td>
<td>0.49</td>
<td>0.46</td>
</tr>
<tr>
<td>p-value(_1N-2p)</td>
<td>(0.07)</td>
<td></td>
<td>(0.16)</td>
<td>(0.24)</td>
<td>(0.32)</td>
<td>(0.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value(_1p-2p)</td>
<td>(0.63)</td>
<td></td>
<td>(0.15)</td>
<td>(0.05)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table represents the results obtained on the whole CRSP database. 18,356 stocks are taken into account. The first column gives the average return of the strategy over the period 1977-2012 (the first 10 years being used for estimation). \( \sigma \) is the standard deviation of each portfolio. \( SR \) stands for Sharpe ratio and \( SR_{at,X} \) is the Sharpe ratio after transaction costs where \( X \) represents the amount of transaction costs in basis point. \( \tau \) is the turnover required by a given strategy. Numbers in parenthesis represents p-value for the test \( H_0: SR_y = SR_z \) where \( y \) and \( z \) are 1/N&1pass for the 1-pass model and 1/N&2pass and 1pass&2pass for the second part. All numbers are annualized.

and a standard deviation of 16.5%, both better or equal to their 1-pass counterparts. It significantly improves Sharpe ratios compared to the 1/N benchmark and to the 1-pass portfolios with a difference larger than 20%.

It clearly appears from this table that the 1-pass methodology does not statistically differ from the 1/N benchmark once we properly account for transaction costs. It also appears that the 2-pass methodology is statistically different from both the 1/N benchmark and the 1-pass methodology once we properly account for transaction costs. In particular at the level of 80bps, which is a conservative estimate based on the literature, when comparing the 1-pass and 2-pass methodology the P-values are 0.07, 0.03 and 0.05 for the 1, 2 and 3 factor models respectively. The economic significance is also quite
important, for example using the 1-factor model at the same 80bps level, the after cost Sharpe ratio goes from 0.28 for the 1-pass approach to 0.58 for the 2-pass approach. For all models, at a transaction cost level of 100 or 120 bps, the 2-pass methodology is statistically superior to the 1-pass methodology at the 5% confidence level, with P-values ranging from 0 to 0.04.

Figure 1.1: Figure representing the cumulative return of tangency portfolios under the assumption of the four factor model based on the whole CRSP dataset (18,356 stocks). The blue line represents the results with a 1-pass estimation procedure. The green line is the results with the 2-pass estimation developed in this paper.

Figure 1.1 shows the cumulative returns of the two strategies. The two-pass regression outperforms the one-pass strategy almost continuously. Following the 2-pass model for tangency portfolio leads to a 50% higher cumulative return by the end of the period. The picture is different with the MVPT algorithm. Sharpe ratios are higher using a 1-pass regression than a 2-pass.

We can see in Table 1.2 that the 1-factor models performs better. In order to study the minimum variance portfolios, we only focus on the 1-factor model with six targets, i.e. 1%, 5%, 9%, 15%, 25% and 40%. Table 1.3 presents the results. The 1-pass approach produces
Table 1.3: All CRSP between 1967 and 2012, MVTP strategy

<table>
<thead>
<tr>
<th>Factor models, 1-pass</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\tau$</th>
<th>$\mu_{at,50}$</th>
<th>$\mu_{at,80}$</th>
<th>$\mu_{at,100}$</th>
<th>$\mu_{at,120}$</th>
<th>$HI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVPT$_{1%}$</td>
<td>0.09</td>
<td>0.11</td>
<td>0.15</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{5%}$</td>
<td>0.09</td>
<td>0.11</td>
<td>0.11</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.09</td>
<td>0.11</td>
<td>0.09</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{15%}$</td>
<td>0.09</td>
<td>0.11</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{25%}$</td>
<td>0.10</td>
<td>0.12</td>
<td>0.17</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
<td>0.002</td>
</tr>
<tr>
<td>MVPT$_{40%}$</td>
<td>0.10</td>
<td>0.14</td>
<td>0.32</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor models, 2-pass</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\tau$</th>
<th>$\mu_{at,50}$</th>
<th>$\mu_{at,80}$</th>
<th>$\mu_{at,100}$</th>
<th>$\mu_{at,120}$</th>
<th>$HI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVPT$_{1%}$</td>
<td>0.09</td>
<td>0.09</td>
<td>0.14</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
<td>0.002</td>
</tr>
<tr>
<td>MVPT$_{5%}$</td>
<td>0.11</td>
<td>0.12</td>
<td>0.11</td>
<td>0.10</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.13</td>
<td>0.18</td>
<td>0.17</td>
<td>0.12</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{15%}$</td>
<td>0.16</td>
<td>0.27</td>
<td>0.10</td>
<td>0.15</td>
<td>0.14</td>
<td>0.13</td>
<td>0.12</td>
<td>0.004</td>
</tr>
<tr>
<td>MVPT$_{25%}$</td>
<td>0.21</td>
<td>0.43</td>
<td>0.52</td>
<td>0.18</td>
<td>0.17</td>
<td>0.15</td>
<td>0.14</td>
<td>0.015</td>
</tr>
<tr>
<td>MVPT$_{40%}$</td>
<td>0.29</td>
<td>0.67</td>
<td>1.13</td>
<td>0.23</td>
<td>0.18</td>
<td>0.16</td>
<td>0.13</td>
<td>0.046</td>
</tr>
</tbody>
</table>

This table presents the results obtained on the whole CRSP database from the minimum variance portfolio with a target excess return. 18,356 stocks are taken into account. The first column gives the average return of the strategy over the period 1977-2012 (the first 10 years being used for estimation). $\sigma$ is the standard deviation of each portfolio. $\tau$ is the turnover required by a given strategy. $\mu_{at,X}$ represents the average excess return after transaction costs of $X$ basis points. $HI$ stands for Herfindahl index. All numbers are annualized.

almost no variation in measured average returns for very different target expected returns, while using the 2-pass approach yields very significant variations.

Using the 1-pass approach the average return ranges from 9% to 10% for the target expected returns ranging from 1% to 40%. Using the 2-pass approach the average return ranges from 9% to 29%. This provides further evidence that the 2-pass methodology helps in estimating the proper risk premia and improves the precision of the estimated expected return. We test $H_0: \mu_1 = \mu_6$ using 1-way ANOVA and Kruskal-Wallis tests and obtain that the difference is statistically significant for the 2-pass approach while it is not for the 1-pass approach.\(^{12}\)

We plot the total excess return of each strategy in Figure 1.2. Using a rolling window of 120 months - 10 years - does not allow avoiding the 2008 crisis. All strategies are hardly hit. We note, however, that naive diversification is no better at avoiding the drawdown of 2008. The larger impact on the Markowitz portfolios is due to their composition. They are more concentrated than naive diversification and are more prone to larger swings due to the state of the economy.

It is by comparing the performance of the $TP$ strategy in Figure 1.2 that we can best

\(^{12}\)Results are similar for 3-factor and 4-factor models.
identify the contribution of our methodology. The estimation of the mean excess return has a direct impact on the performance of the TP strategy and we can see that a richer factor structure significantly improves the estimation and allows the portfolio to perform much better than the $1/N$ benchmark.

Figure 1.3 shows the evolution of the number of stocks that we are invested in throughout the investment period 1977-2012 (the first 10 years are used for calibration when considering a rolling window of 120 months). We invest in more stocks using a naive algorithm than with sophisticated models. The difference between the two curves comes from the trimming device introduced in Section 1.3.1. It amounts to 400 stocks at the beginning of the sample. The difference is increasing from 1967 to 1997 and decreasing thereafter and eventually stabilizes at around 500 stocks between 2001 and 2012. Fewer stocks are available for investment and the trimming device lets a larger proportion of investible stocks to pass in the portfolio. It is worth noting that naive diversification implies investing in all available stocks at all time, the number of stocks included in the portfolio is thus equal to total number of stocks available on CRSP.
One of the main caveats with unconstrained Markowitz portfolios is that the weights tend to be unrealistic. We find that the average proportion of short positions for the period 1967-2012 is around 15% of the total weights and that the largest short position in a single stock is approximately 1%. We believe that large institutional investors are able to take such positions.

1.6 Extension and Robustness

1.6.1 Extension

In this section, we study the performance of different subsamples of the CRSP database, as well as different portfolios found in the literature based on portfolios of stocks. Furthermore, we postulate that the most precise risk premium estimates are computed using the largest data set available, the full CRSP database. It seems realistic to believe that the more information we have, the more precise the estimates. Thus, we replace the risk premia $\lambda$ at each point in time $t$ for all data sets previously studied by the ones computed
Table 1.4: Random size samples

<table>
<thead>
<tr>
<th></th>
<th>100s SR</th>
<th>200s SR</th>
<th>1000s SR</th>
<th>2000s SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.65 0.08</td>
<td>0.67 0.08</td>
<td>0.69 0.08</td>
<td>0.70 0.08</td>
</tr>
<tr>
<td>Factor models, 2-pass</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-factor TP</td>
<td>0.61 0.07</td>
<td>0.63 0.07</td>
<td>0.64 0.07</td>
<td>0.65 0.07</td>
</tr>
<tr>
<td>MVPT_{9%}</td>
<td>0.58 0.45</td>
<td>0.64 0.29</td>
<td>0.69 0.16</td>
<td>0.70 0.16</td>
</tr>
<tr>
<td>3-factor TP</td>
<td>0.37 1.25</td>
<td>0.46 2.18</td>
<td>0.58 0.16</td>
<td>0.60 0.14</td>
</tr>
<tr>
<td>MVPT_{9%}</td>
<td>0.66 0.18</td>
<td>0.70 0.16</td>
<td>0.68 0.16</td>
<td>0.70 0.16</td>
</tr>
<tr>
<td>4-factor TP</td>
<td>0.29 2.99</td>
<td>0.42 0.88</td>
<td>0.57 0.20</td>
<td>0.59 0.18</td>
</tr>
<tr>
<td>MVPT_{9%}</td>
<td>0.70 0.16</td>
<td>0.74 0.15</td>
<td>0.79 0.14</td>
<td>0.80 0.13</td>
</tr>
<tr>
<td>Factor models, (\lambda_{CRSP})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-factor TP</td>
<td>0.61 0.07</td>
<td>0.63 0.07</td>
<td>0.64 0.07</td>
<td>0.65 0.07</td>
</tr>
<tr>
<td>MVPT_{9%}</td>
<td>0.63 0.18</td>
<td>0.66 0.18</td>
<td>0.68 0.17</td>
<td>0.69 0.17</td>
</tr>
<tr>
<td>3-factor TP</td>
<td>0.47 0.55</td>
<td>0.53 0.37</td>
<td>0.60 0.14</td>
<td>0.61 0.13</td>
</tr>
<tr>
<td>MVPT_{9%}</td>
<td>0.65 0.15</td>
<td>0.69 0.15</td>
<td>0.72 0.15</td>
<td>0.73 0.15</td>
</tr>
<tr>
<td>4-factor TP</td>
<td>0.44 0.67</td>
<td>0.50 0.33</td>
<td>0.58 0.18</td>
<td>0.60 0.17</td>
</tr>
<tr>
<td>MVPT_{9%}</td>
<td>0.68 0.14</td>
<td>0.73 0.14</td>
<td>0.78 0.13</td>
<td>0.79 0.13</td>
</tr>
</tbody>
</table>

This table depicts annualized Sharpe ratio (SR) and turnover (\(\tau\)) of ten strategies over four data sets. The numbers in the first row stand for the number of stock in the sample and the \(s\) stands for stocks.

using 18,356 stocks (\(\lambda_{CRSP}\)).

Subsamples analysis

In this section we first study the performance of the 2-pass approach as we increase the sample size, then we analyzes the performance on subsample selected as a function of trading volume. Next, we construct a portfolio containing stocks without missing values. This last dataset allows us to compare 1-pass and 2-pass strategies to alternative strategies such as Ledoit-Wolf portfolio.

First, we randomly select stocks from the CRSP database and repeat this procedure 1,000 times. We choose sample size of 100, 200, 1000 and 2000 stocks. We then compute the average estimates for each measure. Results are displayed in Table 1.4.\(^{13}\)

Sharpe ratios and turnovers are driven by extreme values for samples with a low

\(^{13}\)For clarity, we only display Sharpe ratio and turnover. Other performance measures are available on request.
number of stocks. In particular, tangency portfolios perform poorly and exhibit unrealistic turnovers for portfolios including 100 stocks. It gets better as sample size increases. Sharpe ratios are similar to the benchmark and turnovers, although higher, are reasonable. A comparison with the results obtained using risk premia from the whole CRSP database shows that the later tends to be more stable. Sharpe ratios are significantly higher and turnover much lower for the 100 stocks and 200 stocks portfolios. It is not surprising to find similar results between the two methodology for larger portfolios as the information they englobe is representative of the economy.

The next subsample is based on the average monthly volume traded for each stock. The volume is an impartial criterion to select stocks. Trading volume is not necessarily an appropriate measure of liquidity (see Sarr and Lybek (2002)), but in this section we simply use it as a stock characteristic to generate subsamples. We heuristically select three thresholding values, respectively 5 million, 10 million and 100 million of average volume over the period. The first data set is constituted of 797 stocks, the second of 489 stocks, and the last of 39 stocks. Results are depicted in Table 1.5.

Sharpe ratios are equal or higher in all cases but one once we replace risk premia in the 2-pass setup. The impact is minimal on the first and largest sample. The impact, however, important as the sample size decreases. The increase in Sharpe ratio is between 10% and 15% for the 3- and 4-factor portfolios based on a minimum volume of 10 million. The picture is even better with the smallest sample under scrutiny. Portfolio behavior goes from erratic to stable and Sharpe ratios can be up to 8 times higher (1-factor MVPT). The turnover stays low and is similar with both methods. Certainty equivalents increase in all cases.

The MVPT strategy performs better than TP in general. Results are stable in all cases once we replace risk premia, while the erratic behavior of the TP portfolio in small sample is only dampened. MVPT 1-factor strategy exhibits a turnover of 13.2 before changing risk premia and 0.11 after, which is much more reasonable. The large turnover comes from extreme weights on some stocks in the portfolio. HI goes up to 2758 for TP 4-factor model while the best performing portfolio has a HI of 0.16. We believe that portfolio performance could be strongly enhanced by adding some constraint on portfolio weights, especially for small portfolios. As it is outside of the scope of this paper, we leave it to future research.

We can conclude from this example that replacing risk premia by the ones obtained from a much larger sample enhances the overall performance and stabilize the problem. Sharpe ratios and certainty equivalents increase while the turnover stays low. Hence, we recommend using all the available information even though one cannot invest in the whole investment universe.
Table 1.5: Stock selected by minimum average volume

<table>
<thead>
<tr>
<th>Min. volume: 5mios</th>
<th>1-pass</th>
<th>2-pass</th>
<th>1-pass</th>
<th>2-pass</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR</td>
<td>$SR_{\lambda, vol}$</td>
<td>$SR_{\lambda, CRSP}$</td>
<td>$\tau$</td>
</tr>
<tr>
<td>1-factor TP</td>
<td>0.44</td>
<td>0.55</td>
<td>0.55</td>
<td>0.40</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.56</td>
<td>0.62</td>
<td>0.67</td>
<td>0.10</td>
</tr>
<tr>
<td>3-factor TP</td>
<td>0.39</td>
<td>0.55</td>
<td>0.60</td>
<td>0.41</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.57</td>
<td>0.58</td>
<td>0.65</td>
<td>0.13</td>
</tr>
<tr>
<td>4-factor TP</td>
<td>0.39</td>
<td>0.57</td>
<td>0.59</td>
<td>0.46</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.56</td>
<td>0.61</td>
<td>0.64</td>
<td>0.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Min. volume: 10mios</th>
<th>1-pass</th>
<th>2-pass</th>
<th>1-pass</th>
<th>2-pass</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR</td>
<td>$SR_{\lambda, vol}$</td>
<td>$SR_{\lambda, CRSP}$</td>
<td>$\tau$</td>
</tr>
<tr>
<td>1-factor TP</td>
<td>0.17</td>
<td>0.52</td>
<td>0.52</td>
<td>0.63</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.41</td>
<td>0.57</td>
<td>0.60</td>
<td>0.11</td>
</tr>
<tr>
<td>3-factor TP</td>
<td>0.13</td>
<td>0.48</td>
<td>0.56</td>
<td>0.60</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.42</td>
<td>0.49</td>
<td>0.57</td>
<td>0.16</td>
</tr>
<tr>
<td>4-factor TP</td>
<td>0.14</td>
<td>0.50</td>
<td>0.55</td>
<td>0.71</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.42</td>
<td>0.49</td>
<td>0.54</td>
<td>0.16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Min. volume: 100mios</th>
<th>1-pass</th>
<th>2-pass</th>
<th>1-pass</th>
<th>2-pass</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR</td>
<td>$SR_{\lambda, vol}$</td>
<td>$SR_{\lambda, CRSP}$</td>
<td>$\tau$</td>
</tr>
<tr>
<td>1-factor TP</td>
<td>0.08</td>
<td>0.41</td>
<td>0.41</td>
<td>28.5</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.36</td>
<td>0.06</td>
<td>0.49</td>
<td>0.14</td>
</tr>
<tr>
<td>3-factor TP</td>
<td>0.36</td>
<td>-0.31*</td>
<td>0.27</td>
<td>0.82</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.48</td>
<td>0.33</td>
<td>0.62</td>
<td>0.17</td>
</tr>
<tr>
<td>4-factor TP</td>
<td>0.33</td>
<td>-0.05</td>
<td>-0.08</td>
<td>0.91</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.46</td>
<td>0.27</td>
<td>0.57</td>
<td>0.18</td>
</tr>
</tbody>
</table>

This table represents the gain of using the whole CRSP database to infer risk premia over the standard 1-pass model and a 2-pass model (based on subsample only or on whole sample to compute risk premia). $SR$ stands for Sharpe ratio. $\tau$ is the turnover. All numbers are annualized. The target expected return for MVPT portfolio is 9%. * represents statistical difference at 10% with 1-pass methodology.
Table 1.6 shows the results for a subsample of stocks from the CRSP database, which includes only stocks with a complete return history from 1988 to 2012. The subsample includes 754 stocks and the period spans 25 years.\textsuperscript{14} This sample is clearly not exempt of survivorship bias, but it allows us to implement strategies that require balanced panel data.

Sample estimation gives poor results. It is not surprising as the time series dimension is smaller than the number of stocks leading to a singular covariance matrix. Shrinkage methods\textsuperscript{15} are able to mitigate the problem only for strategies that do not involve an estimate of the average excess returns. Hence, the Ledoit-Wolf algorithm performs best while the Bayes-Stein portfolio cannot produce decent estimates. The LW strategy requires however a turnover five time larger than the benchmark for a Sharpe ratio twice lower. The required shrinkage parameter $\alpha$ has to be large to stabilize the covariance matrix. Hence, most of the information available from the returns is lost. The only strategy that performs better than the benchmark is the Volatility Timing strategy.

It is surprising to note that none of our algorithms is able to outperform the benchmark. This results points out the need of a very large number of stocks to profit from diversification. The 2-pass tangency and $MVPT$ strategies are however better than the ones available in the literature (except for volatility timing) even though those two portfolios involve the estimation of the average excess returns.

Table 1.7 compares the out-of-sample results between two different approaches of the problem. In the first case, we compute risk premia using the subsample in order to compute the expected excess returns while in the second case, we replace those risk premia by the one computed from the whole CRSP database.

Sharpe ratios are not significantly different for both portfolios and all factor models. This observation is similar for turnover and certainty equivalent. Those portfolios do not benefit from the additional information. The explanation lies in the construction of the estimator $\lambda$. The second pass of the model employs a WLS approach that gives more weight to stocks with the longest history. Hence, the risk premia computed from the whole CRSP database are very similar to the ones computed with this subsample.

\textsuperscript{14}We restrict the length of the dataset to 25 years in order to have a large number of stocks. Such restriction on the period 1967-2012 leads to a subsample of 120 stocks, results are similar and available upon request.

\textsuperscript{15}Minimum variance constrained portfolio optimization generates corner solution due to the limited length of the sample (120) relative to the number of stocks (754). We choose not to display these unrepresentative results.
This table represents the results obtained from 754 stocks. They do not have any missing value over the period 1988-2012. The first column gives the average excess return of the strategy over the holding period (180 months). \( \sigma \) is the standard deviation of each portfolio. \( SR \) stands for Sharpe ratio and \( SR_{at} \) is the Sharpe ratio after transaction costs. \( \tau \) is the turnover required by a given strategy. \( CE \) represents Certainty Equivalent and \( ETC \) is the maximum transaction costs sustainable to get the same Sharpe ratio as the benchmark strategy \( 1/N \). * represents statistical difference to \( 1/N \), † to 1-pass, and ** to both. The statistical level is 10%. All numbers are annualized.
Table 1.7: Comparison between restricted and full dataset

<table>
<thead>
<tr>
<th>Factor models, 2 pass</th>
<th>$SR_{\lambda,754}$</th>
<th>$SR_{\lambda,CRSP}$</th>
<th>$\tau_{\lambda,754}$</th>
<th>$\tau_{\lambda,CRSP}$</th>
<th>$CE^5_{\lambda,754}$</th>
<th>$CE^5_{\lambda,CRSP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-factor</td>
<td>TP</td>
<td>0.51</td>
<td>0.07</td>
<td>0.07</td>
<td>0.013</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>MVPT</td>
<td>0.57</td>
<td>0.11</td>
<td>0.13</td>
<td>0.024</td>
<td>0.008</td>
</tr>
<tr>
<td>3-factor</td>
<td>TP</td>
<td>0.47</td>
<td>0.55</td>
<td>0.56</td>
<td>-0.059</td>
<td>-0.114</td>
</tr>
<tr>
<td></td>
<td>MVPT</td>
<td>0.52</td>
<td>0.23</td>
<td>0.23</td>
<td>0.025</td>
<td>0.021</td>
</tr>
<tr>
<td>4-factor</td>
<td>TP</td>
<td>0.39</td>
<td>0.69</td>
<td>0.65</td>
<td>-0.013</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>MVPT</td>
<td>0.39</td>
<td>0.20</td>
<td>0.22</td>
<td>0.014</td>
<td>0.015</td>
</tr>
</tbody>
</table>

This table represents the gain of using the whole CRSP database to infer risk premia instead of using a subsample of data even if the investment universe is smaller. 18,356 stocks are taken into account to compute them. The investment universe is represented by the 754 stocks dataset, which includes stocks having a full return history from 1988 to 2012. $SR$ stands for Sharpe ratio. $\tau$ is the turnover, and $CE$ represents Certainty Equivalent. All numbers, but $CE$, are annualized.

Portfolios of stocks

Table 1.8 depicts the results for three data sets based on stock characteristics. The first sample is based on 20 size and book-to-market value portfolios plus the US equity market (21F1). The second is based on ten industry sectors plus the US equity market (11Ind), the last is constructed by country indices and the world index (Int). Those samples are similar to the ones used in DeMiguel et al. (2009b). It allows us to compare our methodology to past results as those samples are balanced and of small cross-sectional dimension, ensuring invertible covariance matrices.

We reproduce the results for TP, GMVP, MVc with sample estimates and BS strategies from DeMiguel et al. (2009b). We present them, however, in annualized estimates to be consistent with the rest of our paper.

We see that Sharpe ratios are better than the benchmark for the three data sets when computing the minimum variance and the MVPT portfolios. The tangency portfolio gives poor results, small Sharpe ratio and large turnover. Setting a target return (MVPT) helps to get decent estimates. Sharpe ratios before transaction costs are higher than the benchmark for the three samples. Turnover is however between 15 and 44 times higher.

Results for shrinkage models are mitigated. The Bayes-Stein portfolio cannot stabilize the problem enough to perform well, especially in term of turnover. The LW strategy, on the other hand, performs much better. It is able to outperform naive diversification in

---

16 We thank Pr. Garlappi for providing us with the data.
17 To get the same values, the reader should divide Sharpe ratios by $\sqrt{12}$. 
This table represents replication results of existing methods as well as our factor models across data sets commonly used in past literature. 21F1 represents 20 portfolios on size and book-to-market plus the US equity market. 11Ind is a representation of US stocks aggregated by industry plus the excess market return. The last sample contains information about countries and the world index. The table shows annualized Sharpe ratios. The period spans from July 1963 to November 2004 for 21F1 and 11Ind. It covers 1970 to 2001 for the Int sample. * represents statistical difference to 1/N. We thank Pr. Garlappi for kindly providing us with the data. SR stands for Sharpe ratio, $\tau$ is the turnover.

<table>
<thead>
<tr>
<th></th>
<th>21F1</th>
<th></th>
<th>11Ind</th>
<th></th>
<th>Int</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$SR$</td>
<td>$\tau$</td>
<td>$SR$</td>
<td>$\tau$</td>
<td>$SR$</td>
<td>$\tau$</td>
</tr>
<tr>
<td><strong>Benchmark</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.56</td>
<td>0.02</td>
<td>0.47</td>
<td>0.02</td>
<td>0.44</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>Sample estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.04*</td>
<td>350.6</td>
<td>0.24</td>
<td>330.9</td>
<td>-0.11*</td>
<td>328.8</td>
</tr>
<tr>
<td>GMVP</td>
<td>0.96*</td>
<td>0.73</td>
<td>0.54</td>
<td>0.46</td>
<td>0.52</td>
<td>0.21</td>
</tr>
<tr>
<td>MVPT</td>
<td>0.74</td>
<td>0.88</td>
<td>0.50</td>
<td>0.73</td>
<td>0.52</td>
<td>0.46</td>
</tr>
<tr>
<td><strong>Shrinkage methods</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVc</td>
<td>0.54</td>
<td>0.06</td>
<td>0.49</td>
<td>0.05</td>
<td>0.52</td>
<td>0.06</td>
</tr>
<tr>
<td>BS</td>
<td>0.05*</td>
<td>116.8</td>
<td>0.25</td>
<td>264.1</td>
<td>-0.10*</td>
<td>111.6</td>
</tr>
<tr>
<td>LW</td>
<td>0.66</td>
<td>0.29</td>
<td>0.48</td>
<td>0.02</td>
<td>0.52</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>Alternatives options</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VT</td>
<td>0.61</td>
<td>0.03</td>
<td>0.51</td>
<td>0.04</td>
<td>0.49</td>
<td>0.03</td>
</tr>
<tr>
<td>RRT</td>
<td>0.68*</td>
<td>0.07</td>
<td>0.34</td>
<td>0.11</td>
<td>0.51</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>Factor models, 2 pass</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-factor TP</td>
<td>0.40*</td>
<td>0.01</td>
<td>0.40*</td>
<td>0.01</td>
<td>0.53</td>
<td>0.01</td>
</tr>
<tr>
<td>GMVP</td>
<td>0.28</td>
<td>0.34</td>
<td>0.65</td>
<td>0.16</td>
<td>0.44</td>
<td>0.09</td>
</tr>
<tr>
<td>MVPT</td>
<td>-0.03*</td>
<td>155.4</td>
<td>0.25</td>
<td>59.4</td>
<td>0.43</td>
<td>0.08</td>
</tr>
<tr>
<td>3-factor TP</td>
<td>-0.17*</td>
<td>26.4</td>
<td>-0.30*</td>
<td>67.3</td>
<td>0.17</td>
<td>14.4</td>
</tr>
<tr>
<td>GMVP</td>
<td>0.76</td>
<td>0.44</td>
<td>0.60</td>
<td>0.18</td>
<td>0.44</td>
<td>0.09</td>
</tr>
<tr>
<td>MVPT</td>
<td>0.77</td>
<td>0.63</td>
<td>0.06*</td>
<td>1.24</td>
<td>0.35</td>
<td>0.90</td>
</tr>
<tr>
<td>4-factor TP</td>
<td>0.19*</td>
<td>254.9</td>
<td>-0.16*</td>
<td>74.0</td>
<td>-0.21*</td>
<td>101.5</td>
</tr>
<tr>
<td>GMVP</td>
<td>0.74</td>
<td>0.46</td>
<td>0.57</td>
<td>0.20</td>
<td>0.44</td>
<td>0.09</td>
</tr>
<tr>
<td>MVPT</td>
<td>0.88*</td>
<td>0.73</td>
<td>0.06*</td>
<td>1.14</td>
<td>0.38</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Table 1.8: Replication
term of Sharpe ratio before transaction costs in all cases. Additionally, turnover is low for 11Ind and Int but 15 times greater for the 21F1 sample.

The alternative portfolios developed in Kirby and Ostdiek (2009) are quite successful. The VT strategy exceeds 1/N by almost 10% in all cases and does not require a much higher turnover. RRT, on the other hand, is more sensitive to sample selection. It outperforms the benchmark in only two samples and requires a turnover almost four times larger for all of them.

The 1-factor model gives poor results for the tangency and the MVPT portfolios. The turnover is high and the Sharpe ratio is low in both cases. The global minimum variance strategy has a better behavior and gives decent estimates. The turnover is low, especially for the International data set, which has a similar Sharpe ratio as the benchmark. The Sharpe ratio is higher than 1/N and all other portfolios for the 11Ind sample. It is however lower for the 21F1. The 3-factor model gives better results than the 1-factor, the GMVP strategy outperforms the benchmark on every sample. On the other hand, TP has poor results. The 4-factor GMVP outperforms the benchmark in any case, with a relatively low turnover. TP strategy performs poorly on all samples. MVPT results are mixed, the Sharpe ratio is high for 21F1 but low for 11Ind and average for Int.

Out of those samples, it is very difficult to see a clear benefit from factor models. Our strategies are only able to outperform naive diversification in some cases, as in DeMiguel et al. (2009b). It is due, however, to the intrinsic nature of those samples. Grouping stocks into portfolios is known to reduce the amount of information available in the returns. In addition, our choices of factors are well suited to model stock returns. Gagliardini et al. (forthcoming) show that they are not good in capturing portfolio returns. This explanation is the major drawback of factor models as a way of regularization. One needs to know at least some of the structure of the data ex-ante to get interesting results. We are confident that with suitable factors, one would obtain better and more stable results. Overall, we notice a significant decrease in turnover for GMVP strategies once we use a factor model compared to their sample estimation counterparts, and this holds for all data sets independent of our choice of factors.

1.6.2 Robustness

We implement the 2-pass approach on the full CRSP dataset (18,356 stocks) with different rolling window sizes. In particular, we use a 60 months and a 180 months rolling windows. The results are displayed in Table 1.9 and Table 1.10. For the longest rolling window, all portfolios under consideration are outperforming naive diversification in term of Sharpe ratio net of transaction costs.
Table 1.9: All CRSP with 60 months rolling window

<table>
<thead>
<tr>
<th></th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$SR$</th>
<th>$\tau$</th>
<th>$SR_{at}$</th>
<th>$CE_1$</th>
<th>$CE_5$</th>
<th>ETC</th>
<th>HI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.08</td>
<td>0.18</td>
<td>0.43</td>
<td>0.10</td>
<td>0.40</td>
<td>0.062</td>
<td>-0.004</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td>Factor models, 2-pass</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.10</td>
<td>0.18</td>
<td>0.55*</td>
<td>0.10</td>
<td>0.52*</td>
<td>0.082</td>
<td>0.020</td>
<td>0.022</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.03</td>
<td>0.74</td>
<td>0.04*</td>
<td>8.47</td>
<td>-0.66*</td>
<td>-0.245</td>
<td>-1.329*</td>
<td>-0.003</td>
<td>0.070</td>
</tr>
<tr>
<td>3-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.05</td>
<td>0.75</td>
<td>0.07</td>
<td>5.15</td>
<td>-0.35*</td>
<td>-0.231</td>
<td>-1.353</td>
<td>-0.004</td>
<td>0.443</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.08</td>
<td>0.14</td>
<td>0.58</td>
<td>0.28</td>
<td>0.46</td>
<td>0.071</td>
<td>0.032</td>
<td>0.007</td>
<td>0.002</td>
</tr>
<tr>
<td>4-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.02</td>
<td>0.56</td>
<td>0.04*</td>
<td>6.93</td>
<td>-0.71*</td>
<td>-0.138</td>
<td>-0.77</td>
<td>-0.002</td>
<td>0.315</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.07</td>
<td>0.11</td>
<td>0.66*</td>
<td>0.23</td>
<td>0.54</td>
<td>0.068</td>
<td>0.043*</td>
<td>0.011</td>
<td>0.002</td>
</tr>
</tbody>
</table>

This table represents the results obtained on the whole CRSP database. 18,356 stocks are taken into account. The first column gives the average return of the strategy over the period 1977-2012 (the first 10 years being used for estimation). $\sigma$ is the standard deviation of each portfolio. $SR$ stands for Sharpe ratio and $SR_{at}$ is the Sharpe ratio after transaction costs. $\tau$ is the turnover required by a given strategy. $CE$ represents Certainty Equivalent with a risk aversion parameter of 1 and 5, and $ETC$ is the maximum transaction costs sustainable to get the same Sharpe ratio as the benchmark strategy 1/N. $HI$ stands for Herfindahl index. * represents statistical difference to 1/N. All numbers are annualized.

Table 1.10: All CRSP with 180 months rolling window

<table>
<thead>
<tr>
<th></th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$SR$</th>
<th>$\tau$</th>
<th>$SR_{at}$</th>
<th>$CE_1$</th>
<th>$CE_5$</th>
<th>ETC</th>
<th>HI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.08</td>
<td>0.17</td>
<td>0.45</td>
<td>0.11</td>
<td>0.41</td>
<td>0.062</td>
<td>0.004</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td>Factor models, 2-pass</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.10</td>
<td>0.16</td>
<td>0.63*</td>
<td>0.09</td>
<td>0.60*</td>
<td>0.090</td>
<td>0.036</td>
<td>0.033</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.11</td>
<td>0.15</td>
<td>0.70*</td>
<td>0.11</td>
<td>0.66*</td>
<td>0.094</td>
<td>0.049</td>
<td>0.033</td>
<td>0.001</td>
</tr>
<tr>
<td>3-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.10</td>
<td>0.16</td>
<td>0.62*</td>
<td>0.10</td>
<td>0.58*</td>
<td>0.087</td>
<td>0.035</td>
<td>0.028</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.11</td>
<td>0.15</td>
<td>0.72*</td>
<td>0.11</td>
<td>0.67*</td>
<td>0.094</td>
<td>0.051</td>
<td>0.033</td>
<td>0.001</td>
</tr>
<tr>
<td>4-factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.10</td>
<td>0.16</td>
<td>0.60*</td>
<td>0.12</td>
<td>0.55*</td>
<td>0.082</td>
<td>0.031</td>
<td>0.020</td>
<td>0.001</td>
</tr>
<tr>
<td>MVPT$_{9%}$</td>
<td>0.10</td>
<td>0.13</td>
<td>0.73*</td>
<td>0.10</td>
<td>0.69*</td>
<td>0.089</td>
<td>0.053</td>
<td>0.035</td>
<td>0.001</td>
</tr>
</tbody>
</table>

This table represents the results obtained on the whole CRSP database. 18,356 stocks are taken into account. The first column gives the average return of the strategy over the period 1977-2012 (the first 10 years being used for estimation). $\sigma$ is the standard deviation of each portfolio. $SR$ stands for Sharpe ratio and $SR_{at}$ is the Sharpe ratio after transaction costs. $\tau$ is the turnover required by a given strategy. $CE$ represents Certainty Equivalent with a risk aversion parameter of 1 and 5, and $ETC$ is the maximum transaction costs sustainable to get the same Sharpe ratio as the benchmark strategy 1/N. $HI$ stands for Herfindahl index. * represents statistical difference to 1/N. All numbers are annualized.
Outperforming Naive Diversification Using Stock Level Information

ratio before and net of transaction costs as well as in term of certainty equivalent. Similar to our findings with a 120 months window, the turnover is very low. The conclusions are the same. The difference in performance between the benchmark and factor models is statistically significant with this methodology. The Sharpe ratio net of transaction costs for tangency portfolios range from 0.55 to 0.60 while the benchmark is 0.40. They are between 0.66 and 0.69 for MVPT.

Results are erratic in a setting with a rolling of 60 months. The estimation of the variance-covariance matrix benefits from the factor structure. On the other hand, the estimation of the average excess returns is more problematic as the tangency portfolio and the MVPT exhibit inconsistent results. In particular, some portfolios tends to have extreme weights, which is shown through large $HI$ and turnovers.

Results are in line with the literature. There exists a trade-off between a long horizon that provides good estimates but are biased due to too much past information and a short horizon free of old irrelevant data but noisy estimators.

1.7 Conclusion

In this paper, we show the benefits of optimal portfolio allocation under the Markowitz framework. DeMiguel et al. (2009b) challenge the usefulness of mean-variance optimization in practice. They show that not a single strategy is constantly able to outperform naive diversification over a range of data sets. Their samples are, however, constituted of portfolios of stocks. Ang et al. (2008) point out the severe distortion of information introduced by grouping stocks into portfolios. Hence, in our paper, we compute the performance of portfolios constituted of a very large number of stocks. Our largest sample is the whole CRSP database, which includes over 18,000 stocks and spans a period of 45 years.

We use a factor model to construct the two moments of interest in the Markowitz framework, namely the mean and the covariance matrix. We rely on a recent paper by Gagliardini et al. (forthcoming) to estimate the factor loadings of individual stocks and the risk premia of common factors. The authors propose a convenient and tractable two-pass estimation procedure to compute these estimates. They include a missing-at-random design, which is particularly helpful when it comes to empirical analysis. Their methodology significantly improves the estimation of average excess returns, which in turns helps portfolio optimization.

We find that all portfolios under consideration are able to outperform naive diversification on a sufficiently large data set. Sharpe ratios and certainty equivalents are higher than the benchmark. More importantly, turnover remains low, which has been one of the
major problems for Markowitz portfolios.

On the other hand, we confirm the results obtained by DeMiguel et al. (2009b) once the data set under consideration is composed of portfolios of stocks. None of the standard strategies can constantly outperform 1/N on the three samples build on portfolios of stocks. This is also in line with the findings of Gagliardini et al. (forthcoming), who find that the estimated factor loadings are totally different when one uses portfolios of stocks instead of individual stocks.

In addition, we show that performing a 2-pass regression leads to better results than the simple 1-pass setting. The former uses all available information to compute the common part in the alpha parameter while the latter uses only individual information. The 2-pass regression helps to mitigate the estimation errors by including all stocks.

We focus on standard factor models and illustrate the performance of the methodology. Further analysis on the optimal factor structure is left for future research.
Chapter 2

Pairwise Correlation Dynamics and Incomplete Information

Tony Berrada\textsuperscript{1}, Sébastien Coupé, & Thuy-Duong Tô\textsuperscript{2}

\textbf{Abstract.} We consider an economy with a continuum of firms paying out cash-flows which are affected by an aggregate shock and a firm specific shock. The representative agent has incomplete information and must learn about the aggregate cash flow process, which follows a continuous time Markov chain. We study the dynamics of pairwise correlations and show that (i) pairwise correlations are U-shaped as a function of the probability of the good state (ii) correlations are stronger for higher beta stocks, and increase more during bad time for lower beta stocks. In addition, the model calibrated to the US business cycle matches the data in terms of volatility, correlation, and risk premium. We further study the implications of our model for portfolio allocation.

\textbf{JEL Classification:} G12

\textbf{Keywords:} Incomplete information, pairwise correlation, volatility clustering

2.1 Introduction

When it rains, it pours. Correlation is the key to portfolio diversification and therefore portfolio risk management. However, as the market goes into recession, when the need

\textsuperscript{1}University of Geneva and Swiss Finance Institute, Unimail, Boulevard du Pont d’Arve 40, 1211 Geneva–4, Switzerland. tony.berrada@unige.ch, +41(0)22 379 81 26.

\textsuperscript{2}School of Banking and Finance, UNSW Business School, The University of New South Wales, Sydney, NSW 2052, Australia. Email: td.to@unsw.edu.au
to control risk is presumably the most, correlation goes up resulting in the least effect of portfolio risk reduction technique. This observation is true for correlation between individual stocks (Chordia et al. 2011), between individual stocks and the market (Ang and Chen 2002, Hong et al. 2007), and between international stocks (Erb et al. 1994, Longin and Solnik 2001).

This paper aims to provide a rational explanation to the phenomenon. We want to pinpoint the underlying structure that creates the important but intriguing patterns of volatility and correlation, namely time varying, clustering, and asymmetric to market conditions. We want to know whether correlation risk is priced. The answer will help the quest of understanding equity risk premia. Finally, we want to build efficient portfolios for investors.

We propose a rational expectation intertemporal equilibrium model of asset prices with aggregate and idiosyncratic risks, where information about the aggregate component is incomplete. Learning about the unobservable component creates business cycles that are conditioned by the time varying probability of a given state. Shocks to the idiosyncratic component affect stock characteristics conditional on the aggregate state of the economy, creating a dynamic pattern of cross sectional heterogeneity. In this model, a conditional version of the CAPM holds and only aggregate shocks carry a risk premium. However, idiosyncratic volatility is affected by the aggregate state of the economy, and therefore the cross section of correlations varies over time.

Our model generates results that are in line with empirical findings. On the time series dimension, both volatility of and correlations between stocks are higher during extreme market conditions\(^3\). On the cross sectional dimension, cross correlations within low beta stocks are lower than those within high beta stocks\(^4\). Our efficient frontier, built for each investment horizon by proper conditioning, reduces significantly the risk of the global minimum variance portfolio.

Our research relates to literature on equity correlations. For empirical research, Andersen, Bollerslev, Diebold, and Ebens (2001) find strong evidence of positive relations between individual stocks volatilities, and between contemporaneous stock correlations, suggesting a latent factor model. Our model implies the probability of a given state of the economy as the latent factor. Ang and Bekaert (2002) proposes a regime switching process for the international equity market that can produce higher correlations during bad time. The model assumes that the agent knows the regimes at each time, whereas

\(^3\)Extensive research has shown the asymmetric pattern of volatility that arises from bad economic condition (see French et al. (1987), Schwert (1989), Campbell and Hentschel (1992), Bekaert and Wu (2000)

\(^4\)See Ang and Chen (2002) and Chordia et al. (2011) for empirical results of the difference in correlations between low and high beta stocks
in our model, agent needs to learn about the state of the economy. In a general equilibrium setting, Ribeiro and Veronesi (2002) also study an international market with shifts in regime and learning. Our model differs by introducing idiosyncratic shocks. In our model equilibrium systematic and idiosyncratic volatilities at the stock level are both affected by the global market condition modeled by a Markov chain. Correlations change across regime because systematic and idiosyncratic volatilities are not symmetrically affected by the global market conditions. This effect is new and specific to our model. Our model also differs significantly from the setting in Santos and Veronesi (2010), which has a finite number of stocks and each firm’s idiosyncratic risk affects aggregate consumption and is therefore priced in equilibrium. The mechanism we identify which drives the evolution of correlations does not rely on under diversification of idiosyncratic risk.

Our research adds to the extensive research on general equilibrium models of the economy, which cover a range of important issues. Papanikolau (2011) models investment shocks and show their impact on business cycle. Other notable papers are Gomes, Kogan, and Zhang (2003) who build a general equilibrium production economy that links expected stock returns to firm characteristics, and Ai, Croce, and Li (2013) who analyze the value premium in a model where some technology shocks the productivity of old capital. Kogan, Papanikolaou, and Stoffman (2015) construct a general equilibrium model in which benefits of technological innovation are distributed asymmetrically which gives rise to high risk premium on aggregate stock market and return co-movement. Models in these papers have rich implications but do not allow for closed form solution and cross-correlation is not studied. Our study focuses on a simple cash flow risk model and is highly tractable. In addition, our model allows for relatively easy simulations, which in turns ease portfolio allocation.

The paper is organized as follows. In Section 2.2, we introduce the model and explain how aggregate and idiosyncratic shocks affect firms’ cash flows. Section 2.3 presents the equilibrium results. Section 2.4 discusses the calibration procedure and the theoretical implications of the model. Section 2.5 shows evidence from market data supporting the model. Section 2.6 analyzes the implication for portfolio optimization. Section 1.7 concludes.

### 2.2 The model

We consider an economy with a continuum of firms indexed \( i \in \mathcal{I} \). Time is continuous and the horizon is infinite. Sections 2.2.1 and 2.2.2 presents the uncertainty structure and firm specifications, respectively. The representative agent problem and the information structure are described in Section 2.2.3. Section 2.2.4 specifies financial markets.
2.2.1 Uncertainty

Uncertainty is modeled by a triplet \((B^a, B^i, \varepsilon)\) where \(B^a\) is a one-dimensional brownian motion, \(B^i\) is an infinite dimensional brownian motion vector and \(\varepsilon\) is a continuous time Markov chain. The brownian motion \(B^a\) represents aggregate uncertainty in the economy. \(B^i\) represents firm specific risk. The Markov chain \(\varepsilon_t \in \{\varepsilon_1 = \varepsilon_L, \varepsilon_2 = \varepsilon_H\}\) drives the evolution between different aggregate growth regimes. \(^5\) The matrix \(\Lambda = [\lambda_{ij}]\) gives the switching intensity, where \(\lambda_{ij}\) is the instantaneous rate of switching from state \(\varepsilon_i\) to \(\varepsilon_j\). The rates of staying in a given state, also called inertia rates, are diagonal elements and negative. They satisfy \(\lambda_{ii} = -\lambda_{ij}\).

2.2.2 Firms

Ownership of a firm yields payment of a continuous cash flow process denoted \(\text{CF}_i\) which is the product of two terms

\[
\text{CF}_{it} = X_t Y_{it}.
\]

The first term \(X_t\) denotes the aggregate cash flow in the economy and it evolves according to

\[
dX_t = \theta(\varepsilon) X_t dt + \delta^a X_t dB^a_t
\]

for some positive constant \(\delta^a\). \(\theta(\varepsilon)\) is the growth rate that depends on the economic condition modeled by \(\varepsilon\). We assume that \(\varepsilon_H\) is the good state of the economy and \(\theta(\varepsilon_H) > \theta(\varepsilon_L)\). The second term \(Y_{it}\) is a firm specific component that follows a mean-reverting process

\[
dY_{it} = \alpha(\eta - Y_{it}) dt + \delta^i dB^i_{it}
\]

\[Y_{i0} \sim \mathcal{N}\left(\eta, \frac{(\delta^i)^2}{2\alpha}\right)\]

\(\delta^i\) is the firm specific cash flow volatility, \(\alpha\) is the speed of mean reversion, and \(\eta\) the long term mean. The dynamic of \(Y_{it}\) is defined as a stationary process for tractability.

\(^5\) An extension of the usual product measure called the Fubini extension (see Sun (2006)) is rich enough to satisfy all our requirements and definitions. It also guarantees that we can apply a version of the strong law of large number when aggregating across firms.
Aggregating over the set of firms yields the aggregate cash flow in the economy

\[ \int I CF_{it} di = \int I X_i Y_{it} di = \eta X_t. \]

It is important to note that the aggregate economy output does not depend on firm specific characteristics. In particular it is independent of idiosyncratic shocks \( B_i \). We assume that cash flows are immediately perishable and non-storable, it follows that in equilibrium aggregate consumption equals aggregate cash flow.

### 2.2.3 Representative agent

The representative agent has constant relative risk aversion (CRRA) preferences. The relative risk aversion coefficient and the discount rate are positive constants denoted \( \gamma \) and \( \rho \) respectively. The information available to the representative agent is summarized by the augmentation of the natural filtration of the continuum of processes \( CF_{it} \), which we denote \( F_t \). It follows, that by aggregating firm specific cash flows, the aggregate output \( X_t \) is also observable. By observing the continuous evolution of these processes, the representative agent also observes their quadratic variations and co-variations, and therefore \( \delta^a, \delta^i \) are also observable. The representative agent maximizes her expected utility

\[ U(c) = E \left[ \int_0^\infty e^{-\rho s} u(c_s) ds \right] \]

where \( c_s \) is consumption at time \( s \), and

\[ u(c_s) = \frac{c_s^{1-\gamma}}{1-\gamma}. \]

Information is incomplete in this model, as \( \theta(\varepsilon) \) is not observable. The representative agent forms beliefs about the state of the economy \( \varepsilon \) based on initial priors and the observations. We denote \( p_t \) the conditional probability at time \( t \) of being in the high state \( H \). Applying standard results in filtering theory, the conditional probability of \( \varepsilon = \varepsilon_H \) evolves according to

\[ dp_t = (-p_t \lambda_{12} + (1 - p_t)\lambda_{21}) dt + p_t (1 - p_t) \left( \frac{\theta(\varepsilon_H) - \theta(\varepsilon_L)}{\delta^a} \right) d\nu^a_t \]

\[ = p_t^\rho dt + \sigma_t^p d\nu^a_t \]
where the innovation process $\nu_t^a$ is given by

$$\nu_t^a = B_t^a + \int_0^t \frac{\theta(\varepsilon_s) - (p_s \theta(\varepsilon_H) + (1 - p_s) \theta(\varepsilon_L))}{\delta^a} ds.$$ 

The average growth rate is given by $m_t = (p_t \theta(\varepsilon_H) + (1 - p_t) \theta(\varepsilon_L))$. Initial prior about the state of the economy is exogenously given by $p_0$.

We can write the evolution of the aggregate output in the representative agent’s filtration as follows

$$dX_t = m_t X_t dt + \delta^a X_t d\nu_t^a.$$

David (1997) and Veronesi (2000) introduced models of financial markets with this type of uncertainty structure. In our setting, however, the agent does not observe the aggregate cash flow directly but a continuum of cash flow processes. Fortunately this is equivalent to seeing the aggregate signal and no additional information is inferred from the idiosyncratic component. This result is proved in Pastor and Veronesi (2012) and the technicalities of filtering theory covered in Lipser and Shirayev (2001) apply.

### 2.2.4 Financial markets

Firms are financed through equity only. Since cash flows are perishable and non-storable, and we do not model the firm’s investment decision, all cash flows are paid out to the stock holders. Stocks are traded on competitive financial markets and stock prices evolution in equilibrium are given by the following equation

$$\frac{dS_{it} + CF_{it} dt}{S_{it}} = \mu_{it} dt + \sigma^a_{i,t} d\nu_t^a + \sigma^s_{i,t} dB^s_{i,t}$$

where $\mu_{it}$ is the expected return, $\sigma^a_{i,t}$ the systematic, or aggregate, volatility of the return and $\sigma^s_{i,t}$ the idiosyncratic volatility of the return. These three quantities are endogenous and are obtained as part of the competitive equilibrium.

### 2.3 Equilibrium

In this section we characterize the competitive equilibrium of the economy described in section 2.2. The state price density is given in Section 2.3.1, the stock price in section 2.3.2 and the volatility in section 2.3.3. Section 2.3.5 considers the cross-correlation of stocks.
2.3.1 State price density

A state price density $\xi_t$ and a consumption process $c_t$ are an equilibrium if all markets clear and the representative agent consumption is optimal. In equilibrium the representative agent must hold the entire market and consume the aggregate cash flow. It follows that the state price density is proportional to the representative agent’s marginal utility evaluated at the aggregate cash flow level. Standard arguments and simplifications yield the following equation for the state price density process

$$\xi_t = e^{-\rho t} \frac{X_t^{-\gamma}}{X_0^{\gamma}}$$

with dynamic identifying the equilibrium risk free rate $r_t$ and market price of risk $\kappa_t$

$$d\xi_t = \xi_t (-r_t dt - \kappa_t d\nu^a)$$

where

$$r_t = \rho + \gamma m_t - 0.5\gamma (\gamma + 1)(\delta^a)^2$$

$$\kappa_t = \delta^a \gamma.$$ 

The market price of risk is the product of the relative risk aversion coefficient and the volatility of aggregate cash flow, it is one dimensional as the idiosyncratic risks are perfectly diversified in the aggregate market portfolio held by the representative agent. It follows that the representative agent ’s marginal utility is unaffected by the idiosyncratic brownian motion which therefore have no risk premium.

The assumption of CRRA preferences implies that long run risk modeled through the markov chain affecting the growth rate of aggregate consumption is not directly priced. This would be the case with a recursive preference specification. We choose to focus on this simpler model to analyze the role of learning at the aggregate level on the dynamics and cross section of correlations.

Our model also differs significantly from settings with a finite number of stocks where each firm’s idiosyncratic risk affects aggregate consumption and is therefore priced in equilibrium (see for example Santos and Veronesi (2010)). The mechanism we identify which drives the evolution of correlations does not rely on under diversification of idiosyncratic risk.
2.3.2 Stock price

Under some regularity conditions that avoid the emergence of bubbles in asset prices, the stock price of firm $i$ is given by the discounted expected sum of future dividends

$$S_{it} = E \left[ \int_t^\infty \frac{\xi_s}{\xi_t} X_s Y_{is} ds \right] \mid \mathcal{F}_t$$

Our model takes into account several key aspects of the economy. First, it incorporates the expected impact on prices of future news through the learning process. Second, the model is rich enough to generate correlations across assets, allowing us to study the behavior at the cross-section. On the other hand, the model remains simple enough to be tractable and to produce closed-form solutions. It allows studying the mechanism carefully with comparative statics analysis.

**Proposition 1** The stock price of firm $i$ at time $t$ is given by

$$S_{it} = f(X_t, Y_{it}, p_t) = X_t \eta (A + B p_t) + X_t (Y_{it} - \eta) (a + b p_t)$$

where the constants $A, B, a$ and $b$ are defined in the appendix.

The stock price is composed of two distinct parts. The first part is the price that would prevail without heterogeneity across firm cash flows. It is a function of the systematic part of the cash flow, $X_t$, and the business cycle conditions summarized by $p_t$. Heterogeneity affects stock prices along two dimensions that appear in the second term. First, as $Y_{it}$ differs from its average $\eta$, the stock price is decreasing or increasing relative to the average. Second, heterogeneity is also varying across time through its dependence on $p_t$. In other words, heterogeneity differs from good times to bad times. This dual dependence is the key to modeling the dynamics of the cross section of correlations.

2.3.3 Stock returns and volatility

We can use the closed form expression of proposition 1 to derive the expected instantaneous return and volatility of individual stocks.

**Proposition 2** The equilibrium stock return is given by

$$\mu_{it} = (r_t + \sigma_{it}^\alpha \gamma^\alpha)$$
and the systematic and idiosyncratic volatilities by

\[
\sigma^a_{i,t} = \delta^a + \frac{(Y_{it} - \eta)b + \eta B}{(Y_{it} - \eta)[a + bp_t] + \eta[A + Bp_t]}\sigma^p_t \tag{2.1}
\]

\[
\sigma^i_{i,t} = \left[ Y_{it} + \eta (A + Bp_t) - (a + bp_t) \right]^{-1} \delta^i \tag{2.2}
\]

The stock volatility, depends on firms specific components represented by \( Y_{it} \) and on the business cycle condition defined by \( p_t \).

Systematic volatility, given in equation 2.1, is the sum of two terms. The first term is due to the cash flow systematic volatility \( \delta^a \) while the second term depends on the business cycle condition volatility \( \sigma^p_t \). The sensitivity of the stock return to business cycle conditions is affected by the idiosyncratic cash flow component, \( Y_{it} \), and it follows that systematic volatility displays heterogeneity across firms.

Similarly, idiosyncratic stock volatility, given in equation 2.2, is affected by business cycle conditions and the idiosyncratic cash flow component. The dependence are however functionally different. Idiosyncratic and systematic volatility react differently to changes in idiosyncratic shocks and changes in business cycle conditions.

It is important to remark that a conditional version of the CAPM holds in our model. Indeed, consider a claim on the aggregate cash flow \( \int_i CF_{it}di = \eta X_t \), which we call the market \( M_t \), we have

\[
\frac{dM_t + \eta X_t dt}{M_t} - r_t dt = \sigma_{mt} \gamma \delta^a dt + \sigma_{mt} d\nu^a.
\]

where

\[
\sigma_{mt} = \delta^a + \frac{\sigma^p_t B}{A + Bp_t}
\]

A proof of this result is given in the appendix.

We can then write the stock dynamics as a function of the market dynamics

\[
\frac{dS_{it} + CF_{it}dt}{S_{it}} - r_t dt = \frac{\sigma^a_{mt}}{\sigma^2_{mt}} \left( \frac{dM_t + \eta X_t dt}{M_t} - r_t dt \right) + \sigma^i_{it}(p_t, Y_{it})dB^i
\]

\[
= \beta_{it} \left( \frac{dM_t + \eta X_t dt}{M_t} - r_t dt \right) + \sigma^i_{it}(p_t, Y_{it})dB^i
\]

In other terms, we have the excess return of firm \( i \) on the left hand side and the market excess return multiplied by the instantaneous market beta plus an error term on the right hand side.
2.3.4 Market beta

Our framework has direct implication for the ongoing debate about the nature of the underlying mechanism affecting stock returns in the literature (see Daniel and Titman (1997), Daniel and Titman (2012)).

Market betas are given by the following formula

\[ \beta_{it} = \frac{\sigma_{it}^a}{\sigma_{mt}} \]

\( \sigma_{it}^a \) being a function of \( Y_{it} \) it implies that a firm’s market beta is a function of the idiosyncratic shock. Hence, firm characteristics act as proxies for the riskiness of a firm.

An extension of the model presented in this paper with multiple sources of idiosyncratic risk is possible and a closed form solution can be derived easily (e.g. by writing the cash flow process as being \( X_t Y_{i,t}^1 \cdots Y_{i,t}^n \)). The interpretation of the results, however, is complex. Indeed, each idiosyncratic component will affect one another due to the multiplicative nature of the process. We leave this issue for future research. Figure 2.1.a shows the effect of \( Y_{i,t} \) on market betas.

2.3.5 Cross-sectional correlation

We now address the central question of the paper, which relates to the dynamics of the cross-section of correlations. There is ample empirical evidence that correlations tend to increase in bad times (e.g. Chordia et al. (2011), Ang and Bekaert (2002)), therefore significantly affecting the gains from diversification in portfolio allocation. Understanding correlation dynamics is also important for risk management and asset prices.

In this section we study the effect of idiosyncratic risk and learning on the cross section of correlations. It follows from our previous results on stock price and volatility, that the key components of the correlation between stock \((i,j)\) are the idiosyncratic shocks \(Y_{it}\) and \(Y_{jt}\) and the conditional probability \(p_t\). Before formally deriving these implications we provide some intuition. Given that the conditional CAPM holds in our model it follows that

\[ R_{it}^e = \beta_{it} R_{mt}^e + \varepsilon_{it} \]

where \( R_{it}^e \) denotes excess return. In our model the standard deviation of the distribution of \( \varepsilon_{it} \) is affected by the conditional probability \( p_t \) (as shown in Equation 2.2). Note that \( \varepsilon_{it} \) is not priced in equilibrium, as aggregating over stocks provides a perfect hedge against idiosyncratic risk. Nevertheless, some characteristics of idiosyncratic volatility depend on aggregate variables. This means that correlations, which depends on the relative fraction
of risk explained by common factors, will vary in the cross section (for different pairs \((\varepsilon_i, \varepsilon_j)\)) but also through time for different \(p_t\). This is the mechanism that drives the results in our model. The next corollary details the dependence of the correlation on \((Y_{it} \text{ and } Y_{jt})\) and the conditional probability \(p_t\).

**Corollary 3** The instantaneous cross-sectional correlation between two stocks is defined by

\[
\rho(Y_{it}, Y_{jt}, p_t) = \frac{\sigma_{i,t}^a \sigma_{j,t}^a}{\sqrt{(\sigma_{i,t}^a)^2 + (\sigma_{i,t}^l)^2} \sqrt{(\sigma_{j,t}^a)^2 + (\sigma_{j,t}^l)^2}}
\]

where the volatilities are as defined in Equation 2.1 and Equation 2.2.

### 2.4 Theoretical implications of the model

#### 2.4.1 Calibration

We illustrate the quantitative implications of the theory by performing a calibration experiment. Tables 2.1 summaries all parameter values of our model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount rate</td>
<td>(\rho)</td>
<td>0.05</td>
</tr>
<tr>
<td>Risk aversion</td>
<td>(\gamma)</td>
<td>3</td>
</tr>
<tr>
<td>High state</td>
<td>(\theta_H)</td>
<td>0.08</td>
</tr>
<tr>
<td>Low state</td>
<td>(\theta_L)</td>
<td>-0.02</td>
</tr>
<tr>
<td>Intensity high to low</td>
<td>(\lambda_{12})</td>
<td>0.23</td>
</tr>
<tr>
<td>Intensity low to high</td>
<td>(\lambda_{21})</td>
<td>1.15</td>
</tr>
<tr>
<td>Aggregate Cash Flow</td>
<td>(X)</td>
<td>100</td>
</tr>
<tr>
<td>Aggregate CF volatility</td>
<td>(\delta^A)</td>
<td>0.1</td>
</tr>
<tr>
<td>Speed of mean reversion</td>
<td>(\alpha)</td>
<td>0.5</td>
</tr>
<tr>
<td>Long term mean</td>
<td>(\beta)</td>
<td>5</td>
</tr>
<tr>
<td>Idiosyncratic volatility</td>
<td>(\delta^i)</td>
<td>0.8</td>
</tr>
</tbody>
</table>

**Table 2.1:** Parameter values used in all computation in sections 2.4, 2.5 and 2.6.

We consider two states in this model, the economy can be is recession or in expansion. The transition matrix coefficients representing the economy jumping from one state to the other are assessed from NBER post-war data. The average length of a recession is
ten months and the average length of an expansion is close to fifty months, resulting in estimates for $\lambda_{12}$ and $\lambda_{21}$ of 0.23 and 1.15 respectively. The stationary probability of the high state is given by $\bar{p} = \frac{\lambda_{21}}{\lambda_{12} + \lambda_{21}} = 0.83$.

The dynamic of the aggregate cash flow is characterized by a (unknown) drift and volatility. We set the drift in the high state and in the low state of the economy to 8% and −2% respectively, which produce an average in line with observed dividend yield. An increase in the difference between those two measures implies a larger volatility for the filter. The aggregate shock level ($X_t$) is equal to 100. It is worth noting that the only measure affected by this choice is the stock price. Betas, price-earning ratios, volatility coefficients and cross-correlations are independent of the level of $X_t$. We discuss the choice of volatility level below since it is linked to the choice of the risk aversion parameter.

The dynamic of the idiosyncratic cash flow is a mean-reverting process. Hence, the drift term is characterized by two parameters. The first is the long term average $\eta$ and is set to 5. The second is the speed of mean reversion ($\alpha$) and is set to 0.5. We study the impact of different values of $Y_t$ around $\eta$.

The volatility of the idiosyncratic shock $\delta^t$ comes from the estimation of Diavatopoulos and Peterson (2008) and is equal to 80%.

The representative agent in this economy has CRRA type preferences. CRRA type preferences in this framework are only defined by the consumption level at a given time and a constant risk aversion coefficient. A direct implication of our model is the equity risk premium, which is similar to Veronesi (2000). It is not impacted by the addition of the idiosyncratic shock and is given by the product of risk aversion and the systematic volatility of the cash flow. Empirically the Sharpe ratio is around 30% annually (see Constantinides and Ghosh (2011)). In order to fit this measure, we set $\delta^0$ at 10% and the risk aversion coefficient $\gamma$ at 3. The discount factor $\rho$ used by the representative agent in order to maximize her expected utility is set to 5%.

### 2.4.2 Theoretical implications

The model developed in this paper allows computing closed-form solutions for several measures. It dramatically eases the interpretation of the results. The two most important variables impacting this model are the state of the economy, which is represented by $p_t$, and the idiosyncratic shock $Y_{it}$, which is a proxy for stock beta. This section displays graphs of measures that depend on either one of those two variables or both.

Our model differs from the standard case (Veronesi (2000)) due to the addition of an idiosyncratic shock. The model is able to generate betas, price-earning ratio and cross-correlation between assets, which isn’t possible in the canonical case. On the other hand,
the market price of risk, as predicted by standard theory, is not impacted by the addition of the idiosyncratic shock.

Figure 2.1 plots the evolution of three stock characteristics for different state of the economy and different idiosyncratic shock level around the average, namely (a) market betas, (b) price-earning ratios and (c) stock prices.

Market betas are increasing in $Y$. It means that a larger idiosyncratic shock than the average shock increases the covariance between the market and the stock. Conversely, a smaller shock decreases the beta of the stock. The effect of $Y$ is more pronounced when market condition uncertainty is at its maximum, i.e. when $p_t = 0.5$.\(^8\)

![Figure 2.1](image)

(a) beta  
(b) Price Earning ratio  
(c) Stock Price

**Figure 2.1:** Effect of $Y$ and $p$ on different measures

The stock price is increasing in $Y$, more earnings lead to higher stock valuations. It is, however, decreasing in $p_t$. It can come as a surprise to the reader but one should note that this graphs is plotted assuming a constant aggregate cash flow $X_t$. The aggregate cash flow obviously increases in good times, hence, prices are overall increasing in $p_t$. This graph should be read as a representation of the part of the idiosyncratic shock price that is decreasing in economic expansion.

Figure 2.2 is a 3-D view of the (a) total volatility with respect to the economic state

\(^8\)market beta range can be increased by increasing the range of $Y$. 

and the idiosyncratic shock level. It also shows the decomposition of the total volatility between (b) systematic and (c) idiosyncratic volatilities. Both the total volatility and the systematic volatility are affected by the volatility of the conditional probability dynamic of the economy, $\sigma_p$, hence the U-shape. It means that systematic and total volatilities are increasing in extreme market conditions. It has been shown that aggregate (market) volatility tends to be high in recessions (e.g Schwert (1989)).

Idiosyncratic volatility increases in $p_t$. It implies that in normal market conditions, the idiosyncratic volatility will play a more important part on the total volatility. Surprisingly, the idiosyncratic volatility is negatively affected by the idiosyncratic shock. The direct implication is that low beta stocks have a larger volatility than higher beta stocks. The effect of $Y$, a proxy for stock betas, is marginal on the systematic volatility.

![Figure 2.2: Effect of Y on different volatility measures](image)

A breakdown of the total volatility level shows that the major part comes from the aggregate volatility. Indeed, the total volatility ranges from 7% to 10.5% annually in our model and the aggregate volatility ranges from 6% to 10%. Idiosyncratic volatility only has a small impact and has a range between 3% and 4%. These results are close to empirical data (see Section 2.5).

The main goal of this paper is to study the effect of stocks characteristics on the
cross-correlation between assets. Figure 2.3.a shows the entire mapping of the interaction between market condition and stock betas on cross-correlation between assets. Figure 2.3.b plots the effect of the state of the economy on three different stock pairs, namely stocks with a low beta, with a medium beta, and with a large beta. It is worth noting that the combination between a low beta stock and a high beta stock leads to a cross-correlation similar to a medium/medium combination.

(a) Cross-correlation for all (similar) pairs

(b) Cross-correlation for low beta, medium beta and high beta pairs

**Figure 2.3:** Effect of probability on cross-correlation for different beta pairs

First, we see from these two graphs that the cross-correlation is positively affected by
extreme states of the economy. The cross-correlation exhibits a U-shape for all stock pairs. It means that cross-correlation increases as people are more confident about the state of the economy. This result fits the intuition. One would expect that more stocks move in the same direction in a boom or in a burst. The vast majority of stocks tend to move up or down and at a similar pace once everyone agrees about the state of the economy. While the way a stock behaves when $p_t$ is close to 0.5 is more likely to be erratic compared to its peers. A more careful look at the graph shows the cross-correlation increasing slightly faster and higher in in bad times than in good times.

The second observation is that the pair of stocks with low beta exhibits a significantly lower cross-correlation than the medium/medium beta couple, which in turns has a lower cross-correlation than the pair high/high beta. This observation is valid for any state of the economy. The difference between different stock pairs diminishes as the uncertainty related to the state of the economy resolves. The cross-correlation ranges from 0.71 to 0.86 for low beta stocks and 0.82 to 0.91 for high beta stocks. The average correlations across the business cycle, given by $\int_0^1 \rho dp$, are 0.77 and 0.85 for low and high beta stocks respectively. These results are close to what is observed in the market as the following section shows.

### 2.5 Empirical results

#### 2.5.1 Data

We use data from the Center for Research in Security Prices (CRSP) for firm level (daily) return and weighted value market (daily) return. Additional data is obtained from the Kenneth French data library.

We use different data series to validate our predictions. Dataset 1 focuses on the top 500 market capitalization in the US market, divided into 20 portfolios according to beta. Dataset 2 and 3 are from Kenneth French data library, namely the 48 industry portfolios and the 100 BTM portfolios. Dataset 4 is composed of the 10 beta deciles portfolios from CRSP. We construct Dataset 5 by aggregating the 100 BTM portfolios into 10 beta decile portfolios. Table 2.2 lists datasets. The largest sample period for our analysis ranges from January 1963 to December 2014. Some of the analysis is conducted on smaller sample due to data availability and indications are provided when needed.

---

For each quarter, we use the past 1-year daily returns to compute the beta for the stocks. We then rank stocks according to their beta. The ranked stocks are then sorted in 20 bins with ascending order of beta. We form equally weighted portfolios from stocks belonging in each bin. The resulting portfolios represent stocks with beta sorted in ascending order.
Comparative statics are displayed in table 2.3. The difference across portfolios is large, average returns span from 12.4% to 19.8% and volatilities from 13.9% to 20.3% annually. Most datasets are representative of the economy as their average beta is close to one. The pairwise correlations are in the range of 0.54 to 0.89. There is a good variation in pairwise correlations over time.

We use the Aruoba-Diebold-Scotti Business Conditions Index (hereafter ADS index) as a proxy for $p_t$, the conditional probability of the economy being in high state at time $t$. The ADS index is designed to track real business conditions at high frequency. Its underlying (seasonally adjusted) economic indicators (weekly initial jobless claims; monthly payroll employment, industrial production, personal income less transfer payments, manufacturing and trade sales; and quarterly real GDP) blend high- and low-frequency information and stock and flow data. ADS index data is obtained from the website of the Federal Reserve Bank of Philadelphia.
2.5.2 Empirical patterns observed in the data

In this section we want to see whether the theoretical implications of our model (as outlined in Section 2.4) match the patterns observed in the data. The following procedure is done for each dataset. For each year, we use the daily returns of the stocks within that year to compute the beta for the stocks. We then rank stocks according to their beta. The ranked stocks are then sorted in different bins with ascending order of beta. We form equally weighted portfolios from stocks belonging in each bin. The resulting portfolios represent stocks with beta sorted in ascending order. Volatility and cross-correlation are then calculated for each bins. We define extreme market conditions as when the ADS index moves to its 10% tails (2 tails). Other time is defined as normal market conditions.

Table 2.4 shows how volatilities change as the market conditions change, across different beta portfolios. As beta moves up, all volatilities (ie. total, systematic, and idiosyncratic volatilities) increase. Irrespective of beta, volatility is higher when the market moves to extreme conditions. In addition, our model predicts that in normal market conditions, the idiosyncratic part plays a more important part of the total volatility, this is confirmed by the data.

Similarly, Table 2.5 shows how pairwise correlations change as the market conditions change, across different beta portfolios. A typical graphical representation of the time series of correlations can be found in Figure 2.4. Consistent with our model prediction regarding the level of correlations, we observe that higher beta portfolios have higher cross correlations. Irrespective of beta, correlations increase as the market moves into extreme conditions, but the increase is stronger in the lower beta group. The dispersion of the cross correlations also confirms our model prediction. The higher the beta, the lower the dispersion of cross-correlation.

2.5.3 Model implied cross-correlation

In this section, we want to see the ability of our model to generate correlation paths depending on the state of the economy to match the level of cross-correlation in the data (as in Figure 2.4). We test our model on the five datasets presented in Section 2.5.1. It is important to note that the U-shaped pattern is obtained holding beta constant and only varying $P_t$, the conditional probability of the high state. Empirically stock beta

\(^{10}\)For robustness check, we also use moving window technique to form the time series, rather than choosing data based on each calendar year. The conclusions remain unchanged.

\(^{11}\)The proportion of idiosyncratic volatility in the total volatility is calculated as $\frac{\sigma_{\text{idio}}^2}{\sigma_{\text{total}}^2}$

\(^{12}\)Figure 2.4 is for the top 500 firms of S&P where a moving window approach is used, with window size of 1 year and moving step of one quarter.
All portfolios | Low Beta Portfolios | High Beta Portfolios
---|---|---
Dataset 1: Top 500 firms in S&P
Volatility | 0.170 | 0.160 | 0.209 | 0.104 | 0.099 | 0.122 | 0.260 | 0.246 | 0.321
Vol. Sys. | 0.149 | 0.139 | 0.190 | 0.072 | 0.067 | 0.094 | 0.239 | 0.224 | 0.302
Vol. Idio. | 0.074 | 0.072 | 0.080 | 0.068 | 0.067 | 0.072 | 0.099 | 0.097 | 0.107
Dataset 2: 48 Industry portfolios
Volatility | 0.203 | 0.193 | 0.246 | 0.174 | 0.170 | 0.192 | 0.257 | 0.240 | 0.332
Vol. Sys. | 0.145 | 0.135 | 0.192 | 0.090 | 0.083 | 0.123 | 0.203 | 0.189 | 0.268
Vol. Idio. | 0.130 | 0.127 | 0.143 | 0.137 | 0.137 | 0.136 | 0.150 | 0.142 | 0.188
Dataset 3: 100 BTM portfolios
Volatility | 0.172 | 0.163 | 0.211 | 0.134 | 0.127 | 0.166 | 0.217 | 0.205 | 0.267
Vol. Sys. | 0.136 | 0.128 | 0.176 | 0.099 | 0.091 | 0.131 | 0.180 | 0.170 | 0.228
Vol. Idio. | 0.098 | 0.095 | 0.110 | 0.086 | 0.084 | 0.095 | 0.114 | 0.109 | 0.133
Dataset 4: 10 beta decile portfolios
Volatility | 0.139 | 0.127 | 0.190 | 0.100 | 0.092 | 0.136 | 0.178 | 0.163 | 0.243
Vol. Sys. | 0.121 | 0.109 | 0.171 | 0.079 | 0.071 | 0.114 | 0.162 | 0.148 | 0.227
Vol. Idio. | 0.063 | 0.060 | 0.076 | 0.056 | 0.053 | 0.068 | 0.069 | 0.066 | 0.084
Dataset 5: 10 BTM-beta decile portfolios
Volatility | 0.150 | 0.141 | 0.189 | 0.138 | 0.130 | 0.175 | 0.162 | 0.153 | 0.202
Vol. Sys. | 0.137 | 0.128 | 0.176 | 0.122 | 0.113 | 0.160 | 0.151 | 0.143 | 0.191
Vol. Idio. | 0.057 | 0.056 | 0.064 | 0.060 | 0.059 | 0.066 | 0.055 | 0.053 | 0.062

| Table 2.4: Empirical volatility patterns for dataset 1-5. “Vol. Sys.” and “Vol. Idio.” stand for systematic volatility and idiosyncratic volatility respectively. “All”, “Norm.” and “Ex” represent different market conditions, ie. all time, normal time, and extreme time. Extreme market condition is defined as the time when the ADS index lies in its 10% tails. |
54 Pairwise Correlation Dynamics and Incomplete Information

<table>
<thead>
<tr>
<th>Dataset 1: Top 500 firms in S&amp;P</th>
<th>All portfolios</th>
<th>Low Beta Portfolios</th>
<th>High Beta Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Norm.</td>
<td>Ex.</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.742</td>
<td>0.736</td>
<td>0.771</td>
</tr>
<tr>
<td>Std(corr)</td>
<td>0.098</td>
<td>0.099</td>
<td>0.096</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset 2: 48 Industry portfolios</th>
<th>All portfolios</th>
<th>Low Beta Portfolios</th>
<th>High Beta Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Norm.</td>
<td>Ex.</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.542</td>
<td>0.528</td>
<td>0.606</td>
</tr>
<tr>
<td>Std(corr)</td>
<td>0.129</td>
<td>0.130</td>
<td>0.107</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset 3: 100 BTM portfolios</th>
<th>All portfolios</th>
<th>Low Beta Portfolios</th>
<th>High Beta Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Norm.</td>
<td>Ex.</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.668</td>
<td>0.662</td>
<td>0.698</td>
</tr>
<tr>
<td>Std(corr)</td>
<td>0.121</td>
<td>0.119</td>
<td>0.129</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset 4: 10 beta decile portfolios</th>
<th>All portfolios</th>
<th>Low Beta Portfolios</th>
<th>High Beta Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Norm.</td>
<td>Ex.</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.836</td>
<td>0.826</td>
<td>0.885</td>
</tr>
<tr>
<td>Std(corr)</td>
<td>0.071</td>
<td>0.073</td>
<td>0.031</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset 5: 10 BTM-beta decile portfolios</th>
<th>All portfolios</th>
<th>Low Beta Portfolios</th>
<th>High Beta Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Norm.</td>
<td>Ex.</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.886</td>
<td>0.881</td>
<td>0.906</td>
</tr>
<tr>
<td>Std(corr)</td>
<td>0.051</td>
<td>0.053</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Table 2.5: Empirical correlation patterns for dataset 1-5. “Std(corr)” measures the dispersion of all cross-correlations. “All”, “Norm.” and “Ex” represent different market conditions, i.e. all time, normal time, and extreme time. Extreme market condition is defined as the time when the ADS index lies in its 10% tails.

our sample. With this definition $P_t^{imp}$ is by construction between 0 and 1.

Having defined $P_t^{imp}$, we compute the average betas for low $\beta$ bin and high $\beta$ bin. The spread in theoretical betas, as seen in Figure 2.1.a, generated from the model being smaller than the one in the data, we map empirical betas on the range of theoretical betas.\textsuperscript{13} It allows us to get more dispersion in cross-correlation. We then find the closest match between mapped average beta and theoretical beta. Finally, we get the cross-correlation implied by the model between stocks having similar betas.\textsuperscript{14} Repeating this procedure for each year gives theoretical cross-correlation trajectories. We compute yearly empirical correlation by using daily returns. We consider the period 1971-2014 and two sub-periods of equal size, namely 1971-1992 and 1993-2014. To simplify the analysis we only consider two groups of stocks, namely low and high beta stocks. Figure 2.5 through 2.9 illustrates the evolution of the sample and implied cross-correlations for the different datasets.

We can clearly see in most dataset that there is a strong relationship between the observed empirical curve and the theoretical one. This is especially striking over the

\textsuperscript{13} We map betas on the period 1971-2014, whatever the period under consideration.

\textsuperscript{14} Mapping each stocks and then computing average cross-correlation leads to larger correlation dispersion. We chose to use average betas to get less volatile results.
Figure 2.4: Cross-correlation over time for Dataset 1 (top 500 firms of S&P)

second half of the period. We formally test it by performing a regression analysis using non-overlapping data for each dataset and period. We have

\[ \rho^{*}_{ijt} = g_0 + g_1 \rho_{iY_t^{imp}, Y_j^{imp}, P_t^{imp}} + \varepsilon_{ijt} \]

where \( \rho^{*}_{ijt} \) is the sample correlation between beta categories \( i \) and \( j \). We test the null hypothesis of \( g_1 \) being different than 1. Table 2.6 shows the results.

R-squared tends to be higher for low beta stocks than for high beta stocks in most cases, whichever the period. The explanatory power over 1971-2014 period is not significant for all datasets. R-squared ranges from 0% to 7%. Dataset 1 has the poorest performance, with R-squared ranging from 4% to 6% while dataset 4 has R-squared between 15% and 27%. \( g_1 \) is not significantly different from 1 in 2 cases out of 10 at a 5%
Figure 2.5: Comparison of the empirical correlation with the model implied correlation for Dataset 1. The red dashed line represents empirical estimates. The blue continuous line represents simulated values.
Figure 2.6: Comparison of the empirical correlation with the model implied correlation for Dataset 2. The red dashed line represents empirical estimates. The blue continuous line represents simulated values.
Figure 2.7: Comparison of the empirical correlation with the model implied correlation for Dataset 3. The red dashed line represents empirical estimates. The blue continuous line represents simulated values.
Figure 2.8: Comparison of the empirical correlation with the model implied correlation for Dataset 4. The red dashed line represents empirical estimates. The blue continuous line represents simulated values.
**Figure 2.9:** Comparison of the empirical correlation with the model implied correlation for Dataset 5. The red dashed line represents empirical estimates. The blue continuous line represents simulated values.
### Table 2.6: Regression analysis using model implied correlation for Dataset 1-5. The regression estimated is $\rho_{ijt}^* = g_0 + g_1 \rho(Y_{i,\text{imp}}^t, Y_{j,\text{imp}}^t, P_{t,\text{imp}}) + \varepsilon_{ijt}$. We test the null hypothesis $g_1$ is different from 1 and report the t-stat value.

<table>
<thead>
<tr>
<th>Year</th>
<th>Dataset 1</th>
<th>Dataset 2</th>
<th>Dataset 3</th>
<th>Dataset 4</th>
<th>Dataset 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low beta</td>
<td>High beta</td>
<td>Low beta</td>
<td>High beta</td>
<td>Low beta</td>
</tr>
<tr>
<td>1971-2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.06</td>
<td>0.04</td>
<td>0.14</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>$g_1$</td>
<td>0.40</td>
<td>-0.46</td>
<td>0.51</td>
<td>0.50</td>
<td>0.49</td>
</tr>
<tr>
<td>t-stat</td>
<td>2.49</td>
<td>4.08</td>
<td>2.60</td>
<td>2.75</td>
<td>3.45</td>
</tr>
<tr>
<td>1971-1992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>$g_1$</td>
<td>0.24</td>
<td>-0.25</td>
<td>-0.16</td>
<td>-0.03</td>
<td>0.10</td>
</tr>
<tr>
<td>t-stat</td>
<td>3.59</td>
<td>5.13</td>
<td>6.08</td>
<td>3.36</td>
<td>5.75</td>
</tr>
<tr>
<td>1993-2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.13</td>
<td>0.11</td>
<td>0.28</td>
<td>0.24</td>
<td>0.46</td>
</tr>
<tr>
<td>$g_1$</td>
<td>0.88</td>
<td>-1.13</td>
<td>0.90</td>
<td>0.65</td>
<td>0.95</td>
</tr>
<tr>
<td>t-stat</td>
<td>0.23</td>
<td>2.98</td>
<td>0.32</td>
<td>1.37</td>
<td>0.24</td>
</tr>
</tbody>
</table>
level. When looking at sub-periods, we see that all the explanatory power comes from the 1993-2014 period, as suggested in Figures 2.5 to 2.9. R-squared ranges from 11% for dataset 1 to 51% for dataset 5. All but three $g_1$ are not significantly different from 1. The explanatory power of our model is thus quite large for the latest period. It suggests that the model should be able to improve Marokwitz portfolio allocation results.

2.6 Implication for portfolio allocation

Correlation plays a key role in portfolio allocation. Under the Markowitz framework, the global minimum variance portfolio weights are computed using the inverse of the covariance matrix. A correct assessment of the correlation among assets at the investment horizon is crucial for a successful implementation of this approach. The model developed in this paper takes into account the evolution of the state of the economy, enabling us to compute the correlation between stocks for different investment horizons.

![Figure 2.10: Standard deviation of the conditional probability for different horizons.](image)

Figure 2.10 displays the volatility of the high state probability for various horizon conditional on the current state, which is formally defined as

$$\sqrt{E[(P_T - E(P_T))^2 | \mathcal{F}_t]}.$$  

As the horizon, $T$, increases, uncertainty about the future state of the economy decreases. This result is expected since a longer maturity leads to the steady state with a higher probability. This behavior for the volatility of the conditional probability has
Pairwise Correlation Dynamics and Incomplete Information

63

a direct impact on cross-correlation as shown in Figure 2.11. The impact is different depending on stock betas. Low beta stocks exhibit a larger increase in cross-correlation than higher beta stocks, as the investment horizon increases. This is exacerbated when the current state of the economy is uncertain ($p_t = 0.5$).

A portfolio manager using the instantaneous correlation matrix will be mislead by its sample estimation translating to an over/under-weight of the allocation in each asset depending on the state of the economy. Our model takes into account time-variation of cross-correlations. Hence, we can correct the weight estimates in order to improve portfolio performance. Cross-correlation increases for every stock pair with time horizon. In other terms, the efficient frontier computed with past returns, which are approximations of the instantaneous values, is not the true frontier that will be generated at maturity.

Consider for example an investor, at time $t$, using a set of $n$ stocks with different betas satisfying $\beta_1 < \ldots < \beta_n$. With an horizon $T$, the investor should consider the covariance matrix of stock returns defined as

$$R_{it,T} = \frac{S_{iT} + \int_t^T CF_{it} ds - S_{it}}{S_{it}}$$

where $CF_{it}$ corresponds to the dividend gains. The investor must therefore estimate the conditional covariance of stock returns for all pairs $i \neq j$

$$\text{cov} (R_{it,T}, R_{jt,T}|\mathcal{F}_t).$$

If instead, the investor uses the contemporaneous covariance matrix constructed from $\text{cov} (R_{it}, R_{jt})$, he will obtain a biased estimate of the efficient frontier. We quantify this effect in Figure 2.12. As correlation tend to increase with maturity the benefit of diversification are amplified if one considers the contemporaneous covariance matrix. We focus our attention on the global minimum variance portfolio (GMVP), which is entirely defined by the estimated covariance matrix. In Figure 2.13 we analyze the volatility of the GMVP using the instantaneous correlations or the expected correlation at the investment horizon. For all initial states of the economy, the volatility of the GMVP is significantly lower when using the appropriate correlation measure. With our parameterization, for an horizon of one year, the gains are approximately 50 basis points for an average volatility of 10.5 %.

In practice, the instantaneous volatility is not directly observable and must be estimated from past data. Typically the estimation window for a portfolio allocation decision
Figure 2.11: Evolution of correlation between different beta stock pairs
Figure 2.12: Efficient frontier using instantaneous estimates (blue) and using estimates at maturity (red). Circles represent realized tangency portfolio, stars represent minimum variance portfolios. Th indicates measures using correlation at the investment horizon and i indicates measures using instantaneous correlation.

Figure 2.13: Standard deviation of the GMVP return using instantaneous estimates (blue) and using estimates at maturity (red).
ranges from a few months to a few years. It is also quite common to use an estimation window, which correspond to the investment horizon. In order to assess the ability of our model to improve portfolio allocation we propose a methodology to map correlation estimates into expected correlation at the horizon. Because of the strong nonlinearity in our stochastic processes it is not possible to obtain closed form expression for the model implied estimate of correlation based on past data, or the expected correlation of stock return at the horizon. We therefore rely on simulations to construct the mapping between correlation estimates and expected correlation at the horizon.

To simplify the approach we discretize the two main dimensions of the model, the conditional probability and the idiosyncratic shock. We consider 10 beta bins, $\beta_1 < \cdots < \beta_{10}$ and 5 probability bins, $P_1 < \cdots < P_5$. For an estimation window length $\tau$, and an investment horizon $T$ we construct by simulation the following function

$$
\psi(\beta_i, \beta_j, P_k; \tau, T) = \frac{\text{cov} (R_{it,T}, R_{jt,T}|P_k)}{\text{cov}_{t,\tau} (R_i, R_j)}
$$

where $\text{cov}_{t,\tau} (R_i, R_j)$ is the sample estimate of the correlation between stock $i$ and stock $j$ using an estimation window of length $\tau$.

We can use the function $\psi(\cdot)$ to estimate stock correlations for an investment horizon $T$ given a standard sample estimate of stock correlation. In order to limit the effect of estimation errors and maintain the constructed portfolio within reasonable and realistic allocation weights range, we focus our attention on a dataset with small dimension, 10 portfolios constructed by aggregating the Fama-French 100 BTM into beta decile portfolios, and consider the GMVP with a short sale constraint ($cGMVP$).\footnote{It is well known that allowing for short sale constraint yield very negative short positions combined with highly leveraged long position which are not realistic.} Following the procedure of the previous section, we obtain estimates of the conditional probability of the business cycle condition from the ADS index which we map into the interval $[0, 1]$. We construct the cGMVP portfolio using a rolling window approach starting for the period 1971-2014. We compare the performance of the sample based portfolio with the performance of the transformed portfolio using the function $\psi(\cdot)$.

We consider two investment horizons of respectively 6 months and 12 months, and three estimation windows of 72, 144 and 216 days. Table 2.7 compares the standard deviation and mean return of the two approaches. In all cases, we observe a decrease in the standard deviation using the modified estimates. The largest difference, 2.53 percentage points, is obtained for the 12 month investment horizon with the estimation window of 72 days. While it is not part of the objective function of cGMVP and cannot
<table>
<thead>
<tr>
<th>Invest. Horizon</th>
<th>$P_k$</th>
<th>Estimation window $\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>126</td>
<td>Sample</td>
<td>Modified</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>16.91%</td>
<td>15.27%</td>
</tr>
<tr>
<td>$\mu$</td>
<td>7.65%</td>
<td>7.08%</td>
</tr>
<tr>
<td>252</td>
<td>Sample</td>
<td>Modified</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>24.70%</td>
<td>22.17%</td>
</tr>
<tr>
<td>$\mu$</td>
<td>15.94%</td>
<td>14.65%</td>
</tr>
</tbody>
</table>

Table 2.7: Portfolio allocation performance. The sample period is 1971-2014. The dataset is the 100 BTM portfolio from Fama and French. The portfolio are aggregated into 10 beta decile portfolios. For each estimation window the column Sample displays the annualized standard deviation, $\sigma$, and average return, $\mu$, of the cGMVP portfolio using the sample estimate of the covariance matrix. The column Modified uses the modified estimate of the covariance matrix using the function $\psi(\cdot \cdot \cdot)$ defined in section 2.6. Results are presented for the entire sample.

serve as a performance measure, we note that the average return using the modified estimate is usually lower than its counterpart using the sample estimate.

Table 2.8 and Table 2.9 provide more details on the performance of our approach by decomposing the results in the 5 conditional probability levels $P_1, \ldots, P_5$. We can see that the risk reduction mainly occurs when the conditional probability is low (especially at $p = 2$), i.e. during recession periods the correction generated by the model has a more significant effect on the performance of the cGMVP allocation. Table 2.9 shows the difference in performance between the standard model and our augmented model. At a short term investment horizon of six months, the difference in volatility using our treatment is always positive with a 72 days and a 216 days estimation window and three times out of five with a 144 days estimation window. It is important to note that rise in risk is rare and low (< 1%) while risk reduction ranges from +0.02% to +9.06%. At an investment horizon of one year, portfolio performance is significantly improved during period of boom, whatever the choice of estimation window length. Performance during recession is also strong, however, we note a decrease in some cases. Once again, rise in risk is low (from -0.24% to -1.43%) compared to risk reduction (from +0.17% to +6.98%).

The discussion above illustrates the fact that correlation dynamics at any horizon should be conditioned on the current business cycle condition. While our approach tries to link this conditioning to the precise specification of our theoretical model, and does so relatively successfully, a more general form of conditioning on business cycle conditions might be a future avenue of research.
<table>
<thead>
<tr>
<th>Invest. Horizon $P_k$</th>
<th>Estimation window $\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>72</td>
</tr>
<tr>
<td>126</td>
<td>Sample</td>
</tr>
<tr>
<td>1</td>
<td>27.64%</td>
</tr>
<tr>
<td>2</td>
<td>35.36%</td>
</tr>
<tr>
<td>3</td>
<td>29.47%</td>
</tr>
<tr>
<td>4</td>
<td>16.00%</td>
</tr>
<tr>
<td>5</td>
<td>15.00%</td>
</tr>
<tr>
<td>252</td>
<td>Sample</td>
</tr>
<tr>
<td>1</td>
<td>24.06%</td>
</tr>
<tr>
<td>2</td>
<td>28.87%</td>
</tr>
<tr>
<td>3</td>
<td>24.00%</td>
</tr>
<tr>
<td>4</td>
<td>23.25%</td>
</tr>
<tr>
<td>5</td>
<td>23.77%</td>
</tr>
</tbody>
</table>

Table 2.8: Portfolio allocation performance. The sample period is 1971-2014. The dataset is the 100 BTM portfolio from Fama and French. The portfolio are aggregated into 10 beta decile portfolios. For each estimation window the column Sample displays the annualized standard deviation of the cGMVP portfolio using the sample estimate of the covariance matrix. The column Modified uses the modified estimate of the covariance matrix using the function $\psi(\cdots)$ defined in section 2.6. Results are presented for each conditional probability levels $P_1, \ldots, P_5$ as defined in section 2.6.
<table>
<thead>
<tr>
<th>Invest. Horizon</th>
<th>$P_k$</th>
<th>Estimation window $\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>126</td>
<td>1</td>
<td>0.02%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9.06%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.26%</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.40%</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.58%</td>
</tr>
<tr>
<td>252</td>
<td>1</td>
<td>-0.24%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6.98%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-1.43%</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.54%</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>3.24%</td>
</tr>
</tbody>
</table>

Table 2.9: Portfolio allocation performance. The sample period is 1971-2014. The dataset is the 100 BTM portfolio from Fama and French. The portfolio are aggregated into 10 beta decile portfolios. For each estimation window the table shows the difference in annualized standard deviation of the cGMVP portfolio between the approach using the sample estimate of the covariance matrix and the approach using the modified estimate of the covariance matrix using the function $\psi(\cdots)$ defined in section 2.6. Text appearing in red indicates a lower standard deviation for the sample estimate approach. Results are presented for each conditional probability levels $P_1, \ldots, P_5$ as defined in section 2.6.
2.7 Conclusion

The model developed in this article aims at understanding the dynamics of the correlations between assets. There is ample evidence in the literature that stock correlations are not constant through time. In particular, they tend to increase when the economy is facing major shocks. Most of the literature on the topic points out that negative shocks imply larger increase in cross-correlation than positive shocks.

Correlation plays a crucial role for portfolio allocation. If the benefits of diversification vanish in case of extreme market conditions, investors are not protected as they assume they are. While this should not be a serious concern when the economy is in a boom, it is, however, a major drawback when the economy enters a recession. It is therefore important to carefully consider the evolution of the cross-correlations.

In this paper we provide a framework that allows studying both the time series and cross sectional dimension of correlations. We propose an equilibrium model of asset prices with aggregate and idiosyncratic risks, where information about the aggregate component is incomplete. Learning about the unobservable component creates business cycles that are conditioned by the time varying probability of a given state. Shocks to the idiosyncratic component affect stock characteristics conditional on the aggregate state of the economy, creating a dynamic pattern of cross sectional heterogeneity. It allows us to sort stocks according to their market beta and to infer their pairwise correlation dynamics over time.

We provide empirical evidence that support the predictions of our model along several dimension. First, we find that cross-correlations within the low beta stocks are lower than for the high beta stocks. Second, our model generates higher average correlation between stocks during extreme market condition. Finally we consider the implications of our model for portfolio allocation, and we quantify the performance loss associated with an inadequate estimation of the stock return correlations. We show that conditioning on current market conditions, in particular conditional recession probabilities, and using our model’s structure generates important changes in the efficient frontier.

The CRRA preference structure in this paper does not allow studying the dynamics of correlation risk premia, further research in that direction seems warranted.
Conclusion

In this dissertation, I analyze portfolio allocation problems on large data sets as well as correlation modeling. In Chapter 1 Tony Berrada and I construct mean-variance portfolios using a factor model approach. We show the importance of portfolio allocation for large unbalanced equity data sets. We consider the complete CRSP database, which includes all the US stocks in the past 50 years. We find that these portfolios are still subject to large losses during crisis.

Motivated by our findings, we analyzed more deeply the cause of the losses and find that during crisis, the pairwise correlation between stocks tends to increase significantly. Hence, the benefits of diversification are lost.

In Chapter 2, we build a general equilibrium model that is able to generate the time variation and the cross-sectional variation observed in the data. We show that correcting the sample variance-covariance matrix with the prediction of our model significantly reduces portfolio risk for the global minimum variance portfolio.
Appendix to Pairwise Correlation Dynamics and Incomplete Information

**Proof.** Proposition 1

\[ e^{-\rho t} X_t^{-\gamma} S_t = E \left[ \int_t^\infty e^{-\rho s} X_s^{1-\gamma} Y_s ds \mid \mathcal{F}_t \right] \]

We have:

\[ dX_t = m_t X_t dt + \delta_t X_t d\nu_t^\alpha \]

\[ X_s = X_t \exp \left( \int_t^s (m_u - 0.5\delta_u^2) du + \int_t^s \delta_u d\nu_u^\alpha \right) \]

Simple algebra and replacing terms give

\[ \frac{S_t}{X_t} = E \left[ \int_t^\infty e^{-\rho(s-t)} Y_s \exp \left( (1 - \gamma) \int_t^s (m_u - 0.5\delta_u^2) du + (1 - \gamma) \int_t^s \delta_u d\nu_u^\alpha \right) ds \mid \mathcal{F}_t \right] \]

From the law of iterated expectation we can write

\[ \frac{S_t}{X_t} = E \left[ \int_t^\infty e^{-\rho(s-t)} [\eta + (Y_{it} - \eta)e^{-\alpha(s-t)}] \times \exp \left( (1 - \gamma) \int_t^s (m_u - 0.5\delta_u^2) du + (1 - \gamma) \int_t^s \delta_u d\nu_u^\alpha \right) ds \mid \mathcal{F}_t \right] \]

Applying Girsanov theorem, we have that under the new probability measure \( d\nu_t^* = \)
\[ \frac{d\nu_t^a}{\nu_t^a} - (1 - \gamma)\delta^a dt \]

\[
\frac{S_t}{X_t} = E^*[\int_t^\infty e^{-\rho(s-t)[\eta + (Y_{it} - \eta)e^{-\alpha(s-t)}]} \times \exp \left( (1 - \gamma) \int_t^s (m_u - 0.5\delta^2_a) du + (1 - \gamma)^2/2 \int_t^s \delta^2_a(s-t) \right) ds | \mathcal{F}_t] \]

Replacing \( m_u \) by its expression and rearranging terms leads to

\[
\frac{S_t}{X_t} = E^*[\int_t^\infty [\eta + (Y_{it} - \eta)e^{-\alpha(s-t)}] \times \exp \left[ (0.5\gamma(\gamma - 1)\delta^2_a - \rho + (1 - \gamma)\theta_L)(s-t) + (1 - \gamma)(\theta_H - \theta_L) \int_t^s p_u du \right] ds | \mathcal{F}_t] \]

\[
= E^*[\int_t^\infty [\eta + (Y_{it} - \eta)e^{-\alpha(s-t)}] \exp \left[ - \int_t^s (\rho - \bar{\theta}_L - (\bar{\theta}_H - \bar{\theta}_L)p_u) du \right] ds | \mathcal{F}_t] \]

\[
= \eta E^*[\int_t^\infty \exp \left[ - \int_t^s (\rho - \bar{\theta}_L - (\bar{\theta}_H - \bar{\theta}_L)p_u) du \right] ds | \mathcal{F}_t] + (Y_{it} - \eta)E^*[\int_t^\infty \exp \left[ - \int_t^s (\rho + \alpha - \bar{\theta}_L - (\bar{\theta}_H - \bar{\theta}_L)p_u) du \right] ds | \mathcal{F}_t] \]

where \( \bar{\theta}_H = (1 - \gamma)\theta_H + 0.5\gamma(\gamma - 1)\delta^2_a \) and \( \bar{\theta}_L = (1 - \gamma)\theta_L + 0.5\gamma(\gamma - 1)\delta^2_a \).

Applying Feynman-Kac theorem in conjunction with the Ansatz \( u(t, p_t) = A + Bp_t \) for the first part, the following PDE is satisfied

\[
\mu_p B - V(t, p_t)u(t, p_t) + 1 = 0 \]

where \( \mu_p = -\lambda_{12}p_t + \lambda_{21}(1 - p_t) + p_t(1 - p_t)(\bar{\theta}_H - \bar{\theta}_L) \) and \( V(t, p_t) = \rho - \bar{\theta}_L - (\bar{\theta}_H - \bar{\theta}_L)p_t \).

The solution of this PDE is

\[
A = \frac{\rho + \lambda_{12} + \lambda_{21} - \bar{\theta}_H}{\rho(\rho + \lambda_{12} + \lambda_{21}) - \theta_H(\rho + \lambda_{21} - \bar{\theta}_L) - \bar{\theta}_L(\rho + \lambda_{12})} \]

\[
B = \frac{\bar{\theta}_H - \bar{\theta}_L}{\rho(\rho + \lambda_{12} + \lambda_{21}) - \theta_H(\rho + \lambda_{21} - \bar{\theta}_L) - \bar{\theta}_L(\rho + \lambda_{12})} \]

with \( A, B > 0 \) and the denominator

\[
\rho(\rho + \lambda_{12} + \lambda_{21}) - \theta_H(\rho + \lambda_{21} - \bar{\theta}_L) - \bar{\theta}_L(\rho + \lambda_{12}) \neq 0 \]

Replacing \( \rho \) by \( (\rho + \alpha) \) gives the solutions for \( a \) and \( b \). This completes the proof. ■

**Proof.** Proposition 2 Volatilities can be derived directly by applying Itô’s lemma on the
equation of the stock price. We have

\[
dS_{it} = df(X_t, Y_{it}, p_t)
\]

\[
= \left[ \eta(A + B p_t) + (Y_{it} - \eta) (a + b p_t) \right] dX_t + X_t (a + b p_t) dY_{it}
\]

\[
+ \left[ X_t \eta B + X_t (Y_{it} - \eta) b \right] dp_t + \left[ \eta B + (Y_{it} - \eta) b \right] d\langle X, p \rangle_t
\]

\[
= \left[ \cdots \right] dt
\]

\[
+ \left[ \delta^a \eta (A + B p_t) + \delta^a (Y_{it} - \eta) (a + b p_t) + \eta B \sigma_t^p + \sigma_t^p (Y_{it} - \eta) b \right] X_t dv_i^a
\]

\[
+ X_t (a + b p_t) \delta^i d B^i_{it}
\]

By isolating \( S_{it} \), it follows that the volatilities are given by

\[
\sigma_{t, i}^a = \delta^a + \frac{(Y_{it} - \eta) b + \eta B}{(Y_{it} - \eta) (a + b p_t) + \eta (A + B p_t)} \sigma_t^p
\]

\[
\sigma_{t, i}^1 = \left[ Y_{it} + \eta (A + B p_t) - (a + b p_t) \right]^{-1} \delta^i
\]

We can compute the drift term easily by noticing that the following process is an \( F_t \)-Martingale

\[
M_{it} = \xi_t S_{it} + \int_0^t \xi_s C F_{is} ds
\]

\[
= \mathbb{E} \left[ \int_0^\infty \xi_s C F_{is} ds | F_t \right]
\]

It follows from the Martingale representation theorem (see Duffie (2001)) that

\[
M_{it} = M_{i0} + \int_0^t \phi_{is} dB_{is}
\]

where \( \phi_s = [\phi_{s,1}, \phi_{s,2}] \) is a two-dimensional \( F_t \)-adapted vector process and \( B_{is} \) is a two-dimensional Brownian motion representing aggregate and idiosyncratic shocks. We have that the dynamic of the stochastic discount factor is given by

\[
d\xi_t = \xi_t (-r dt - \kappa_{i} dv_i^a)
\]

Taking the derivative on both side of the following equation

\[
\xi_t S_{it} + \int_0^t \xi_s C F_{is} ds = M_{i0} + \int_0^t \phi_{is} dB_{is}
\]
Pairwise Correlation Dynamics and Incomplete Information

gives

\[ d(\xi_t S_{it}) + \xi_t CF_{it} dt = \phi_{it} dB_{it} \]

\[ \Leftrightarrow \xi_t [dS_{it} + CF_{it} dt] = -S_{it} d\xi_t - d(\xi_t, S_{it}) + \phi_{i,1} d\nu^{a}_t + \phi_{i,2} dB_{i,t}^i \]

\[ \Leftrightarrow \frac{dS_{it} + CF_{it} dt}{S_{it}} = [r_t + \kappa_t \sigma^a_{it}] dt + \phi_{i,1} d\nu^{a}_t + \phi_{i,2} dB_{i,t}^i \]

Hence, the drift term is given by

\[ \mu_{it} = r_t + \kappa_t \sigma^a_{it} \]

**Proof.** Here we prove the result obtained for the dynamic of the aggregate market return. Define \( \int_i S_{it} di = M_t \) as being the market. We have

\[ \int_i S_{it} di = \int_i [X_t \eta(A + B p_t) + X_t (Y_t - \eta)(a + b p_t)] di \]

\[ = X_t \eta(A + B p_t) = M_t \]

We have that

\[ dS_{it} + CF_{it} dt = \mu_{i,t} S_{i,t} dt + \sigma^a_{i,t} S_{i,t} d\nu^{a}_t + \sigma^i_{i,t} S_{i,t} dB_{i,t}^i \]

Replacing and rearranging terms gives

\[ dS_{it} + CF_{it} dt = r_t S_{i,t} dt + [\delta^a S_{i,t} + \sigma^p_{i} X_t ((Y_{i,t} - \eta)b + \eta B)] \gamma \delta^a dt \]

\[ + [\delta^a S_{i,t} + \sigma^p_{i} X_t ((Y_{i,t} - \eta)b + \eta B)] d\nu^{a}_t \]

\[ + \delta^i X_t (a + b p_t) dB_{i,t}^i \]

Taking the integral over a continuum of firms gives

\[ \int_i (dS_{it} + CF_{it} dt) di = r_t X_t \eta(A + B p_t) dt + \gamma \delta^a [\delta^a X_t \eta(A + B p_t) + \sigma^p X_t \eta B] dt \]

\[ + (\delta^a X_t \eta(A + B p_t) + \sigma^p X_t \eta B) d\nu^{a}_t \]

\[ = r_t M_t dt + \gamma \delta^a \sigma_{m,t} M_t dt + \sigma_{m,t} M_t d\nu^{a}_t \]

Result follows.
## List of Tables

1.1 List of models ............................................. 15
1.2 All CRSP between 1967 and 2012, tangency portfolio ................. 19
1.3 All CRSP between 1967 and 2012, MVTP strategy ...................... 21
1.4 Random size samples ................................... 24
1.5 Stock selected by minimum average volume ........................ 26
1.6 754 Stocks .................................................. 28
1.7 Comparison between restricted and full dataset .......................... 29
1.8 Replication .................................................. 30
1.9 All CRSP with 60 months rolling window .......................... 32
1.10 All CRSP with 180 months rolling window ........................ 32

2.1 Parameter values used in all computation in sections 2.4, 2.5 and 2.6. . . 45
2.2 This list depicts all datasets under consideration in the paper. $N$ represents the number of portfolios in the dataset. .......................... 51
2.3 Comparative statics for Dataset 1-5. Average returns and volatilities are annualized. $\rho$ represents correlation and $\beta$ sensibility to the market. ........ 51
2.4 Empirical volatility patterns for dataset 1-5. “Vol. Sys.” and “Vol. Idio.” stand for systematic volatility and idiosyncratic volatility respectively. “All”, “Norm.” and “Ex” represent different market conditions, ie. all time, normal time, and extreme time. Extreme market condition is defined as the time when the ADS index lies in its 10% tails. ............ 53
2.5 Empirical correlation patterns for dataset 1-5. “Std(corr)” measures the dispersion of all cross-correlations. “All”, “Norm.” and “Ex” represent different market conditions, ie. all time, normal time, and extreme time. Extreme market condition is defined as the time when the ADS index lies in its 10% tails. ................ 54
2.6 Regression analysis using model implied correlation for Dataset 1-5. The regression estimated is $\rho_{ijt}^s = g_0 + g_1 \rho(Y_{i}^{imp}, Y_{j}^{imp}, P_{t}^{imp}) + \epsilon_{ijt}$. We test the null hypothesis $g_1$ is different from 1 and report the t-stat value. ........ 61
2.7 Portfolio allocation performance. The sample period is 1971-2014. The dataset is the 100 BTM portfolio from Fama and French. The portfolio are aggregated into 10 beta decile portfolios. For each estimation window the column Sample displays the annualized standard deviation, \( \sigma \), and average return, \( \mu \), of the cGMVP portfolio using the sample estimate of the covariance matrix. The column Modified uses the modified estimate of the covariance matrix using the function \( \psi(\cdots) \) defined in section 2.6. Results are presented for the entire sample. 

2.8 Portfolio allocation performance. The sample period is 1971-2014. The dataset is the 100 BTM portfolio from Fama and French. The portfolio are aggregated into 10 beta decile portfolios. For each estimation window the column Sample displays the annualized standard deviation of the cGMVP portfolio using the sample estimate of the covariance matrix. The column Modified uses the modified estimate of the covariance matrix using the function \( \psi(\cdots) \) defined in section 2.6. Results are presented for each conditional probability levels \( P_1, \ldots, P_5 \) as defined in section 2.6. 

2.9 Portfolio allocation performance. The sample period is 1971-2014. The dataset is the 100 BTM portfolio from Fama and French. The portfolio are aggregated into 10 beta decile portfolios. For each estimation window the table shows the difference in annualized standard deviation of the cGMVP portfolio between the approach using the sample estimate of the covariance matrix and the approach using the modified estimate of the covariance matrix using the function \( \psi(\cdots) \) defined in section 2.6. Text appearing in red indicates a lower standard deviation for the sample estimate approach. Results are presented for each conditional probability levels \( P_1, \ldots, P_5 \) as defined in section 2.6.
List of Figures

1.1 Figure representing the cumulative return of tangency portfolios under the assumption of the four factor model based on the whole CRSP dataset (18,356 stocks). The blue line represents the results with a 1-pass estimation procedure. The green line is the results with the 2-pass estimation developed in this paper. ......................................................... 20

1.2 Figure representing the cumulative return of four strategies, namely 1/N, tangency portfolio (TP), minimum variance portfolio (GMVP) and minimum variance portfolio with a target return of 9% based on a 2-pass estimation, for the four factor model. The dataset is composed of 18,356 stocks. The four factors are the excess market return (ERM), the Fama-French factors (SMB,HML), and momentum........................................... 22

1.3 Figure representing the number of stocks with non zero weight in the portfolio through time. The investment universe is constituted of all stocks from CRSP database (18,356 stocks). ......................................................... 23

2.1 Effect of \( Y \) and \( p \) on different measures ................................................................. 47

2.2 Effect of \( Y \) on different volatility measures ............................................................... 48

2.3 Effect of probability on cross-correlation for different beta pairs .................... 49

2.4 Cross-correlation over time for Dataset 1 (top 500 firms of S&P) ............ 55

2.5 Comparison of the empirical correlation with the model implied correlation for Dataset 1. The red dashed line represents empirical estimates. The blue continuous line represents simulated values........................................... 56

2.6 Comparison of the empirical correlation with the model implied correlation for Dataset 2. The red dashed line represents empirical estimates. The blue continuous line represents simulated values........................................... 57

2.7 Comparison of the empirical correlation with the model implied correlation for Dataset 3. The red dashed line represents empirical estimates. The blue continuous line represents simulated values........................................... 58
2.8 Comparison of the empirical correlation with the model implied correlation for Dataset 4. The red dashed line represents empirical estimates. The blue continuous line represents simulated values. ................................................................. 59
2.9 Comparison of the empirical correlation with the model implied correlation for Dataset 5. The red dashed line represents empirical estimates. The blue continuous line represents simulated values. ................................................................. 60
2.10 Standard deviation of the conditional probability for different horizons. . 62
2.11 Evolution of correlation between different beta stock pairs ................. 64
2.12 Efficient frontier using instantaneous estimates (blue) and using estimates at maturity (red). Circles represent realized tangency portfolio, stars represent minimum variance portfolios. Th indicates measures using correlation at the investment horizon and i indicates measures using instantaneous correlation. ................................................................. 65
2.13 Standard deviation of the GMVP return using instantaneous estimates (blue) and using estimates at maturity (red). ................................. 65


