

Separating the Wheat from the Chaff: Understanding Portfolio Returns in an Emerging Market

Nicolas A. Eterovic^{a,1,*}, Dalibor S. Eterovic^{b,2}

^a *Wivenhoe Park, Colchester CO4 3SQ, UK.*

^b *Diagonal Las Torres 2640, Peñalolen, Santiago, Chile*

Abstract

In this paper we apply Random Matrix Theory (RMT) to study daily returns correlations of 83 companies that are part of the Chilean stock market during the period 2000 to 2011. We find that using RMT to identify statistically significant correlations within our sample of stocks significantly improves the efficiency of a family of Markowitz Portfolios. Moreover, by using Vector Autoregressive analysis we identify global risk aversion as the main driver of the Chilean equity market returns followed in importance by shocks to the monthly rate of inflation and the country's monetary policy rate. By studying the effects of macroeconomic variables on the constructed portfolio returns we reach a better understanding of the true risks involved in an emerging market portfolio.

Keywords:

Random matrix theory, Portfolio Optimization, Financial Markets and the Macroeconomy

JEL Classification G10; G11; E44.

[☆]We would like to thank the participants of the economic seminars at the Chilean Central Bank and the economic faculties of the Universidad Catolica de Chile and the Universidad de Chile for helpful comments. The usual disclaimer applies.

*Corresponding author

Email addresses: netero@essex.ac.uk (Nicolas A. Eterovic),
dalibor.eterovic@uai.cl (Dalibor S. Eterovic)

¹PhD Student, Department of Economics, University of Essex

²Grupo Security and School of Government, Universidad Adolfo Ibañez

1. Introduction

The Great Recession of 2008-2009 was a vivid reminder that financial correlations breakdown during periods of high volatility. It also highlighted the importance of identifying stable correlations in order to quantify the underlying risk of diversified portfolios. This paper contributes to current debates on portfolio optimization by examining the statistical significance of the correlations across the Chilean stock market.

Two questions frame our investigation. First, we ask if all the return correlations across the Chilean stock market are equally statistically significant. Second, we ask which are the main macroeconomic drivers affecting the Chilean stock market returns. To answer these questions, we use Random Matrix Theory (RMT) to study the daily returns of 83 Chilean stocks that are part of the IPSA and IGPA indices from January 2000 to January 2011. The RMT helps us to separate the wheat from the chaff in the correlation matrix. Using Markowitz's Portfolio Theory (MPT), we then compare the efficiency of portfolios constructed under RMT with others constructed under the standard approach which considers all covariances in the correlation matrix equally significant. Finally, we use Vector Autorregresion approach (VAR) to determine the impact of a set of macroeconomic and financial variables on the optimal portfolios derived from our significant eigenvalues.

We focus on the Chilean stock market as we think provides a good case study for other emerging markets due to a number of reasons. First, the Chilean equity market is one of the most developed within the emerging market world with a market capitalization of 120% of its GDP. Other markets such as the Brazilian one have much lower market capitalizations as percentage of their GDP (58%) while Mexico has a market capitalization of 39% of GDP. Second, Chile is a small open economy with almost no restrictions to the access of international investors. Therefore, both idiosyncratic and global factors are likely to be important determinants of the stock market returns. Assessing the relative importance of domestic versus international factors in explaining domestic market volatility is key for developing public policy and market regulation in other emerging economies that are following the path of financial liberalization.

Our main findings can be summarized as follows: First, applying Random Matrix Theory to a portfolio composed of Chilean equities improves its efficiency compared to a portfolio constructed using a standard MPT approach

by at least 48%.³ Second, using VAR analysis, we identify global risk aversion as the main macroeconomic determinant of the Chilean equity market returns followed in importance by shocks to the monthly rate of inflation and the country's monetary policy rate. Third, it is possible to diversify away some of the market portfolio risk by adding positions on the portfolios constructed by the second and third largest eigenvalues. Fourth, the three smallest eigenvalues produce portfolio returns that are mostly uncorrelated with macroeconomic shocks. These portfolios are also uncorrelated with the market portfolio.

In general, the stability and statistical significance of empirical correlations are crucial for risk and portfolio management, since the probability of large losses in a portfolio is mainly driven by the way the correlations between the assets in it behave. For example, a position which is simultaneously long in stocks and short in bonds is riskier than one holding any of those assets exclusively. This is because bonds and stocks usually move in opposite directions during periods of crisis (Laloux et al., 1999).

The use of correlation matrices has a long history in portfolio management and is one of the main ingredients of Markowitz's Portfolio Theory, which solves the following dual problem: given a set of financial assets characterized by their average return and risk; what is the optimal weight on each asset such that the overall portfolio provides the best return for a fixed level of risk, or equivalently, the lowest risk for a given return? The solution to this problem entails the resolution of a system of non-linear equations where the correlation matrix C has to be inverted. The final outcome of this exercise is a risk-return relation, the so-called "efficient frontier".

Despite the apparent success of this theory in explaining optimal portfolios, it has been widely criticized by its inability to provide good risk estimates. In particular, it tends to underestimate *realized* risk, in such a way that the approach has been abandoned in the financial industry and embraced by academicians who use it as an illustrative theory to explain optimal portfolios.

A possible explanation for this failure in the Markowitz's Portfolio Theory, has found room in the field of Statistical Physics, which states that the

³Throughout the paper, efficiency is measured by the Mean Percentage Error, which measures how close the risk of our portfolios are from the "true" realized risk. This is an average result over the sample analyzed in the most restrictive scenario, that is, without allowing for a "short-sale" strategy.

estimation of a correlation matrix may be difficult in case the length of the time series (henceforth, T) is not very large compared to the total number of assets (henceforth, N). In these cases we should expect the covariances to contain a lot of noise and therefore be, to a large extent, random. In particular, the smallest eigenvalues are precisely those that are the most sensitive to this noise and their associated eigenvectors the ones that determine the least risky portfolios (Laloux et al., 1999, 2000). With all these problems in mind, it becomes apparent the need to seek for new methods to separate out “signals”, which contain economically relevant information, from “noise” which do not. From this perspective, it turns out interesting to compare the properties of an empirical correlation matrix C with respect to a “null hypothesis” purely random, that one could obtain from a time series of completely independent assets. In this context, deviations from this random counterpart could indicate evidence of relevant economic information.

The Econophysics literature deals with this issue by using the Random Matrix Theory (RMT), a technique taken from Statistical Physics. RMT has been recently applied to the study of financial correlations (Laloux et al., 1999, 2000). A number of studies show the advantages of using the RMT approach to remove noise from correlation matrices, with significant improvements in portfolio choice (Plerou et al., 1999, 2000a,b, 2001, 2002; Gopikrishnan et al., 2001).

Our paper contributes to the academic literature in two important ways. First, by applying RMT to the Chilean stock market we provide further evidence of the benefits of using this technique in building efficient portfolios. Second, to the best of our knowledge we are one of the first to explore the relationship between macroeconomic variables and portfolio returns constructed using statistically significant correlations. The rest of the paper is organized as follows. Section 2 introduces the relevant literature on Random Matrix Theory and applies it to the Chilean Stock Market. In Section 3, we look for the macroeconomic and financial drivers of stock returns. In Section 4, we improve upon some of our results by removing the market effect from our portfolio. Finally, in Section 5, we conclude and propose lines for future research.

2. Random Matrix Theory (RMT)

Random Matrix Theory (RMT) was developed by Wigner (1955, 1958), in the field of quantum physics to explain the energy levels of complex nuclei.

Recently, starting with the work developed by Laloux et al. (1999, 2000), Plerou et al. (1999, 2000a,b, 2001, 2002) and Gopikrishnan et al. (2001), has found applications in a financial context.

The understanding of correlations between different stocks has been an ongoing quest in the last decades, not only because of a better understanding of the economy as a complex dynamical system, but also for practical reasons such as asset allocation and risk forecast. As opposed to most physical systems, where correlations between subunits can be related to basic interactions, in the case of financial markets, the underlying interactions are not known.

In order to quantify the cross-correlations between financial assets, we first need to calculate the price returns of stocks over a period of time Δt :

$$G_i(t) = \ln S_i(t + \Delta t) - \ln S_i(t) \quad (1)$$

where $S_i(t)$ denotes the price of stock i . As different stocks have different volatility levels, we proceed to define the normalized returns:

$$g_i = \frac{G_i(t) - \langle G_i \rangle}{\sigma_i} \quad (2)$$

where $\sigma_i = \sqrt{\langle G_i^2 \rangle - \langle G_i \rangle^2}$ is the standard deviation of G_i and $\langle \cdot \rangle$ denotes a time average over the period considered. We then compute the equal-time cross-correlation matrix C with elements:

$$C_{ij} = \langle g_i(t)g_j(t) \rangle \quad (3)$$

By construction, the elements C_{ij} are bounded to the domain $-1 \leq C_{ij} \leq 1$, where $C_{ij} = 1$ corresponds to perfect correlations, $C_{ij} = -1$ corresponds to perfect anti-correlations, and $C_{ij} = 0$ corresponds to uncorrelated pairs of stocks.

As pointed out by Plerou et al. (1999, 2000a, 2002), when estimating the cross-correlation matrix we have the following dilemma: (i) if we use a long time series of returns to calculate the correlation matrix, we will have that as market conditions change over time, the cross-correlations between any pair of stocks may not be stationary, which lead us to use a short time-series, but (ii) if we use a short time-series, the finite length of the time series available to estimate cross-correlations introduces “measurement noise”.

This poses the following question: How can we identify from C_{ij} those stocks that remained significantly correlated (on average) over the period

studied? To answer this question, the statistics of C have to be analyzed against the “null hypothesis” of a random matrix, that is, a correlation matrix constructed from mutually uncorrelated time series. If the properties of C agree with those of a random correlation matrix, then it follows that the contents of the empirically measured C are random. Conversely, deviations of the properties of C from those of a random correlation matrix convey information about “genuine” correlations (Plerou et al., 2002). While above we defined the way the empirical correlation matrix C is constructed, now we turn to define its random counterpart:

$$R = \frac{1}{T}AA^T \quad (4)$$

where R has the same dimension of C , containing random elements with zero mean and unit variance, that are mutually uncorrelated. By construction R belongs to the type of matrices often referred to as Wishart Matrices in multivariate statistics. The analysis of Wishart matrices is carried out on an eigenvalue basis, through the following decomposition:

$$R = V\Lambda V^{-1} \quad (5)$$

where the columns of V are represented by each eigenvector and the diagonal matrix Λ contains each eigenvalue associated to each eigenvector. An interesting feature of these random matrices is that their properties are known. In particular, in the limit ; $N \rightarrow \infty$, $T \rightarrow \infty$ such that $T/N > 1$ is fixed, the distribution $P_{rm}(\lambda)$ of eigenvalues λ of the random correlation matrix R is given by:

$$P_{rm} = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \quad (6)$$

This distribution is known as the Marčenko-Pastur Distribution. In the limiting case, eigenvalues λ are bounded in the interval $\lambda_- \leq \lambda_i \leq \lambda_+$ (minimum and maximum eigenvalues of R) which are given by:

$$\lambda_{\pm} = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}} \quad (7)$$

We then compare the eigenvalue distribution $P_c(\lambda)$ of C with $P_{rm}(\lambda)$. Plerou et al. (1999) and Laloux et al. (1999) find the presence of a well-defined “bulk” of eigenvalues of C that falls within the bounds $\lambda_- \leq \lambda_i \leq \lambda_+$

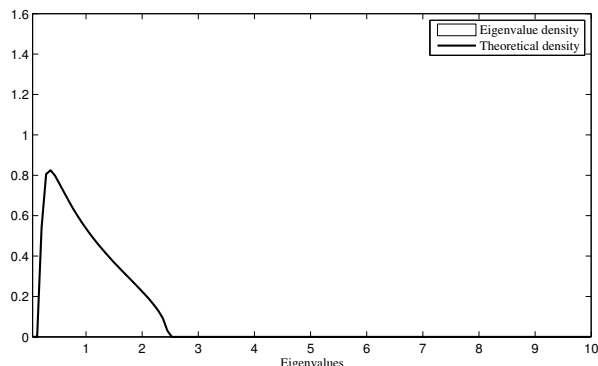


Figure 1: The Marčenko-Pastur Distribution

for $P_{rm}(\lambda)$. This in turn suggests that the contents of C that agree with $P_{rm}(\lambda)$ are mostly random except for the values that deviate. These latter eigenvalues would therefore be the ones that contain economically relevant information. There are two strands of literature on applications of the RMT to financial markets that were carried out over the same period. The first was Laloux et al. (1999, 2000) who applied the RMT approach to 406 time series returns belonging to the S&P 500, using daily data covering the period 1991-1996. The second was Plerou et al. (1999, 2000a,b, 2001, 2002), who used two datasets for the U.S. The first dataset consisted in 30-minute returns for 1000 US stocks with large market capitalization, for a 2-year period from 1994-1995 and 1996-1997. The second dataset contained daily returns for 422 U.S. stocks over 35 years, from 1962 to 1996. In both studies, the authors find that the “bulk” of the eigenvalues of the empirical correlation matrix well agree with the properties of the RMT, with the exception of a few deviating eigenvalues. Based on this latter, they develop a filter to retain only the information of the deviating eigenvalues and construct a *filtered* correlation matrix.⁴ They then construct a family of portfolios using the *filtered* and *unfiltered* correlation matrices. In both cases, the authors find

⁴The filtering approach applied by Plerou et al. (1999, 2000a,b); Gopikrishnan et al. (2001); Plerou et al. (2001, 2002) is slightly different to the one applied by Laloux et al. (1999, 2000) in that the former replaces the random eigenvalues by zeros, while the latter replaces the random eigenvalues by the average of the noisy eigenvalues. The consequences of these different filter approaches have been discussed in a number of studies (Wilcox and Gebbie, 2004, 2007).

that the forecast error is reduced, in some cases by a factor of 1.5.

Emerging markets have also been studied. For example, studies applying the RMT approach to emerging markets, namely, Nilantha et al. (2007) for Sri Lanka and Kulkarni and Deo (2007); Pan and Sinha (2007) for India, Medina and Mansilla (2008) for Mexico, find good agreement with the universal properties of the RMT, with fewer deviations. In particular, Pan and Sinha (2007), using 201 stocks of the National Stock Exchange (NSE), found fewer deviations from the bulk distribution than those found for the NYSE. The authors argue that this was due to the relative weakness of intra-sector interactions between stocks, compared to the market mode. Based on this, they suggested that the emergence of an internal structure of multiple groups of strongly correlated components is an indication of market development.

Wilcox and Gebbie (2004, 2007) analysed the Johannesburg stock exchange (JSE). In this case, they had to deal with difficulties of missing data and illiquid stocks. In the presence of such data, they still found good agreement with the RMT. For all data cleaning methods tested, the authors reported that most eigenvalues agreed with RMT, while there were a small number of large deviating values. They also noted that the choice of data cleaning method affected the results, with some introducing more noise than others. The portfolios considered here were also large, containing between 250 and 350 shares.

While these studies have been very successful in establishing the broad applicability of the RMT approach to a number of different economic context and financial markets, most of the existing literature using RMT to forecast realized risk has focused on relatively tranquil times in financial markets. One of the few recent studies focusing on more turbulent periods is that of Sandoval et al. (2012), who used the RMT approach to clean the correlation matrix of Brazilian stocks contained in the BOVESPA Index. They use this procedure to construct portfolios of stocks based on the Markowitz framework to build better predictions of future risk based on past data. They do this for years of low and high volatility for the Brazilian market during the years 2004-2010. They find that the *filtered* correlation matrix using the RMT approach is certainly better than the *unfiltered* one, but it fails to provide a better forecast during high volatility periods.

In this context, our study examines whether the predictions of the RMT also hold for the Chilean case. Following Sandoval et al. (2012) we look into low and high volatility periods to test the relative performance of the RMT approach under different volatility regimes. In addition, our study

contributes to the existing literature by exploring the macroeconomic determinants of the statistically significant eigenvalues, with the use of a Vector Autoregressive Model (VAR), in order to shed some light on what are the main drivers of systemic risk and their effects on portfolio management.

3. Eigenvalue Analysis for the Chilean Stock Market

In this section, to introduce our methodology we apply the random matrix approach to 83 records of Chilean stocks, using daily data for the year 2005.⁵ We therefore have $T = 260$ and $N = 83$, which implies $Q = T/N = 3.132$.

We then construct the correlation matrix C^{2005} based on the average time series returns and normalized variance equal to 1. Using equation 6, we can obtain the upper and lower bounds of the theoretical Marčenko-Pastur (MP) distribution. In this case, we have $\lambda_+ = 2.4492$ and $\lambda_- = 0.1892$.

Similarly, we generate a Wishart correlation matrix R composed by cross-correlations of mutually independent returns (bounded by the aforementioned edges). We then plot the MP density of the eigenvalues for 10000 trials and compare it with the eigenvalue density of the empirical correlation matrix C^{2005} . The results of this exercise are presented in figure 2.

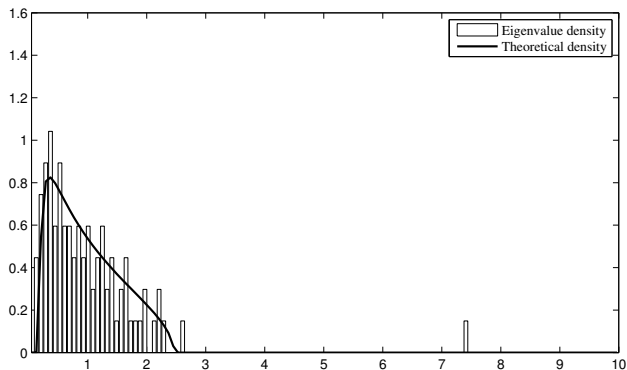


Figure 2: Histogram of eigenvalues (sample 2005) vs. The Marčenko-Pastur Distribution

⁵Throughout the paper, we will be using these 83 stocks for every year of the sample (2000-2010). The selection of these stocks was based on two criteria: (i) their availability over the entire sample, which allows to make reliable comparisons and (ii) were the most frequently traded.

As expected, the “bulk” of the eigenvalues lies within the bounds of the MP distribution. However, there are obvious deviations from the RMT predictions. The largest eigenvalue is well above the maximum predicted by the RMT predictions ($\lambda_+ = 2.4599$). This value as well as the remaining deviating eigenvalues, are the ones we expect to carry economically relevant information, while those within the distribution bounds can be considered just noise.

For 2005, there are two large deviations from the upper bound predicted by the RMT, $\lambda_1 = 7.4196$ and $\lambda_2 = 2.5942$. We also find three clear deviations in the lower bound of the MP distribution, $\lambda_{83} = 0.1419$, $\lambda_{82} = 0.1543$ and $\lambda_{81} = 0.1572$. From a portfolio management perspective, the list of eigenvalues can be seen as different levels of risk of a certain portfolio. Likewise, the eigenvector components associated to each eigenvalues can be regarded as the weights of the different stocks in a certain portfolio. In order to better clarify this point, in figure 3, we plot the components corresponding to the eigenvector 1 and 2 from the two largest eigenvalues.

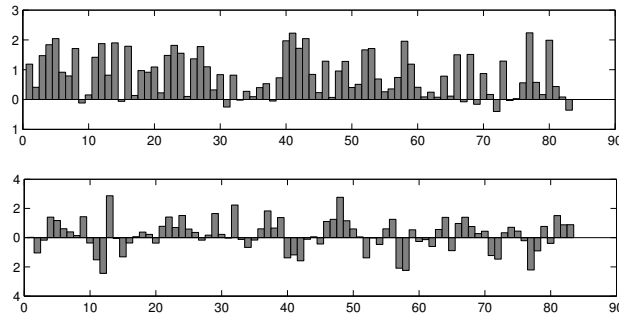


Figure 3: Eigenvector components of the two largest eigenvalues: Year 2005

In figure 3, regarding the eigenvector components associated to the largest eigenvalue, we can see that most of these components are positive and contribute in a similar way to the eigenvectors. This suggests a strong collective behavior of stocks. From a portfolio perspective, eigenvector components can be seen as the weights of the different stocks in a certain portfolio, where some stocks are *long* while others are *short*. In this line, the eigenvalue λ_1 represents the risk of the market portfolio. In general, large eigenvalues correspond to a mix of risky assets. In stock markets, the largest eigenvalue corresponds to the risk of a portfolio whose weights are the same for each

assets. There is no diversification on this portfolio: the only bet is whether the market as a whole goes up or down, this is why the risk is large. Conversely, if two stocks move very tightly together (e.g., Coca-Cola and Pepsi), then buying one and selling the other leads to a portfolio that hardly moves, being only sensitive to events that strongly differentiate the two companies (Bouchaud and Potters, 2009).

In order to prove this latter point, we construct a time series returns using the weights contained in the largest eigenvector u^1 , using the following expression:

$$G_t^1 = \sum_{j=1}^n u_j^1 G_j(t) \quad (8)$$

This basically defines the returns on the portfolio defined by u^1 . We then proceed to normalize these returns and run an OLS regression against the normalized returns of the market index G_t^{ipsa} , represented by the Indice de Precios Selectivo de Acciones (IPSA), which is a market index that measures the price variations of the 40 most traded stocks.

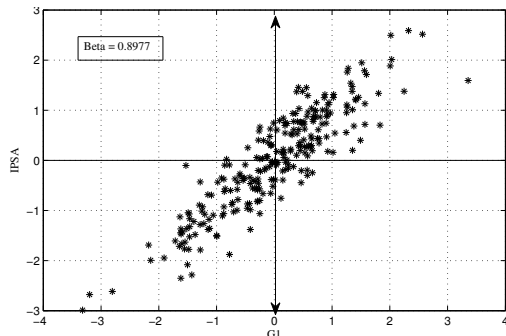


Figure 4: Scatter plot of the returns of G_t^1 and the returns of the IPSA

We plot the results of this regression in figure 4.⁶ We can see that there is a clear linear and positive relationship between the returns G_t^1 and G_t^{ipsa} , with a strong correlation coefficient $\langle G_t^{ipsa} G_t^1 \rangle = 0.8977$. From this, we can therefore conclude that the largest eigenvalue corresponds to the risk of the market mode. We repeat the same exercise for the smallest eigenvector u^{83} . We plot the correlation coefficients in figure 5.

⁶Due to their high correlation, using the IGPA would not change the results.

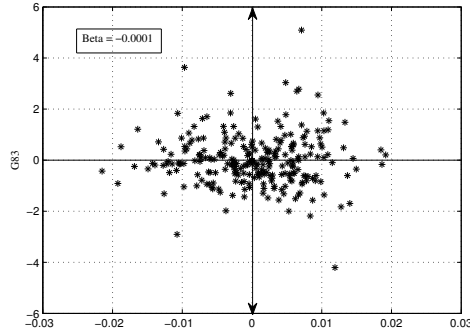


Figure 5: Scatter plot of the returns of G_t^{83} and the returns of the IPSA

From a portfolio perspective, the smallest eigenvalue determines the least risky portfolio one can build. This can be corroborated by the nearly zero beta coefficient of the regression against the market index. This in turn implies that the portfolio constructed from the smallest eigenvalues is almost totally immune to the systemic risk emanating from the collective behaviour of stocks represented by the largest eigenvalue.

Figure 6 plots the eigenvector components corresponding to the smallest eigenvalues, λ_{83} and λ_{82} . Here we can see that the eigenvector corresponding to the smallest eigenvalue exhibits clear preference for some pairs of stocks with the highest correlations. Finally, we plot the eigenvector components of two eigenvalues inside the bounds of the MP distribution in figure 7.

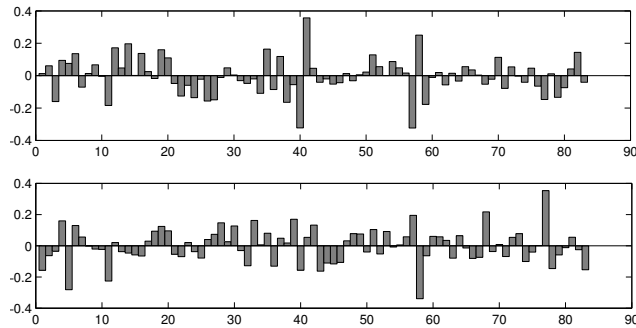


Figure 6: Eigenvector components of the two smallest eigenvalues: Year 2005

As we mentioned above, the eigenvalues inside of the bulk of the MP distribution are assumed to be random and therefore should carry no eco-

nomically relevant information. This appears to be the case, as the eigenvector components for both cases, do not seem to show major preferences for any stocks. Overall, our results are in line with the findings of most of the existing literature on the subject (Laloux et al., 1999, 2000; Plerou et al., 1999, 2000a,b; Gopikrishnan et al., 2001; Plerou et al., 2001, 2002; Bouchaud and Potters, 2009). In the next section, we look deeper into the significant eigenvalues and identify the individual stocks that participate in each of them.

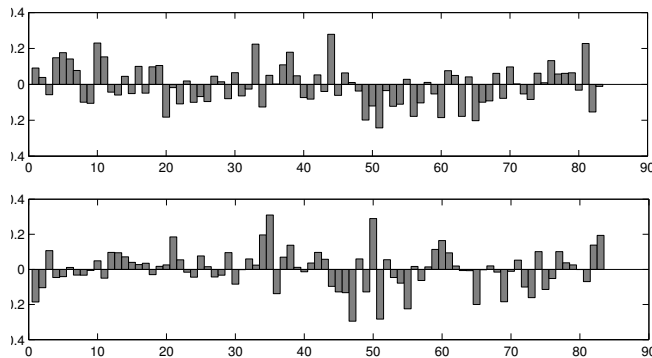


Figure 7: Eigenvector components of two eigenvalues inside of the bulk: Year 2005

4. The number of significant components in the deviating eigenvectors

Since the eigenvalues that are close to the bounds of the MP distribution are more likely to be random, as opposed to the largest eigenvalues which are clearly deviated from the upper edge of the spectrum, we quantify the number of components that participate significantly in each eigenvector, which in turn shows the level of deviation from the RMT for the distribution of eigenvector components. For this purpose, we use the notion of inverse participation ratio (IPR), often applied in localization theory. The IPR of the eigenvector u^k is defined as

$$I^k = \sum_{l=1}^N [u_l^k]^4 \quad (9)$$

where u_l^k , $l = 1, \dots, 83$ are the components of eigenvector u^k . The meaning of I^k can be illustrated by two limiting cases: (i) a vector with identical components $u_l^k = 1/\sqrt{N}$ has $I^k = 1/N$, whereas (ii) a vector with one component $u_1^k = 1$ and a remainder zero has $I^k = 1$. Thus, the IPR quantifies the reciprocal of the number of eigenvector components that contribute significantly. In our case, we have that the eigenvector associated to the largest eigenvalue has a $IPR1 = 0.0287$, which implies that we have $PR1 = 34.86 \approx 35$ significant stocks in the largest eigenvector. This is a very important result, which together with the high correlation of the portfolio returns built from the largest eigenvector with the market returns embedded in the IPSA index, is basically telling us that most of the volatility of the IPSA-IGPA index is driven by these 35 stocks. The significant firms in the largest eigenvector are presented in Table 6 in the appendix of this paper. By inspecting the largest eigenvector we can see that it has very strong components in the mining sector, with companies such as SQM-A, Oro Blanco and CalicheraA; energy sector, in companies such as Enersis and Endesa, as well as in the retail sector represented by Falabella. Also very important the contribution of the food and beverage industry, represented by firms such as Andina A and B, CCU, San Pedro and Conchatoro. In general, the largest eigenvector would seem to include most sectors of the Chilean stock market.

Regarding the rest of the deviating eigenvectors, the interpretation of their meaning is not as straightforward as in the former case. This is so because the largest eigenvalue is of order of magnitude larger than the others, hence constraining the remaining $N - 1$ since $TrC = N$. Therefore, in order to avoid the effect of the largest eigenvalue, we first need to remove it from the rest. For doing this, we regress, $G^1(t)$, calculated in the previous section, against each return and compute the residuals $\epsilon_i(t)$.⁷ Then, we calculate again the correlation matrix using $\epsilon_i(t)$ and compute the eigenvectors u^k of C . The number of significant participants in the second largest eigenvector for the year 2005, can be found in Table 6. In this case we have $PR2 \approx 15$ significant components with strong emphasis in the mining sector with companies such as CalicheraA, Oro Blanco and SQM-A/B, in the energy and combustibles sector, with companies like ECL, Colbun and COPEC as

⁷Alternatively, we could have used the returns on the largest eigenvector, $G^1(t)$ and the results would not change stem from the high correlation between both of an order of 0.89.

well as in the forest and agro-industrial sector with companies like CMPC, IANSA and Campos. Also, the removal of the first component allowed to identify a third deviating eigenvalue in the upper bound, with $PR3 \approx 28$, with again strong weights on the mining, energy and banking and investment sectors.

Finally, the smallest eigenvectors, each has $PR83 \approx 24$ and $PR82 \approx 24$ significant participants. Recall that these eigenvectors determine the least risky portfolios of stocks one can build when using the Markowitz Portfolio Theory (MPT). Typically, it has been found in the literature that the components in the smallest eigenvectors represent a very small number of stocks with large correlations.⁸ As we can see in Table 6, this is not the case for the Chilean stock market. For the smallest eigenvector, u^{83} , although, we do find strong negative correlations between some pairs of stocks, instead of finding a small portfolio we rather find a quite diversified portfolio of stocks composed by a mix of sectors, with strong preferences for stocks in the food and beverage industry, metal-mechanic industry, investment and real estate and to a lower extent mining. The portfolio composed by these stocks has a correlation with the IPSA, $\langle G_{83}^{2005} G^{ipsa} \rangle = -0.0081$. Regarding the second smallest eigenvalue (not shown), u^{82} , we notice it has strong weights again in the food and beverage, mining, energy sectors and to a lower extent on public services and private pension funds (AFPs). Therefore, the results for the Chilean equity market are different to what has been found in the literature for the U.S, where the portfolio defined by the smallest eigenvector contain very few stocks with very large negative correlations. A similar phenomenon has nonetheless been found in recent literature on emerging markets (Pan and Sinha, 2007), where there is a tendency for relatively weak intra-sector interactions between stocks. In this context, the emergence of an internal structure of multiple groups of strongly correlated components would be sign of market development.

The aforementioned results are very interesting from an asset allocation perspective, where the use of the RMT approach allows us to identify groups of stocks that can be used to construct portfolios with different degrees of exposure to the systemic risk embedded in the market index.

In the next section, we will apply these notions in the context of Markowitz's

⁸See Laloux et al. (1999, 2000); Plerou et al. (2000a); Gopikrishnan et al. (2001); Plerou et al. (2002); Bouchaud and Potters (2009).

portfolio theory and show that the RMT approach can be helpful for constructing minimum variance portfolios, far superior than using standard correlation matrix.

5. Portfolio Analysis for the Chilean Stock Market

The underlying randomness contained in the “bulk” of eigenvalues seen in the previous section has important implications in optimal portfolio selection. In the context of a Markowitz’s optimal portfolio theory, the effect of noise has a strong weight on the smallest eigenvalues, which are precisely those that determine the least risky portfolios. This is why the Random Matrix Approach allows reconciling the mean-variance approach of Markowitz as a tool for providing better risk predictions. Consider a portfolio $\Pi(t)$ of stocks with prices S_i . The return on $\Pi(t)$ is given by

$$\Phi = \sum_{i=1}^N w_i G_i \quad (10)$$

where $G_i(t)$ is the return on stock i and w_i is the fraction of wealth invested in stock i . The fractions w_i are normalized such that $\sum_{i=1}^N w_i = 1$. The risk of holding the portfolio $\Pi(t)$ can be quantified by the variance

$$\Omega^2 = \sum_{i=1}^N \sum_{j=1}^N w_i C_{ij} w_j \sigma_i \sigma_j \quad (11)$$

where σ_i is the standard deviation (average volatility) of G_i , and C_{ij} are elements of the cross-correlation matrix C . In order to find an optimal portfolio, we must minimize Ω^2 under the constraint that the return on the portfolio is some fixed value Φ . In addition, we also have the constraint that $\sum_{i=1}^N w_i = 1$. The resolution of this minimization can be implemented by using two Lagrange multipliers, which leads to a system of linear equations for w_i , which can then be solved. The family of optimal portfolios therefore can be represented as a plot of the return Φ as a function of risk Ω^2 , the so-called, *efficient frontier*.

To find the effect of randomness of C on the selected optimal portfolio, we carry out our analysis for the years 2005-2006. We first calculate the cross-correlation matrix C_{2005} and using the returns G_i for 2006, we construct a family of 100 optimal portfolios, and plot Φ as a function of the predicted risk Ω_p^2 for 2006. For this family of portfolios, we also compute the realized risk

Ω_r^2 during 2006 using C_{2006} . We do this in order to compare the performance of the RMT approach isolated from any additional source of noise that could arise, for example, in case we had to make a forecast of returns.

Since the meaningful information in C is contained in the deviating eigenvectors (whose eigenvalues are outside of the RMT bounds), we use the filtered approach carried out in the study of Plerou et al. (2002). They construct the filtered matrix C^f , by retaining only the deviating eigenvectors, using the eigenvector decomposition $C = V\Lambda V^{-1}$, where V contains the eigenvectors and Λ is a diagonal matrix containing the eigenvalues. Plerou et al. (2002) construct the filtered matrix C^f , by setting a new diagonal matrix Λ^f , with elements $\Lambda_{ii}^f = (0, \dots, 0, 0, \dots, 0, \lambda_2, \lambda_1)$. Then they transform Λ^f back to the basis of C , resulting in the *filtered* cross-correlation matrix C^f .⁹ In addition, they set the diagonal elements $C_{ii}^f = 1$, to preserve $Tr(C) = Tr(C^f) = N$. We use this filtering technique and we repeat the above approach for finding the optimal portfolio using C^f instead of C . We plot the efficient frontiers for both approaches, for the *filtered* and *unfiltered* matrices, in figure 8.¹⁰

In order to measure the relative performance of our *filtered* portfolios vs. the *unfiltered* ones, we calculate the difference between both curves by using the Mean Average Percentage Error (MAPE), defined as:

$$MAPE = \frac{1}{n} \sum_{i=1}^N \left| \frac{\Omega_i^2 r - \Omega_i^2 p}{\Omega_i^2 p} \right| \quad (12)$$

In this case, we have that the *unfiltered* approach, which uses the standard covariance matrix, underestimates the realized risk in 28%, while when using the RMT approach it only underestimates the realized risk in 16%. There is certainly a sharp contrast between the use of the *filtered* and *unfiltered* matrices, with the *filtered* matrix performing far better than the *unfiltered* matrix for forecasting portfolio risk.

⁹There are a number of reasons why the smallest eigenvalues are not considered in this filter. The most relevant is that while the two largest eigenvalues are clearly separated from the MP bounds, the same does not always applies for the smallest eigenvalues. In general, the existence of small eigenvalues outside the bounds is consistent with the fact that N and T are finite in practice.

¹⁰For this example, we also allow for “short-sale”. Although, “short-selling” is not always possible as increases the risk of the portfolio.

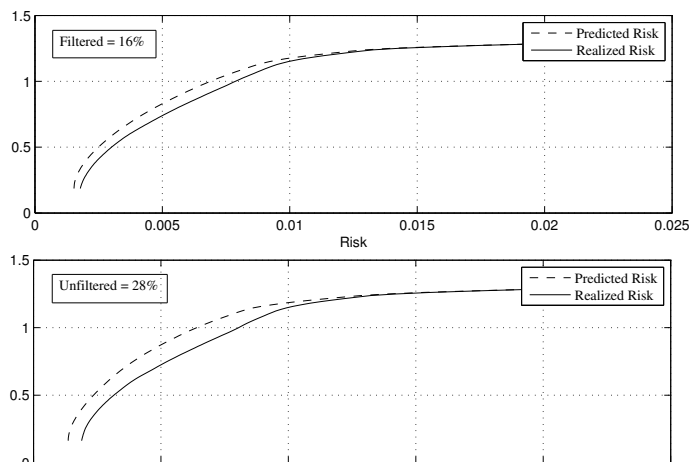


Figure 8: Efficient Frontier during 2006

In this context, also an interesting question would be to assess how our *filtered* approach performs relative to the *unfiltered* one, during different volatility regimes in the stock market. To do this, we apply the methodologies presented in sections 2, 3 and 4 to consecutive one-year periods from 2000-2001 to 2009-2010. In columns (1)-(2) we compare the MAPE using the *filtered* and *unfiltered* approach, not allowing for “short sale”, while in columns (3)-(4) we compare again both approaches, but this time allowing for “short-sale”. Finally, in columns (5) and (6) we present the volatility of the IPSA calculated as the mean of the absolute value of returns for the *predicted* year and the VIX index as a measure of international market volatility. The results are presented in Table 1.

By comparing columns (1) and (2) we clearly see that in most cases the *filtered* approach performs far better than the *unfiltered*, with only one exception; when using the correlation matrix of 2007 to predict the correlation matrix of 2008. During 2007 the burst of the housing bubble in the U.S led to the start of a crisis whose zenith was reached during 2008. This result could be an indication that the RMT approach may fail to predict the correlation matrix during periods of high volatility.

Nonetheless, the results for 2007-2008 are reversed when we allow for “short- sale”. From columns (3) and (4), we can see that the *filtered* approach is far superior to the *unfiltered* one in all cases. “Short-sale” involves selling assets that are borrowed in expectation of a fall in the asset’s price. In case the price declines, the investor covers his position by buying an equiv-

Previous-Predicted	No short sale		Short sale		Vol. IPSA	VIX
	Unfiltered	Filtered	Unfiltered	Filtered		
	(1)	(2)	(3)	(4)	(5)	(6)
2000-2001	0.307	0.100	0.345	0.066	0.745	23.3
2001-2002	0.029	0.027	0.344	0.096	0.615	25.7
2002-2003	0.238	0.189	0.458	0.206	0.708	27.2
2003-2004	0.255	0.220	0.735	0.375	0.695	21.9
2004-2005	0.058	0.019	0.428	0.108	0.567	15.4
2005-2006	0.288	0.163	0.412	0.179	0.581	12.7
2006-2007	0.333	0.276	0.487	0.169	0.545	12.7
2007-2008	0.073	0.130	0.367	0.173	0.888	17.4
2008-2009	0.127	0.041	0.367	0.192	1.168	32.5
2009-2010	0.052	0.023	0.379	0.186	0.754	31.4
Average	0.176	0.119	0.439	0.175	0.727	22.07
Improvement	0.481		3.71			

Table 1: Mean Average Percentage Error for each year of the sample

alent number of assets at the new lower price and returns to the lender the assets that were borrowed. In mathematical terms, when we allow for short sale, portfolio weights are allowed to be negative, and investors can continuously sell low return assets and reinvest in higher yield assets and generate an infinite expected return. Conversely, the investor could “short-sell” a high yield asset and reinvest the proceeds into a low yield asset, therefore achieving an infinite negative return. In this context, “short-selling” would seem to expand the opportunity set for investment decisions. The use of the RMT allows us to exploit these opportunities to construct portfolios that can be useful when they are most needed, this is during high volatility periods. However, “short-sale” is not always allowed in financial markets due to the high volatility that brings to portfolios. Regardless of these potential shortcomings, these results are rather encouraging from an institutional investor’s perspective.

Our results are different from those of Sandoval et al. (2012) for the Brazilian stock market, who finds that the RMT approach fails to predict the correlation matrix in scenarios of high volatility. They apply the RMT, with and without “short-sale” and find negative results for both cases. In particular, they find that the RMT approach performs relatively well, de-

pending on the volatility of the forecasted year being smaller or larger than the year used for the forecast. While this would make sense from a statistical point of view, we only find this phenomenon in one case in our sample; when using the correlation matrix of 2007 to predict the one of 2008, with “short-sale” constraints. Conversely, when “short-sale” is allowed, the *filtered* approach outperforms the *unfiltered* one in all cases.

6. Macroeconomic Determinants of the Chilean Stock Market Returns

In the previous section, we applied the RMT to the Chilean stock market and obtained optimal portfolios. These portfolios show improved efficiency compared to a benchmark of standard Markowitz portfolios. In this section, we now turn into the macroeconomic and financial determinants of the Chilean stock market returns. The question we try to answer here is which are the main macroeconomic and financial factors affecting the Chilean stock market returns? To answer this question we rely on Vector Autorregression analysis (VAR). Our main dependent variables are the monthly returns of optimal portfolios created using the 3 largest and the 3 smallest statistically significant eigenvalues with equation 12 in section 3 for the whole sample.¹¹ We denominate the portfolio returns as G_t^1 , G_t^2 and G_t^3 for the returns of the portfolios constructed using the first, second and third largest eigenvalues and G_t^{81} , G_t^{82} and G_t^{83} for the returns of the portfolios constructed using the first, second and third smallest eigenvalues.

To account for macroeconomic factors affecting the returns, we include 3 domestic macroeconomic variables in our estimations: the seasonally adjusted monthly change of industrial production (x_t^{IMACEC}), the seasonally adjusted monthly inflation (π_t) and the monetary policy rate (r_t).¹² Representing more than 60% of Chilean exports, copper prices are likely to be one of the main channels through which international shocks translate to

¹¹Monthly portfolio returns are constructed based on the same analysis of section 3, but using the significant eigenvectors calculated for the whole sample. This resulted in 6 deviating eigenvalues; 3 deviating on the upper edge and 3 deviating eigenvalues on the lower edge of the spectrum. We constructed returns based on their corresponding eigenvectors.

¹²We treated the seasonality of time-series using the X12 methodology for seasonally adjusted time-series.

the Chilean economy. We include a variable with the monthly change of the price of copper (P_t^{copper}) to account for this effect. Finally, to account for international financial factors we include the VIX . This variable is commonly used as a gauge for international stock market volatility and investors risk aversion.

We estimate an unrestricted VAR with the following structure:

$$VAR(L) : \left(VIX, P_t^{copper}, x_t^{IMACEC}, \pi_t, r_t, G_t^1, G_t^2, G_t^3, G_t^{81}, G_t^{82}, G_t^{83} \right) \quad (13)$$

Producing impulse responses from this econometric model requires an identification strategy. To deal with this issue, we follow Pesaran and Shin (1998) and use a generalized impulse response function for an unrestricted VAR. This identification method has the advantage over the Choleski decomposition in that it produces impulse responses, which are invariant to the ordering of the variables in the system.

We run a battery of tests to get the optimal number of lags for our VAR (Table 2). The majority of the tests recommend the use of 1 lag in the regressions.

0	-641.522	NA	1.94e-09	11.154	11.413	11.259
1	-265.279	675.309	2.49e-11*	6.791*	9.907*	8.056*
2	-177.902	140.400	4.64e-11	7.365	13.338	9.790
3	-69.583	153.682*	6.58e-11	7.582	16.412	11.167
4	28.544	120.773	1.27e-10	7.973	19.659	12.718
5	150.565	127.235	2.02e-10	7.956	22.498	13.860
6	296.965	125.128	2.90e-10	7.521	24.921	14.585

Table 2: VAR Lag Order Selection Criteria Notes: *indicates lag order selected by the criterion; LR: sequential modified LR test statistic (each test at 5% level); FPE: Final prediction error; FPE: Final prediction error; AIC: Akaike information criterion; SC: Schwarz information criterion; HQ: Hannan-Quinn information criterion; Endogenous variables: $G_t^1, G_t^2, G_t^3, G_t^{83}, G_t^{82}, G_t^{81}, x_t^{IMACEC}, \pi_t, r_t, P_t^{copper}$ and VIX . Sample: 2000M10 2011M1.

Since we are mainly interested in the effects of macroeconomic and financial variables on the returns of our portfolios we focus on this set of impulse responses. Figure 9 presents the impulse responses for the VAR(1).

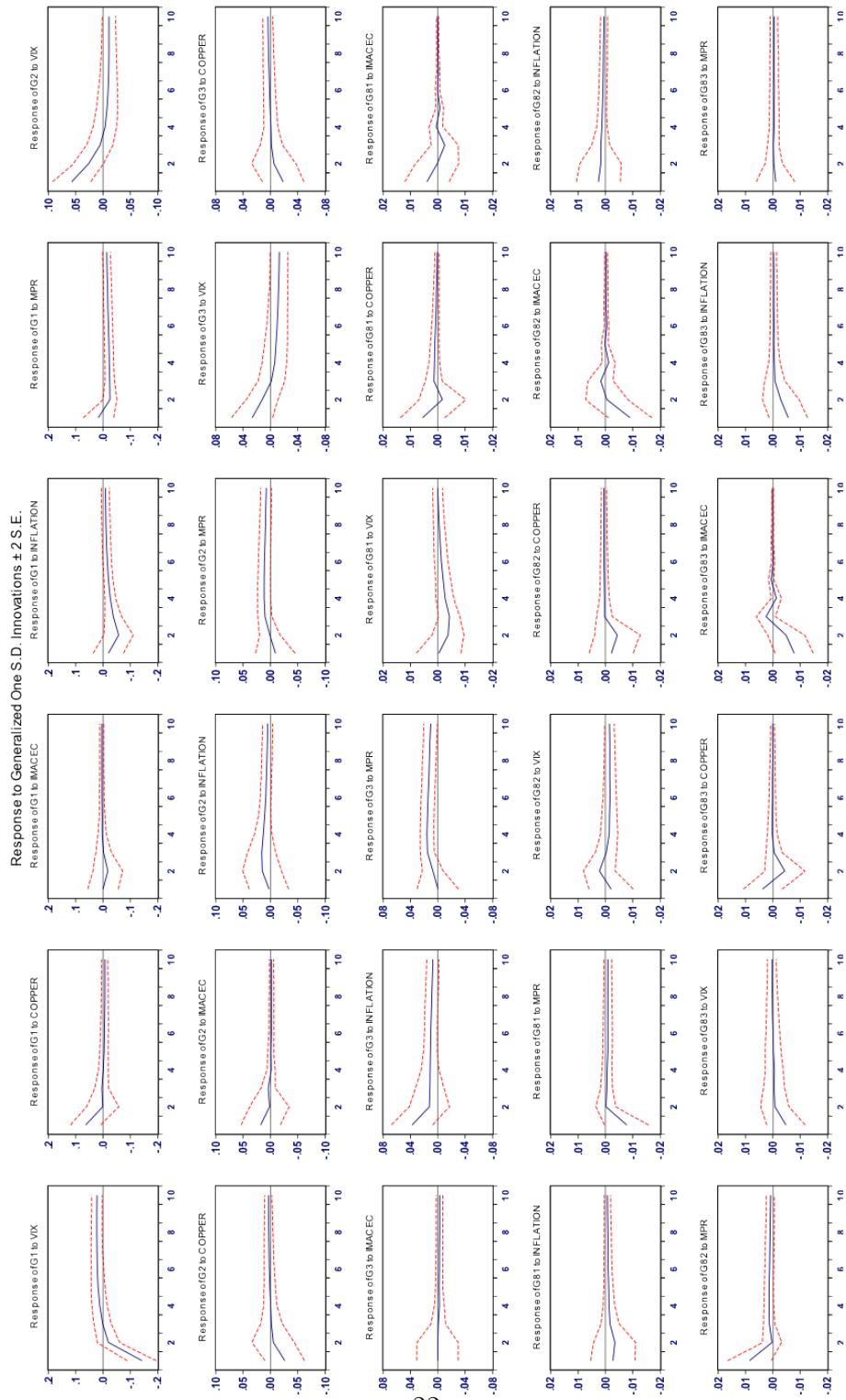


Figure 9: Generalized Impulse Responses

The impulse response functions suggest that shocks to the macroeconomic variables have important and varied effects on the portfolio returns:

- VIX : A positive shock to the VIX (an increase of the international volatility) has a transitory negative effect on the returns of the portfolio constructed using the first largest eigenvalue and a transitory positive effect on the returns of the portfolio constructed with the second largest eigenvalue. The negative impact on returns happens during the first and second months for G_t^1 . The positive effects on G_t^1 also happen during the first and second months. On the other hand, unexpected shocks to the VIX have no effects on the portfolio returns constructed with the smallest eigenvalues.
- P_t^{copper} : A shock to the monthly change of the price of copper has a positive effect on G_t^1 but no effect on the remaining portfolios.
- x_t^{IMACEC} : A positive shock to the economic dynamism measured as the monthly change of x_t^{IMACEC} , does not have any impact on the returns of the portfolios constructed with the 3 largest eigenvalues nor the returns of the portfolio constructed with the third smallest eigenvalue. However, it has a transitory negative impact on the returns of the portfolios constructed with the first and second smallest eigenvalues. These effects occur during the 1st month of the shock.
- π_t : A positive shock on the seasonally adjusted monthly inflation has a transitory negative impact on the returns of the portfolio constructed with the largest eigenvalue and a transitory positive one on the returns of the portfolio constructed with the third largest eigenvalue. On the other hand, inflation shocks would not seem to have an impact on G_t^{81} , G_t^{82} and G_t^{83} .
- r_t : A positive shock to the monetary policy rate (an unexpected increase of the r_t) has a transitory negative impact on the returns of the portfolio constructed with the largest eigenvalue and a transitory positive one on the returns of the portfolio constructed with the third largest and the second smallest eigenvalues.

To better understand the relative importance of each macroeconomic shock on the portfolio returns we carry out a Variance Decomposition Analysis. To do this, we need a different identification strategy for the residuals.

G^1	$S.E.$	VIX	P_t^{copper}	x_t^{JMACEC}	π_t	r_t	G_t^1	G_t^2	G_t^3	G_t^{81}	G_t^{82}	G_t^{83}
3 Months	0.323	19.653 (6.191)	0.163 (1.269)	0.450 (1.568)	4.334 (3.015)	1.006 (1.357)	71.455 (7.213)	0.117 (1.267)	0.762 (1.55761)	1.280 (1.592)	0.256 (1.400)	0.518 (1.815)
12 Months	0.336	21.375 (5.735)	0.198 (1.225)	0.436 (1.483)	6.323 (3.986)	2.529 (1.978)	65.805 (7.052)	0.132 (1.207)	0.858 (1.581)	1.369 (1.483)	0.491 (1.340)	0.478 (1.628)
G^2	$S.E.$	VIX	P_t^{copper}	x_t^{JMACEC}	π_t	r_t	G_t^1	G_t^2	G_t^3	G_t^{81}	G_t^{82}	G_t^{83}
3 Months	0.213	8.799 (4.958)	0.100 (1.704)	0.317 (1.505)	0.909 (1.986)	0.328 (1.381)	38.605 (6.027)	50.031 (7.077)	0.214 (2.079)	0.652 (1.257)	0.001 (0.972)	0.039 (1.145)
12 Months	0.218	10.169 (5.208)	0.131 (1.647)	0.317 (1.443)	2.321 (2.964)	1.404 (1.739)	36.810 (5.749)	47.622 (7.046)	0.307 (2.129)	0.728 (1.195)	0.147 (0.948)	0.038 (1.099)
G^3	$S.E.$	VIX	P_t^{copper}	x_t^{JMACEC}	π_t	r_t	G_t^1	G_t^2	G_t^3	G_t^{81}	G_t^{82}	G_t^{83}
3 Months	0.177	2.694 (3.085)	0.287 (1.610)	0.031 (1.435)	4.341 (3.849)	0.878 (0.968)	15.459 (5.496)	0.781 (1.676)	75.204 (6.429)	0.273 (1.307)	0.040 (1.068)	0.006 (0.993)
12 Months	0.186	5.684 (4.791)	0.401 (1.500)	0.072 (1.347)	6.568 (4.720)	3.690 (1.943)	13.961 (5.047)	0.730 (1.489)	68.072 (7.243)	0.468 (1.191)	0.335 (1.170)	0.015 (1.019)
G^{81}	$S.E.$	VIX	P_t^{copper}	x_t^{JMACEC}	π_t	r_t	G_t^1	G_t^2	G_t^3	G_t^{81}	G_t^{82}	G_t^{83}
3 Months	0.046	1.477 (2.320)	2.117 (2.712)	0.943 (2.091)	0.797 (1.964)	1.998 (2.541)	1.594 (2.222)	1.524 (2.638)	1.060 (1.767)	87.205 (6.044)	0.019 (1.010)	1.259 (2.112)
12 Months	0.046	2.084 (4.003)	2.120 (2.651)	0.987 (2.145)	1.041 (2.205)	2.1487 (2.596)	1.620 (2.174)	1.531 (2.576)	1.060 (1.727)	86.122 (6.820)	0.033 (1.024)	1.249 (2.072)
G^{82}	$S.E.$	VIX	P_t^{copper}	x_t^{JMACEC}	π_t	r_t	G_t^1	G_t^2	G_t^3	G_t^{81}	G_t^{82}	G_t^{83}
3 Months	0.045 (1.405)	0.452 (2.423)	1.196 (3.143)	3.895 (2.123)	0.573 (2.163)	2.718 (2.859)	2.504 (2.782)	1.852 (2.449)	1.346 (2.436)	3.222 (6.853)	81.552 (1.866)	0.684
12 Months	0.046	1.397 (2.301)	1.179 (2.344)	3.884 (3.121)	0.9735 (2.232)	2.997 (2.164)	2.468 (2.751)	1.824 (2.710)	1.377 (2.417)	3.212 (2.342)	80.011 (7.143)	0.671 (1.822)
G^{83}	$S.E.$	VIX	P_t^{copper}	x_t^{JMACEC}	π_t	r_t	G_t^1	G_t^2	G_t^3	G_t^{81}	G_t^{82}	G_t^{83}
3 Months	0.040	1.473 (1.984)	1.870 (2.142)	5.160 (3.955)	1.985 (2.331)	0.016 (0.888)	2.150 (2.399)	2.098 (2.073)	1.932 (2.498)	1.858 (2.402)	0.410 (1.630)	81.043 (6.743)
12 Months	0.040	1.521 (2.420)	1.869 (2.096)	5.259 (3.937)	2.067 (2.369)	0.105 (1.029)	2.150 (2.363)	2.093 (2.018)	1.932 (2.463)	1.856 (2.355)	0.433 (1.586)	80.710 (6.936)

Table 3: Variance Decomposition of Portfolio Returns.

We opted for the Choleski decomposition with the ordering presented in equation 13. The international volatility (VIX) is the most exogenous variable, followed by the copper price (P_t^{copper}). Then, we included the domestic macroeconomic variables beginning with the x_t^{IMACEC} , π_t and r_t . Finally, we include our monthly portfolio returns ordered from the one constructed by largest eigenvalue to the portfolio return constructed by the smallest eigenvalue. Table 3 presents the results of this exercise.

Our results present important implications for risk diversification across Chilean equities. First, the main macroeconomic determinant of the market portfolio volatility (G_t^1) is the international volatility (VIX). This variable explains more than 20% of the market portfolio returns variance after 12 months. Domestic inflation followed by the monetary policy rate (r_t) are also important macroeconomic determinants explaining 6.3% and 2.5% of the market portfolio returns variance after 12 months, respectively. Second, it would be possible to diversify away some of the market portfolio risk by adding positions on the portfolios constructed with the second and third largest eigenvalues. The VIX explains 10% of the returns variance of G_t^2 and 5.7% of the return variance of G_t^3 . But, the impact is of different sign than for the market portfolio returns. Third, as expected, the three smallest eigenvalues produce portfolio returns that are mostly uncorrelated with macroeconomic shocks. These portfolios are also uncorrelated with the market portfolio.

In sum, the market portfolio (which is very close to the IPSA-IPGA indices) is mostly affected by international shocks and therefore should be used by investors willing to increase their exposure to global risks or when they are expecting a decrease in international volatility. The portfolios constructed with the second and third largest eigenvalues, on the other hand, could be used by investors to increase exposure to global risks but when they are expecting an increase on global volatility (due to their positive correlation with the VIX). Finally, as theoretically expected, the portfolios constructed using the smallest eigenvalues are uncorrelated to macroeconomic variables and global risk aversion.

In the next section, we will try to separate out diversifiable from non-diversifiable risk with the use of a simple one-factor model, to extract the market effect and assess the performance of the RMT approach for the filtered and unfiltered matrices.

7. Improving our results by getting rid of the non-diversifiable Risk

In this section we extend our previous results by removing the effect of the market portfolio from our time series returns and verifying how the RMT filter performs under this new scenario. In the previous analysis we were not able to fully distinguish between diversifiable and non-diversifiable risk. Diversifiable risk is directly related to correlations between stocks. While non-diversifiable risk is mainly driven by market momentum and/or breaking news that may affect all stocks at the same time and inducing new correlations that could not be distinguished from the real correlations between stocks. This is particularly true, during drawdowns. To remove the market effect, we run an OLS regression of the returns of the IPSA index on each time series returns, retrieve the residuals of that regression and build again our correlation matrix. This new correlation matrix would be to some extent immune from the market effect. Recent emerging markets literature has made use of this approach to distinguish between co-movements in financial assets from “pure contagion” (Bunda et al., 2009).¹³

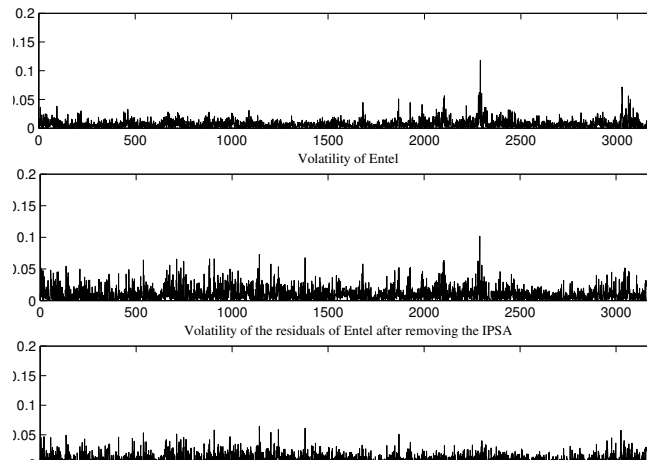


Figure 10: Volatility after removing the market mode

In figure 10, we plot the volatility of the IPSA index, measured as the absolute value of its returns, together with the volatility of a Telecommuni-

¹³Alternatively, we could have used the time series built from the largest eigenvector, with similar results given the high correlation between the largest eigenvector returns and the IPSA-IGPA indices.

Previous-Predicted	No short sale		Short sale		Vol. IPSA	VIX
	Unfiltered	Filtered	Unfiltered	Filtered		
	(1)	(2)	(3)	(4)	(5)	(6)
2000-2001	0.319	0.083	0.347	0.076	0.745	23.3
2001-2002	0.031	0.012	0.364	0.089	0.615	25.7
2002-2003	0.229	0.145	0.454	0.201	0.708	27.2
2003-2004	0.240	0.219	0.706	0.391	0.695	21.9
2004-2005	0.049	0.026	0.422	0.131	0.567	15.4
2005-2006	0.308	0.090	0.450	0.154	0.581	12.7
2006-2007	0.420	0.241	0.527	0.210	0.545	12.7
2007-2008	0.105	0.184	0.443	0.227	0.888	17.4
2008-2009	0.109	0.061	0.377	0.040	1.168	32.5
2009-2010	0.091	0.017	0.418	0.139	0.754	31.4
Average	0.190	0.108	0.451	0.166	0.727	22.07
Improvement	0.763		3.71			

Table 4: Mean Average Percentage Error for each year of the sample without the market mode.

cation Company (ENTEL), while in the third plot we include the volatility of ENTEL after removing the IPSA index. We can see that most peaks in the time series are now smoother after removing the effects of the market. Table 4 presents the performance of the RMT approach, for each year of the sample when the market effect is removed. As we can see by observing columns (7) and (8), again the *filtered* approach outperforms the *unfiltered* one in every year with the exception of 2007-2008. As in the case without removing the non-diversifiable risk, when we allow for “short-sale”, the *filtered* approach performs unequivocally better than the *unfiltered* one. It would seem that the removal of the market mode allow us to improve our risk estimates compared to the risk estimates of the previous section. Overall, these results are encouraging since it tells us that we can use the RMT approach also during periods of high volatility, when good investment strategies are most needed.

This is particularly true, when investors can exploit a full opportunity set, giving room for a “short-sale” strategy. However, “short-sale” is not always possible in financial markets. Despite this potential constraint, the use of the RMT approach could be used for medium/long term investment

decisions, as oppose to active strategies that try to “beat the market” second by second, implying greater chances of large losses. This is quite evident when observing the outstanding performance of the strategy suggested by the RMT approach in horizons greater than one year.

8. Conclusion

In this paper we apply Random Matrix Theory to study the correlations of 83 Chilean stocks that are part of the IPSA and IGPA indices during the period 2000 to 2011. We find that using Random Matrix Theory to identify statistically significant correlations within our sample of stocks significantly improves the efficiency of a family of Markowitz Portfolios.

Moreover, by using Vector Autoregressive analysis we identify the global risk aversion as the main driver of the market returns followed in importance by shocks to the monthly rate of inflation and the country’s monetary policy rate. With the use of statistically significant eigenvalues we can construct portfolios, which are uncorrelated to macroeconomic and financial factors such as global risk aversion.

Despite the recent success of Random Matrix Theory in financial applications, challenges remain. Recent evidence has suggested that meaningful correlations can be measured in the bulk of the MP eigenvalue spectrum (Burda et al., 2004; Burda and Jurkiewicz, 2004; Malevergne and Sornette, 2004; Kwapien et al., 2006). One possible reason for this is the fact that we are assuming that our returns are normally distributed and under this premise we carry out our comparisons with the empirical correlation matrix. In reality, returns in financial markets are not normally distributed and exhibit a number of features, such as fat tails, volatility clustering and non-stationarity. Recent literature has been improving these matters by developing extensions of the MP distribution to account for these phenomena (Potters et al., 2005; Bouchaud and Potters, 2009). However, still one of the main difficulties has been to generate predictions during critical periods, where there is a high collectivity between stocks, especially during draw-downs. In this context, an interesting line of future research is to extend this approach focusing on turbulent periods in financial markets. Another interesting line of future research would be extending our analysis to regional stocks such as the MILA or the MSCI LATAM. Also, it would be interesting to apply RMT to cross assets within the Chilean financial market including the return of equities, fixed income and currency. This would give us a more

complete picture of the factors that account for the volatility of the Chilean financial market.

All in all, the results obtained for the Chilean case provide important evidence of the broad applicability of the RMT approach to equity returns in emerging markets portfolios. The Chilean stocks market is characterized by the existence of large institutional investors such as private pension funds that maintain important structural positions on this market. Applying RMT to their portfolio optimization could be useful in diversifying some of the systemic risk of the Chilean market.

References

- Bouchaud, J. P., Potters, M., Oct. 2009. Financial Applications of Random Matrix Theory: a short review.
- Bunda, I., Hamann, A. J., Lall, S., 2009. Correlations in emerging market bonds: The role of local and global factors. *Emerging Markets Review* 10 (2), 67 – 96.
- Burda, Z., Görlich, A., Jarosz, A., Jurkiewicz, J., 2004. Signal and noise in correlation matrix. *Physica A: Statistical Mechanics and its Applications* 343 (0), 295 – 310.
- Burda, Z., Jurkiewicz, J., 2004. Signal and noise in financial correlation matrices. *Physica A: Statistical Mechanics and its Applications* 344, 67 – 72, applications of Physics in Financial Analysis 4 (APFA4).
- Gopikrishnan, P., Rosenow, B., Plerou, V., Stanley, H. E., 2001. Quantifying and interpreting collective behavior in financial markets. *Physical Review E* 64, 035106.
- Kulkarni, V., Deo, N., 2007. Correlation and volatility of an Indian stock market: A random matrix approach. *The European Physical Journal B* 60.
- Kwapien, J., Drożdż, S., Oświecimka, P., 2006. The bulk of the stock market correlation matrix is not pure noise. *Physica A: Statistical Mechanics and its Applications* 359 (0), 589 – 606.

- Laloux, L., Cizeau, P., Bouchaud, J. P., Potters, M., Aug. 1999. Noise Dressing of Financial Correlation Matrices. *Physical Review Letters* 83 (7), 1467–1470.
- Laloux, L., Cizeau, P., Potters, M., Bouchaud, J. P., 2000. Random matrix theory and financial correlations. *Int. J. Theor. Appl. Finance* 3(3), 391–397.
- Malevergne, Y., Sornette, D., 2004. Collective origin of the coexistence of apparent random matrix theory noise and of factors in large sample correlation matrices. *Physica A: Statistical Mechanics and its Applications* 331, 660–668.
- Medina, M., Mansilla, R., 2008. Teoria de matrices aleatorias y correlacion de series financieras: El caso de la bolsa mexicana de valores. *Revista de Administracion, Finanzas y Economia* 2(2), 125–135.
- Nilantha, K., Ranasinghe, Malmini, P., 2007. Eigenvalue density of cross-correlations in sri lankan financial market. *Physica A: Statistical Mechanics and its Applications* 378 (2), 345 – 356.
- Pan, R., Sinha, S., 2007. Collective behavior of stock price movements in an emerging market. *Physical Review E* 76 (4), 046116.
- Pesaran, H., Shin, S., 1998. Generalized impulse response analysis in multivariate models. *Economic Letters* 58, 17–29.
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., Stanley, H., 2000a. A random matrix theory approach to financial cross-correlations. *Physica A: Statistical Mechanics and its Applications* 287, 374 – 382.
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., Stanley, H., 2001. Collective behavior of stock price movements: a random matrix theory approach. *Physica A: Statistical Mechanics and its Applications* 299, 175 – 180.
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A., Stanley, H., 2000b. Econophysics: financial time series from a statistical physics point of view. *Physica A: Statistical Mechanics and its Applications* 279, 443 – 456.

- Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L. A., Guhr, T., Stanley, H. E., Jun. 2002. Random matrix approach to cross correlations in financial data. *Physical Review E* 65 (6).
- Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L. A., Stanley, H. E., Aug. 1999. Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series. *Physical Review Letters* 83 (7), 1471–1474.
- Potters, M., Bouchaud, J. P., Laloux, L., Jul. 2005. Financial Applications of Random Matrix Theory: Old Laces and New Pieces.
- Sandoval, L., Bruscatto, A., Venezuela, M. K., 2012. Building portfolios of stocks in the Sao Paulo Stock Exchange using Random Matrix Theory. arXiv:1201.0625v1.
- Wigner, E., 1955. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics* 62, 548–564.
- Wigner, E., 1958. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics* 67, 325–327.
- Wilcox, D., Gebbie, T., 2004. On the analysis of cross-correlations in south african market data. *Physica A: Statistical Mechanics and its Applications* 344, 294 – 298, applications of Physics in Financial Analysis 4 (APFA4).
- Wilcox, D., Gebbie, T., 2007. An analysis of cross-correlations in an emerging market. *Physica A: Statistical Mechanics and its Applications* 375 (2), 584 – 598.

9. Appendix

(A) *Variable Definitions*

1. VIX : Is the Chicago Board Options Exchange Market Volatility Index. Source: Bloomberg.
2. P_t^{Copper} : Is the seasonally adjusted monthly change of the London Metal Exchange copper price. Source: Bloomberg.
3. x_t^{IMACEC} : Is the monthly change of the IMACEC (monthly indicator of economic activity which cover 90% of goods and services included in the GDP). Source: Chilean Central Bank.

4. r_t : Is the Chilean Central Bank Monetary Policy Rate. Source: Chilean Central Bank.

(B) *Tables*

	VIX	P_t^{copper}	x_t^{IMACEC}	π_t	r_t	G_t^1	G_t^2	G_t^3	G_t^{81}	G_t^{82}	G_t^{83}
Mean	22.028	1.512	0.526	0.258	3.803	0.110	-0.053	-0.015	-0.005	0.005	-0.009
Median	21.090	0.558	0.651	0.220	3.500	0.090	-0.031	0.004	-0.003	0.006	-0.005
Maximum	59.890	35.424	7.215	1.110	8.250	1.212	0.698	0.674	0.109	0.102	0.089
Minimum	10.420	-36.471	-5.830	-1.029	0.500	-0.864	-0.770	-0.674	-0.138	-0.114	-0.151
Std. Dev.	9.139	8.954	1.587	0.383	2.004	0.325	0.210	0.182	0.046	0.045	0.038
Skewness	1.415	-0.158	-0.197	-1.029	0.185	0.082	-0.353	0.341	-0.185	-0.283	-0.519
Kurtosis	5.772	5.883	7.176	3.587	2.329	4.159	5.425	5.866	3.321	2.657	4.346
Jarque-Bera	80.447	46.275	90.184	2.863	3.010	7.021	32.687	44.472	1.233	2.243	14.795
Probability	0.000	0.000	0.000	0.238	0.222	0.030	0.000	0.000	0.540	0.326	0.001
Sum	2709	199.665	64.674	34.388	467.750	13.476	-6.499	-1.796	-0.623	0.581	-1.061
Sum Sq. Dev.	10190	10502.89	307.350	19.399	489.907	12.904	5.370	4.050	0.253	0.243	0.180
Observations	123	123	123	123	123	123	123	123	123	123	123

Table 5: Descriptive statistics of Variables used in the VAR estimation. Note: Sample period from 2000M11 to 2011M01. Authors calculations.

Company	Business Sector
SQM A	Mining
ENERSIS	Energy
ANDINA B	Food and beverage
FALABELLA	Commerce and retail
VAPORES	Shipping
ENDESA	Energy
ORO BLANCO	Mining
CAP	Metalmechanics
CALICHERA A	Mining
ANDINA A	Food and beverage
CMPC	Forest products, cellulose
CCU	Food and beverage
COPEC	Energy and natural resources
ENTEL	Telecommunications
MADECO	Metalmechanics
BCI	Banking and Financial Services
LAN	Airline and cargo
COLBUN	Energy
SAN PEDRO	Food and beverage
QUINENCO	Conglomerate
CINTAC	Metalmechanics
ALMENDRAL	Various investments
BSANTANDER	Banking and Financial Services
CONCHATORO	Food and beverage
SMCHILE B	Banking and Financial Services
GASCO	Energy and natural resources
INFORSA	Forest products, cellulose
PARAUCO	Commerce and retail
AGUASA	Water resources
CRISTALES	Plastics and glasses
CHILE	Banking and Financial Services
CEMENTOS	Construction
IANSA	Food and beverage
CGE	Energy

(a) Largest Eigenvector

Company	Business Sector
CAMPOS	Agro-Industrial products
IANSA	Agro-Industrial products
NORTEGRAN	Investments and Real Estate
ORO BLANCO	Mining
CALICHERAA	Mining
SM CHILE E	Mining
ECL	Energy
QUINTEC	Investments and Real Estate
NAVARINO	Transport and Logistics
EMBONOR B	Food and Beverage
PUCOBRE A	Mining
SQM B	Mining
GENER	Energy
PROVIDA	Pension Fund
SOMELA	Metalmechanics
CINTAC	Metalmechanics
ENERSIS	Energy
COPEC	Energy and Natural Resources
ENDESA	Energy
ENTEL	Telecommunications
CMPC	Forest Products related
MADECO	Mining Related
BSANTANDER	Banking and Financial Services
FALABELLA	Commerce and Retail
LAN	Airline Services
CGE	Energy
ALMENDRAL	Investments and Real Estate
CAP	Metalmechanics

(a) Third Largest Eigenvector

Company	Business Sector
CALICHERAA	Mining
ORO BLANCO	Mining
SQM A	Mining
NORTEGRAN	Investments and Real Estate
SQM B	Mining
ANDINA A	Food and Beverage
ANDINA B	Food and Beverage
CMPC	Forest Products related
IANSA	Agro-Industrial products
ECL	Energy
COLBUN	Energy
CAMPOS	Agro-Industrial products
CINTAC	Metalmechanics
INFORSA	Forest Products related
COPEC	Energy and Natural Resources

(b) Second Largest Eigenvector

Company	Business Sector
CGE	Energy
ELECMETAL	Metalmechanics
KOPOLAR	Food and Beverage
ORO BLANCO	Mining
EMBONOR B	Food and Beverage
MARINSA	Investments and Real Estate
SECURITY	Banking and Financial Services
CAP	Metalmechanics
ANDINA A	Food and Beverage
CALICHERAA	Mining
ENERSIS	Energy
CHILE	Banking and Financial Services
BANMEDICA	Investments (health related)
NORTEGRAN	Investments and Real Estate
PARAUCO	Commerce and Retail
BSANTANDER	Banking and Financial Services
EMBONOR A	Food and Beverage
EMILIANA	Food and Beverage
SM CHILE E	Banking and Financial Services
ENDESA	Energy
TELCOY	Telecommunications
CINTAC	Metalmechanics
CONCHATORO	Food and Beverage
FOSFOROS	Matches producers

(b) Smallest Eigenvector

Table 6: Significant participating firms in each eigenvector