Four Essays in Econometrics and Macroeconomics

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Benjamin Born
aus Siegen

Bonn 2011

Dekan:	Prof. Dr. Klaus Sandmann
Erstreferent:	Prof. Dr. Jörg Breitung
Zweitreferent:	Prof. Dr. Gernot J. Müller

Tag der mündlichen Prüfung: 23.09.2011

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Acknowledgments

Many people have contributed to this thesis with their comments and support. Special thanks go to my main advisor Jörg Breitung, who has been a great supervisor, providing invaluable guidance and steady support during my dissertation. Not only have I greatly benefited from being able to work with him on joint projects, but he has also always been supportive of my own research interests.

The interaction with Gernot Müller kindled my interest in macroeconomic research, and I am thankful for his support throughout my doctoral studies. Especially the third chapter would not exist in its present form without his constant feedback.

I would like to thank my coauthor and office mate of many years, Johannes Pfeifer, for countless discussions ranging from *Anticipated Tax Shocks* to *Zombie Movies*. I hope that our cooperation continues for many years to come.

Large parts of the fourth chapter were written during my stay at the *Monetary Policy Research Division* of the *European Central Bank*. I am very grateful for the hospitality I experienced during my stay in Frankfurt and I am especially indebted to my coauthors Michael Ehrmann and Marcel Fratzscher, who gave me the possibility to collaborate with them.

I am also indebted to Urs Schweizer, Silke Kinzig, and Pamela Mertens for their tireless efforts in managing the *Bonn Graduate School of Economics* (BGSE) and in creating a truly stimulating place for research. In this regard, I would like to thank the German Research Foundation (DFG) for financial support.

The four years in Bonn would not have been such a great experience without the colleagues and friends at the Institute of Econometrics, especially Heide Baumung, Jörg Breitung, Norbert Christopeit, Matei Demetrescu, Uli Homm, Christian Pigorsch, and Jörn Tenhofen; and my fellow grad students Rafael Aigner, Inga van den Bongard, Deniz Dizdar, Tilman Drerup, Sebastian Ebert, Andreas Esser, Patrick Hürtgen, Stephan Kurka, Matthias Lang, Juliane Parys, Gregor Schwerhoff, Christian Seel, Mirko Seithe, Marco Sorge, Christoph Wagner, and Florian Zimmermann. Thank you, not only for stimulating scientific discussions, but also for great conference trips, hiking tours, Diablo and Starcraft games, and numerous parties.

Finally, I am deeply indebted to my parents for their patience and continuous support. Alexandra has accompanied this dissertation since our first joint presentation in the first-year Macro PhD class, and she has endured all the ups and downs along the way. Thank you!

Contents

List of Figures

List of Tables

Introduction

In their entry "Econometrics" for *The New Palgrave Dictionary of Economics*, John Geweke, Joel Horowitz, and Hashem Pesaran (2008) write that

> Econometric theory and practice seek to provide information required for informed decision-making in public and private economic policy. This process is limited not only by the adequacy of econometrics but also by the development of economic theory and the adequacy of data and other information. Effective progress, in the future as in the past, will come from simultaneous improvements in econometrics, economic theory and data.

In the spirit of the above quote, this thesis contains four chapters contributing to the improvement of econometrics, economic theory, and data. The first two chapters deal with econometric theory and specifically with the development of diagnostic tests for spatial dependence in cross-sectional data (Chapter 1) and serial correlation in panel data (Chapter 2). Chapter 3 improves on economic theory by developing a structural framework in which the influence of policy risk on the business cycle can be modeled and estimated. In Chapter 4, we tackle the data aspect by constructing, and subsequently analyzing, a unique and novel database on central bank communication about financial stability. A more detailed description of each chapter is given in the remainder of the introduction.

Chapter 1.[1] This chapter proposes simple and robust diagnostic tests for spatial dependence, specifically for spatial error autocorrelation and spatial lag dependence. Spatial dependence is a form of cross-sectional dependence where the correlation between cross-sectional units depends on their relative position in space. Over the last 30 years, this issue has received a growing amount of attention as, driven by an increased availability of geo-referenced data and the development of easy-to-use software, the whole field of spatial econometrics has moved "from the margins to the mainstream of applied econometrics and social science methodology" (Anselin, 2010).

Because spatial dependence can render ordinary least squares (OLS) estimation and inference inefficient or even biased and inconsistent, depending on the specific form of the dependence (Anselin, 1988b), it is of utmost importance to have reliable diagnostic tests. We therefore propose diagnostic tests for spatial error autocorrelation and spatial lag dependence as simple and robust alternatives to the existing Moran's I statistic (Cliff and Ord, 1972, 1981; Moran, 1948) and the Lagrange Multiplier (LM) test statistics by Burridge (1980) and Anselin (1988a).

The idea of our tests is to reformulate the testing problem such that the outer product of gradients (OPG)-variant of the LM test (see e.g. Davidson and MacKinnon, 2004, p. 427) can be employed. Our versions of the tests are based on simple auxiliary regressions, where ordinary regression t and F-statistics can be used to test for spatial autocorrelation and lag dependence. Monte Carlo simulations show that while, under homoskedasticity, our tests perform similarly to the established LM tests, the latter suffer from severe size distortions under heteroskedasticity. Therefore our approach gives practitioners an easy to implement and robust alternative to existing tests.

Chapter 2.[2] Here, we move from the cross-section to a panel data framework. With the additional time series dimension comes a number of advantages (see e.g. Hsiao, 2003, p. 3ff), but also additional possibilities of misspecification. One issue, well-known from the time series literature, is serial correlation of the errors. This chapter proposes various tests for serial correlation in *fixed-effects panel data regression models* with a small number of time periods.

[1] The chapter is based on the paper "Simple Regression Based Tests for Spatial Dependence", jointly written with Jörg Breitung (Born and Breitung, 2011a).

[2] This chapter is based on joint work with Jörg Breitung, "Testing for Serial Correlation in Fixed-Effects Panel Data Models" (Born and Breitung, 2011b).

First, a simplified version of the test for serial correlation suggested by Wooldridge (2002) and Drukker (2003) is considered. The second test is based on the LM statistic suggested by Baltagi and Li (1995), and the third test is a modification of the classical Durbin-Watson statistic (Bhargava et al., 1982). Under the null hypothesis of no serial correlation, all tests possess a standard normal limiting distribution as $N \to \infty$ and T is fixed. Analyzing the local power of the tests, we find that the LM statistic has superior power properties. Furthermore, a generalization to test for autocorrelation up to some given lag order and a test statistic that is robust against time dependent heteroskedasticity are proposed.

The first two chapters dealt with developing econometric tools. In the final two chapters, the focus now shifts to economic theory and the application of econometric tools. The common theme is the analysis of the influence of government actions – by fiscal and monetary authorities, respectively – on economic activity.

Chapter 3.[3] In this chapter, we analyze the role of policy risk in explaining business cycle fluctuations by using an estimated New Keynesian model featuring policy risk as well as uncertainty about technology. The aftermath of the financial and economic crisis is clearly characterized by extraordinary uncertainty regarding U.S. economic policy. Hence, the argument that policy risk, i.e. uncertainty about monetary and fiscal policy, has been holding back the economic recovery in the U.S. during the Great Recession has a large popular appeal. But the empirical literature is still inconclusive with respect to the aggregate effects of (mostly TFP) uncertainty. Studies using different proxies and identification schemes to uncover the effects of uncertainty producing a variety of results.

We analyze the role of policy risk in explaining business cycle fluctuations by using an estimated New Keynesian model featuring policy risk as well as uncertainty about technology. We directly measure uncertainty from aggregate time series using *Sequential Monte Carlo Methods*. While we find considerable evidence of policy risk in the data, we show that the "pure uncertainty"-effect of policy risk is unlikely to play a major role in business cycle fluctuations. In the estimated model, output effects are relatively small due to i) dampening general equilibrium effects that imply a low amplification and ii) counteracting partial effects of uncertainty. Finally, we

[3] The work in this chapter, "Policy Risk and the Business Cycle", has been conducted jointly with Johannes Pfeifer, (Born and Pfeifer, 2011).

show that policy risk has effects that are an order of magnitude larger than the ones of uncertainty about aggregate TFP.

CHAPTER 4.[4] Central banks regularly communicate about financial stability issues, by publishing *Financial Stability Reports* (FSRs) and through speeches and interviews. The chapter asks how such communications affect financial markets. For that purpose, we construct a unique and novel database on CB communication comprising more than 1000 releases of FSRs and speeches/interviews by central bank governors from 37 central banks over a time period from 1996 to 2009, i.e. spanning nearly one and a half decades. The degree of optimism that is expressed in these communications is determined using a computerized textual-analysis software. We then use an event study approach to analyze how financial sector stock indices react to the release of such communication.

The findings suggest that FSRs have a significant and potentially long-lasting effect on stock market returns. At the same time, they tend to reduce stock market volatility. Speeches and interviews, in contrast, have little effect on market returns and do not generate a volatility reduction during tranquil times. However, they had a substantial effect during the 2007-10 financial crisis. It seems that financial stability communication by central banks are perceived by markets to contain relevant information, underlining the importance of differentiating between communication tools, their content, and the environment in which they are employed.

[4]The chapter is based on the ECB working paper "Central Bank Communication on Financial Stability", jointly written with Michael Ehrmann and Marcel Fratzscher (Born et al., 2011a).

Chapter 1

Simple Regression Based Tests for Spatial Dependence

1.1 Introduction

Recent years have seen an increasing availability of regional datasets, leading to a growing awareness of spatial dependence (see Anselin, 2007), an issue that can render ordinary least squares (OLS) estimation and inference inefficient or even biased and inconsistent (see e.g. Anselin, 1988b; Krämer, 2003; Krämer and Donninger, 1987). Arguably the most commonly used test for spatial dependence is Moran's I (see Cliff and Ord, 1972, 1981; Moran, 1948), which is based on regression residuals and which has been shown to be best locally invariant by King (1981). In a Gaussian maximum likelihood framework, Lagrange Multiplier (LM) test statistics were proposed by Burridge (1980) against a spatial error alternative and Anselin (1988a) against a spatial lag alternative and against the joint alternative of spatial lag and spatial error.

We show how to compute the outer product of gradient (OPG) variants of these LM tests (see e.g. Davidson and MacKinnon, 2004, p. 427) based on a simple transformation of the spatial weight matrix. This allows us to compute the test statistics as n (the sample size) times the R^2 from an auxiliary regression. An important advantage of the OPG variant is that it is robust against heteroskedastic and non-normal disturbances. In an alternative regression based approach, Baltagi and Li (2001) use Davidson and MacKinnon's (1984; 1988) double length artificial regression approach to test for spatial error and spatial lag dependence but this is computationally more demanding and not robust to heteroskedasticity.

Monte Carlo simulations demonstrate that (under standard assumptions) our versions of the tests perform similarly to the original LM tests. However, if the errors are heteroskedastic, the latter tests suffer from severe size distortions, whereas the OPG variants turn out to be robust against heteroskedastic errors processes.

The remainder of the chapter is organized as follows. Section 2 reviews the existing maximum likelihood-based test procedures. The regression-based OPG variants of the LM test are presented in Section 3. Section 4 analyzes the asymptotic properties of these tests. Sizes and powers in finite samples are compared in section 5. Section 6 concludes.

1.2 LM Test Statistics

Consider the linear spatial first order autoregressive model with spatially autocorrelated disturbances (see e.g. Anselin, 1988b) given by

$$
\begin{aligned}
y &= \phi W_1^n y + X\beta + u \\
u &= \rho W_2^n u + \varepsilon \ ,
\end{aligned}
\tag{1.1}
$$

where y is an $n \times 1$ vector of observations on a dependent variable, X is an $n \times k$ matrix of regressors, β is the associated $k \times 1$ vector of coefficients, ϕ and ρ are spatial autoregressive parameters, and ε is a vector of independent and identically normally[1] distributed random variables. W_1^n and W_2^n are spatial weight matrices of known constants with zero diagonals.

The spatial error model is obtained by setting $\phi = 0$, yielding

$$
y = X\beta + u, \quad \text{where} \ \ u = (I_n - \rho W_2^n)^{-1}\varepsilon \ . \tag{1.2}
$$

Setting $\rho = 0$, the linear spatial autoregressive model (1.1) with first order autore-

[1]The normality assumption is only required to derive the test statistics from the LM principle.

gressive disturbances becomes the spatial lag model

$$y = \phi W_1^n y + X\beta + \varepsilon \ . \tag{1.3}$$

Accordingly, we will consider the three null hypotheses:

$$\begin{aligned} H_0^a: &\quad \rho = 0 \text{ in } (1.2) \\ H_0^b: &\quad \phi = 0 \text{ in } (1.3) \\ H_0^c: &\quad \rho = 0 \text{ and } \phi = 0 \text{ in } (1.1). \end{aligned}$$

Burridge (1980) shows that the LM statistic[2] for H_0^a results as

$$LM^a = \frac{(\widehat{u}' W_2^n \widehat{u})^2}{\widehat{\sigma}^4 \operatorname{tr}[(W_2^n)^2 + W_2^{n\prime} W_2^n]} \ , \tag{1.4}$$

where $\widehat{u} = y - X\widehat{\beta}$, $\widehat{\beta}$ is the OLS estimator of β in the regression $y = X\beta + u$, and $\widehat{\sigma}^2 = n^{-1}\widehat{u}'\widehat{u}$.

To test hypothesis H_0^b, Anselin (1988a) derives the LM test statistic for the null hypothesis $\phi = 0$:

$$LM^b = \frac{(\widehat{u}' W_1^n y)^2}{\widehat{\sigma}^4 \operatorname{tr}\left[(W_1^n)^2 + W_1^{n\prime} W_1^n\right] + \widehat{\sigma}^2 \widehat{y}' W_1^{n\prime} M W_1^n \widehat{y}} \ , \tag{1.5}$$

where $\widehat{y} = X\widehat{\beta}$, and $M = I_n - X(X'X)^{-1}X'$.

The LM test of the joint null hypothesis H_0^c is obtained as (Anselin, 1988a)

$$\begin{aligned} LM^c = & \frac{1}{\widehat{\sigma}^2} \begin{pmatrix} \widehat{u}' W_2^n \widehat{u} \\ \widehat{u}' W_1^n y \end{pmatrix}' \\ & \begin{pmatrix} \operatorname{tr}[(W_2^n)^2 + W_2^{n\prime} W_2^n] & \operatorname{tr}[(W_2^n + W_2^{n\prime}) W_1^n] \\ \operatorname{tr}[(W_2^n + W_2^{n\prime}) W_1^n] & \operatorname{tr}[(W_1^n)^2 + W_1^{n\prime} W_1^n] + \widehat{\sigma}^2 \widehat{y}' W_2^{n\prime} M W_2^n \widehat{y} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{u}' W_2^n \widehat{u} \\ \widehat{u}' W_1^n y \end{pmatrix} \ . \end{aligned} \tag{1.6}$$

Although these test statistics are derived by applying the LM principle, they cannot

[2]Note that the square of the well-known Moran's I-statistic is asymptotically equivalent to LM^a.

be computed as $n\,R^2$ from a regression of a vector of ones on the gradients of the log-likelihood function (see Engle, 1982). For illustration, consider the gradient of the spatial error model (1.2) with respect to the parameter ρ:

$$g(\beta,\sigma^2,\rho) = -\mathrm{tr}\left[(I_n - \rho W_2^n)^{-1} W_2^n\right] - \frac{1}{2\sigma^2}(y - X\beta)' W_2^n (y - X\beta) \ .$$

Inserting the estimates under the null hypothesis H_0^a, we obtain

$$g(\widehat{\beta},\widehat{\sigma}^2,0) \ \equiv \ \widehat{s}^a \ = \ \sum_{i=1}^n \widehat{s}_i^a \ = \ \frac{1}{\widehat{\sigma}^2}\sum_{i=1}^n \widehat{u}_i \widehat{z}_i^n \ , \tag{1.7}$$

where $\widehat{z}_i^n$ denotes the i'th element of the vector $\widehat{z}^n = W_2^n \widehat{u}$ and $\widehat{s}_i^a = \widehat{\sigma}^{-2}\widehat{u}_i \widehat{z}_i^n$. It is important to note that in general $\widehat{s}_i^a$ is (asymptotically) correlated with $\widehat{s}_j^a$ for $i \neq j$ and, therefore, $n^{-1}\sum_{i=1}^n (\widehat{s}_i^a)^2$ does not converge in probability to the information of the likelihood function. Hence, the usual OPG variant of the LM test is invalid.

1.3 Regression Variants

To compute the OPG variants of the LM tests, the scores are decomposed into uncorrelated components. Let us first consider the spatial error model. To focus on the main issues, we assume that β is known so that $\widehat{u}$ is replaced by $u = y - X\beta$. Let

$$u'z^n = \sum_{i=1}^n \sum_{j\neq i} w_{ij,2}^n u_i u_j \ = \ \sum_{i=2}^n \sum_{j=1}^{i-1} (w_{ij,2}^n + w_{ji,2}^n) u_i u_j \ = \ \sum_{i=2}^n u_i \xi_i^n \ ,$$

where $\xi_i^n = \sum_{j=1}^{i-1}(w_{ij,2}^n + w_{ji,2}^n) u_j$ and $w_{ij,2}^n$ is the (i,j) element of the matrix W_2^n. Defining $\xi^n = (0, \xi_2^n, \ldots, \xi_n^n)'$, we have

$$u'z^n = u'(C_2^n + D_2^{n'})u \ = \ u'(C_2^n + D_2^n)u \ = \ u'\xi^n \ ,$$

where C_2^n and D_2^n are lower triangular matrices such that $W_2^n = C_2^n + D_2^{n'}$ and $\xi^n = (C_2^n + D_2^n)u$. However, there is an important difference between the two formulations of the sum $u'z^n$. Whereas ξ_i^n is associated with an increasing σ-field generated by $\{u_1, \ldots, u_{i-1}\}$, this is not the case for $z_i^n = \sum_{j\neq i} w_{ij,2}^n u_j$, as this variable depends on $\{u_j | j \neq i\}$. This has important consequences for the variance of $u'z^n$. Specifically, under the null hypothesis we have

$$\text{Var}(u'\xi^n) = \sigma^2 E(\xi^{n\prime}\xi^n) \ , \qquad \text{but} \qquad \text{Var}(u'z^n) \neq \sigma^2 E(z^{n\prime}z^n) \ .$$

If W_2^n is symmetric, it is not difficult to show that $\text{Var}(u'z^n) = 2\sigma^2 E(z^{n\prime}z^n)$. The factor 2 results from the fact that, due to the symmetric nature of the sum, the product $u_i u_j$ occurs two times for each combination of i and j. We therefore suggest to use ξ^n instead of $z^n = W_2^n u$ for constructing the test statistic.

Using these results, the scores (1.7) are represented as

$$\widehat{s}^a = \frac{1}{\widehat{\sigma}^2}\widehat{u}'W_2^n\widehat{u} \ = \ \sum_{i=2}^{n}\widetilde{s}_i^a \ = \ \frac{1}{\widehat{\sigma}^2}\sum_{i=2}^{n}\widehat{u}_i\widehat{\xi}_i^n \ , \tag{1.8}$$

where $\widehat{\xi}_i^n$ is the i'th element of the vector $\widehat{\xi}^n = (C_2^n + D_2^n)\widehat{u}$ and $\widetilde{s}_i^a = \widehat{\sigma}^{-2}\widehat{u}_i\widehat{\xi}_i^n$. Since $\widetilde{s}_i^a$ is (asymptotically) uncorrelated with $\widetilde{s}_j^a$ for $i \neq j$, we can construct the OPG variant of the LM statistic as

$$\widetilde{LM}^a = \frac{\left(\sum\limits_{i=1}^{n}\widetilde{s}_i^a\right)^2}{\sum\limits_{i=1}^{n}(\widetilde{s}_i^a)^2} \ . \tag{1.9}$$

The test statistic $\widetilde{LM}^a$ can be seen as a heteroskedasticity robust version of the squared t-statistic in the regression $\widehat{u}_i = \rho^*\widehat{\xi}_i^n + e_i$, where the estimated variance of the least-squares estimator $\widehat{\rho}^*$ is replaced by the estimator

$$\widehat{Var}(\widehat{\rho}^*) = \frac{\sum\limits_{i=2}^{n}(\widehat{u}_i\widehat{\xi}_i^n)^2}{\left(\sum\limits_{i=1}^{n}(\widehat{\xi}_i^n)^2\right)^2} \ .$$

This estimator is similar to the heteroskedasticity robust variance estimator suggested by Eicker (1963, 1967) and White (1980), where the residuals are estimated under the null hypothesis.

For H_0^b the scores result as

$$\widehat{s}^b \;=\; \frac{1}{\widehat{\sigma}^2}\widehat{y}'MW_1^n y \;=\; \frac{1}{\widehat{\sigma}^2}(\widehat{u}'W_1^n\widehat{u} + \widehat{u}'W_1^n\widehat{y}) \;=\; \frac{1}{\widehat{\sigma}^2}\widehat{u}'\widehat{\zeta}^n \;,$$

where

$$\widehat{\zeta}^n = (C_1^n + D_1^n)\widehat{u} + MW_1^n\widehat{y} \tag{1.10}$$

and C_1^n and D_1^n are lower triangular matrices such that $W_1^n = C_1^n + D_1^{n\prime}$. Note that we have introduced the projection matrix M in the last term of (1.10). Due to the idempotency of M, this matrix does not affect the product $\widehat{y}'MW_1y$. However, introducing the matrix M yields a consistent estimator of the asymptotic variance (see the proof of Proposition 1 for more details). We can now form the OPG variant of the score statistic as

$$\widetilde{LM}^b = \frac{\left(\sum\limits_{i=1}^{n}\widehat{u}_i\widehat{\zeta}_i^n\right)^2}{\sum\limits_{i=1}^{n}(\widehat{u}_i\widehat{\zeta}_i^n)^2} \;. \tag{1.11}$$

Finally, for the hypothesis H_0^c the scores are given by

$$\widehat{s}^c \;=\; \frac{1}{\widehat{\sigma}^2}\begin{pmatrix}\widehat{u}'W_2^n\widehat{u}\\ \widehat{u}'W_1^n y\end{pmatrix} \;=\; \frac{1}{\widehat{\sigma}^2}\sum_{i=2}^{n}\widehat{u}_i\widehat{\Upsilon}_i^n \;,$$

where $\widehat{\Upsilon}_i^n = [\widehat{\xi}_i^n, \widehat{\zeta}_i^n]'$. The OPG variant of the score statistic results as

$$\widetilde{LM}^c = \left(\sum_{i=1}^{n}\widehat{u}_i\widehat{\Upsilon}_i^n\right)'\left(\sum_{i=1}^{n}\widehat{u}_i^2\widehat{\Upsilon}_i^n\widehat{\Upsilon}_i^{n\prime}\right)^{-1}\left(\sum_{i=1}^{n}\widehat{u}_i\widehat{\Upsilon}_i^n\right) \;, \tag{1.12}$$

which is equivalent to nR^2 obtained from a regression of a constant on $\widehat{u}_i\widehat{\Upsilon}_i^n$.

1.4 Asymptotic Properties

In the previous section, some regression variants of the LM test statistics were suggested. These test statistics are equivalent to heteroskedasticity-robust t and

F-statistics of the following regressions:

$$H_0^a: \ \rho^* = 0 \ \text{ in } \ \widehat{u}_i = \rho^* \widehat{\xi}_i^n + e_i \tag{1.13}$$

$$H_0^b: \ \phi^* = 0 \ \text{ in } \ \widehat{u}_i = \phi^* \widehat{\zeta}_i^n + e_i \tag{1.14}$$

$$H_0^c: \ \rho^* = 0 \ \text{ and } \ \phi^* = 0 \ \text{ in } \ \widehat{u}_i = \rho^* \widehat{\xi}_i^n + \phi^* \widehat{\zeta}_i^n + e_i \ , \tag{1.15}$$

where $\widehat{u}_i = y_i - x_i'\widehat{\beta}$, and $\widehat{\beta}$ denotes the OLS estimator of β. The regressors $\widehat{\xi}_i^n$ and $\widehat{\zeta}_i^n$ are defined in Section 1.3. To analyze the asymptotic properties, we make the following assumptions:

Assumption 1. *(i) The errors ε_i are independent random variables with $E(\varepsilon_i) = 0$, $E(\varepsilon_i^2) = \sigma_i^2 < c < \infty$ and $E(|\varepsilon_i|^{4+\delta}) < \infty$ for all i and some $\delta > 0$. (ii) The vector x_i is a $k \times 1$ vector of constants with $\lim_{n\to\infty} n^{-1} \sum_{i=1}^n x_i x_i' \to C_X$ (positive definite).*

Assumption 2. *(i) The diagonal elements of $W_h^n = (w_{ij,h}^n)$ are zero. (ii) All row and column sums of W_h^n and $W_h^{n\prime} W_h^n$ are uniformly bounded for all n and $h \in \{1, 2\}$.*

These assumptions are standard in the asymptotic analysis of spatial models (e.g. Kelejian and Prucha, 2001; Lee, 2007).

The following proposition states that the OPG variants of the LM tests suggested in Section 3 possess the usual asymptotic distributions if the errors are heteroskedastic. Furthermore, if the errors are homoskedastic, the test statistics are asymptotically equivalent to the original LM tests suggested by Burridge (1980) and Anselin (1988a).

Proposition 1. *(a) Under Assumptions 1 – 2 and hypotheses H_0^a, H_0^b, and H_0^c, we have*

$$\widetilde{LM}^a \xrightarrow{d} \chi_1^2, \quad \widetilde{LM}^b \xrightarrow{d} \chi_1^2 \ , \quad \text{and} \ \ \widetilde{LM}^c \xrightarrow{d} \chi_2^2 \ .$$

(b) If $\sigma_1^2 = \cdots = \sigma_n^2$ (homoskedastic errors), it follows that

$$\widetilde{LM}^a - LM^a \xrightarrow{p} 0, \quad \widetilde{LM}^b - LM^b \xrightarrow{p} 0, \quad \text{and} \ \ \widetilde{LM}^c - LM^c \xrightarrow{p} 0 \ .$$

The proof of this proposition can be found in the appendix.

An important implication of this proposition is that if the errors are homoskedastic, the LM tests can be performed by using the ordinary t-statistics for $\rho^* = 0$ or $\phi^* = 0$

in (1.13) and (1.14). The hypothesis H_0^c can be tested by computing the F-statistic of the joint hypothesis in (1.15).

1.5 Monte Carlo Simulations

In this section, we conduct a small Monte Carlo study to demonstrate the finite sample properties of our new tests and investigate their relative performance compared to the original LM approaches of Burridge (1980) and Anselin (1988a).[3] We simulate three different models. Models (1.2) and (1.3) are employed to evaluate the spatial error test and the spatial lag test, respectively. The test for the joint hypothesis is based on model (1.1). The matrix of exogenous regressors, X, contains two regressors, x_1 and x_2, with associated parameters β_1 and β_2, where $\beta = (1, 1)'$. x_1 is a vector of ones and the elements of x_2 are drawn independently from a standard normal distribution. The elements of the vector ε are generated as independent normally distributed random variables such that $E(\varepsilon\varepsilon') = I$. Furthermore, we set $W_1 = W_2$.

Our weight matrix design closely follows Arraiz et al. (2010).[4] The authors use a setup that mimics the spacing of US states, i.e., units located in the northeast portion of their model space are closer to each other and have more neighbors than the units in the other three quadrants. They refer to a weight matrix defined in such a way as *north-east modified rook matrix*. We choose a specification where the share of units located in the northeast is approximately 75%.[5] The distance between any two units is defined as the Euclidean distance

$$d(i_1, i_2) = \left[(x_1 - x_2)^2 + (y_1 - y_2)^2\right]^{1/2}.$$

[3]There are alternative tests developed for the heteroskedastic case. Kelejian and Robinson (1998) propose a joint test for spatial error dependence and heteroskedasticity, where the variance is a (possibly unknown) function of explanatory variables. Kelejian and Robinson (2004) propose a heteroskedasticity robust version of Moran's I that is based on a consistent estimator of σ_i^2. Since we focus on tests that do not require any knowledge about the variance function, we do not include these alternative tests in our Monte Carlo experiment.

[4]See the online appendix (`http://www.ect.uni-bonn.de/spatialtest_webappendix.pdf`) for additional results using alternative weight matrices.

[5]See Arraiz et al. (2010) for a detailed description of this weight matrix design.

The elements of the row normalized weighting matrix are then defined as

$$w_{ij} = \frac{w_{ij}^*}{\sum_{j=1}^n w_{ij}^*}\ , \text{ where } \quad w_{ij}^* = \begin{cases} 1 & \text{if } \ 0 < d(i_1, i_2) \leq 1 \\ 0 & \text{else} \end{cases} .$$

The original LM tests remain valid under heteroskedasticity as long the heteroskedasticity is not itself spatially correlated (Kelejian and Robinson, 2004). In practice, however, it is reasonable to assume that the heteroskedasticity possesses a spatial pattern. We therefore introduce a disturbance $\psi_i = \varepsilon_i x_{2i}$ with a "medium" extent of heteroskedasticity (see Kelejian and Robinson, 1998), where the spatial correlation in the heteroskedasticity is induced by the sorted vector x_2.

Table 1.1: Empirical sizes under heteroskedasticity, 5% level

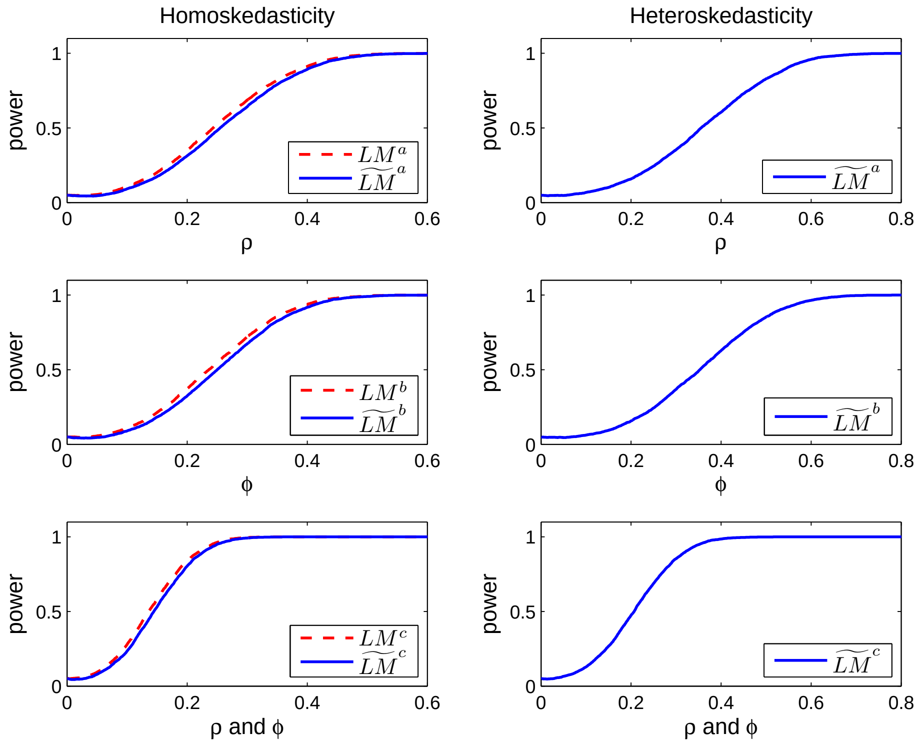
n	LM^a	$\widetilde{LM}^a$	LM^b	$\widetilde{LM}^b$	LM^c	$\widetilde{LM}^c$	LM^a_{boot}	LM^b_{boot}	LM^c_{boot}
105	0.229	0.050	0.243	0.045	0.399	0.047	0.052	0.049	0.066
166	0.235	0.044	0.243	0.046	0.406	0.041	0.051	0.053	0.057
241	0.265	0.044	0.267	0.050	0.457	0.045	0.046	0.054	0.063
486	0.305	0.048	0.310	0.048	0.494	0.047	0.050	0.049	0.053
974	0.353	0.051	0.355	0.050	0.581	0.042	0.051	0.049	0.050

Note: Empirical sizes are calculated using 5000 replications.

Table (1.1) presents the empirical sizes obtained from the Monte Carlo simulations under heteroskedasticity. Not shown here are the empirical sizes under homoskedasticity which are all close to the nominal size. The results change considerably in the presence of heteroskedastic errors as the original LM tests are now strongly oversized. Our OPG variants, on the other hand, do not exhibit any notable size distortions. An alternative approach to produce heteroskedasticity-robust test statistics is to bootstrap the original LM tests.[6] We report the empirical sizes of the bootstrapped LM tests in the last three columns of table (1.1).

[6] We thank an anonymous referee for suggesting the bootstrap test. Specifically, we employ the wild bootstrap (Liu, 1988). In this approach, the true OLS residuals $\widehat{u}_i$ are replaced in the bootstrap DGP by $\widehat{u}_i^* = \widehat{u}_i \varepsilon_i$, where $\varepsilon_i = 1$ with probability 0.5 and $\varepsilon_i = -1$ with probability 0.5 (see Davidson and Flachaire, 2008).

Figure 1.1: Size corrected power under homo- and heteroskedasticity (n=241)



Size corrected power curves[7] for the tests are depicted in figure (1.1). The left column of plots shows size corrected power curves of original and OPG versions of the LM tests under homoskedasticity. The OPG variants are nearly as powerful as the original LM tests. In the right column, we only plot the size corrected power curves of our proposed OPG variants, as the original LM tests suffer from massive size distortions under heteroskedasticity.

1.6 Conclusion

In this chapter, we propose simple and robust diagnostic tests for spatial error autocorrelation and spatial lag dependence. We reformulate the testing problem such

[7]As pointed out by Krämer (2005) and Martellosio (2010), the power of spatial autocorrelation tests can drop to zero for some combinations of X and W_2^n. Since the regression based test is asymptotically equivalent to Moran's I, our test suffers from the same deficiency.

that the outer product of gradients (OPG) variant of the LM tests can be employed. Our versions of the tests are based on simple auxiliary regressions, where ordinary regression t and F statistics can be used to test for spatial autocorrelation and lag dependence. We show that these tests are asymptotically equivalent to the existing LM tests, yet simpler to implement. An important advantage of the proposed test statistics is that they are robust against heteroskedastic errors.

Monte Carlo simulations suggest that our new tests have good size properties, even under heteroskedasticity, where the original LM tests suffer from size distortions. Hence, we believe that the proposed tests will give researchers a robust and easily implementable tool for their applied work.

Appendix to Chapter 1

Proof of Proposition 1:

(a) We first assume that β is known such that $u_i = y_i - x_i'\beta$, $u = [u_1, \ldots, u_n]'$, and

$$u'W_2^n u = \sum_{i=2}^{n} \left(\sum_{j=1}^{i-1} (w_{ij,2}^n + w_{ji,2}^n) u_j \right) u_i = \sum_{i=2}^{n} u_i \xi_i^n \ ,$$

where ξ_i^n is the i'th element of the vector $\xi^n = (C_2^n + D_2^n)u$. It is important to note that although the sequence $\{(u_2\xi_2), \ldots, (u_n\xi_n^n)\}$ is different for a different ordering of the cross section units and the respective weights, the sum $\sum u_i\xi_i^n$ is invariant to the ordering of the units i. Thus, since we are only interested in the distribution of $\sum u_i\xi_i^n$, the ordering of the units does not matter for our asymptotic results.[8]

Let $\mathcal{F}_n$ be the increasing σ-algebra generated by $\{u_1, \ldots, u_n\}$ and $\mathcal{Z}_n = \sum_{i=2}^n u_i\xi_i^n$. Note that $\mathcal{Z}_n$ is a martingale difference sequence with respect to the filtration $\mathcal{F}_n$. From Assumptions 1 and 2 it follows that[9]

$$\begin{array}{rl} (i) & \dfrac{1}{n}\mathcal{Z}_n \ \xrightarrow{p} \ 0 \\ (ii) & \dfrac{1}{n}\displaystyle\sum_{i=2}^{n} u_i^2(\xi_i^n)^2 \ \xrightarrow{p} \ \lim_{n\to\infty} \dfrac{1}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}(w_{ij,2}^n + w_{ji,2}^n)^2\sigma_j^2\sigma_i^2 \ \equiv \ s_{\mathcal{Z}}^2 \\ (iii) & \dfrac{1}{\sqrt{n}}\mathcal{Z}_n \ \xrightarrow{d} \ N(0, s_{\mathcal{Z}}^2) \ . \end{array}$$

It is not difficult to see that the limiting distribution does not change if u_i is replaced by the OLS residual $\widehat{u}_i$. To see this, consider

$$\widehat{\mathcal{Z}}_n = \sum_{i=2}^{n} \widehat{u}_i \widehat{\xi}_i^n = \sum_{i=2}^{n} \left(\sum_{j=1}^{i-1} (w_{ij,2}^n + w_{ji,2}^n)\widehat{u}_j \right) \widehat{u}_i \ .$$

[8] As pointed out by a referee, this invariance property may be lost if the weight matrices are renormalized for different orderings.

[9] Using different techniques, a similar result is derived by Kelejian and Prucha (2001, Theorem 1).

Using $\widehat{u} = u - X(\widehat{\beta} - \beta)$, we obtain

$$\widehat{\mathcal{Z}}_n = \mathcal{Z}_n + (\widehat{\beta} - \beta)' X' W_2^n X (\widehat{\beta} - \beta) - (\widehat{\beta} - \beta)' X' (W_2^n + W_2^{n\prime}) u \ .$$

Assumptions 1 and 2 imply that $\widehat{\beta} - \beta$ is $O_p(n^{-1/2})$ and, therefore,

$$\frac{1}{\sqrt{n}} \widehat{\mathcal{Z}}_n = \frac{1}{\sqrt{n}} \mathcal{Z}_n + O_p(n^{-1/2}) \ .$$

In a similar manner, it can be shown that

$$\frac{1}{n} \sum_{i=2}^{n} \widehat{u}_i^2 (\widehat{\xi}_i^n)^2 \ \xrightarrow{p} \ s_{\mathcal{Z}}^2 \ .$$

It follows that

$$\widetilde{LM}^a = \frac{\widehat{\mathcal{Z}}_n^2}{\sum\limits_{i=2}^{n} \widehat{u}_i^2 (\widehat{\xi}_i^n)^2} \xrightarrow{d} \chi_1^2 \ .$$

Regarding the asymptotic distribution of $\widetilde{LM}^b$, we first consider the case that β is known. Let ζ^n be constructed as $\widehat{\zeta}^n$, where $\widehat{\beta}$ is replaced by β:

$$\begin{aligned} \mathcal{Z}_n^* = u'\zeta^n &= u' W_1^n u + u' W_1^n X \beta = u'(C_1^n + D_1^n) u + u' W_1^n X \beta \\ &= \sum_{i=1}^{n} u_i (\xi_i + \mu_i) \ , \end{aligned}$$

where μ_i is the i'th element of the vector $\mu = W_1^n X \beta$. It follows that

$$\frac{1}{n} \mathrm{Var}(\mathcal{Z}_n^*) \ \rightarrow \ s_{\mathcal{Z}}^2 + \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mu_i^2 \sigma_i^2 \ \equiv \ s_{\mathcal{Z}^*}^2$$

$$\text{and} \qquad \frac{1}{\sqrt{n}} \mathcal{Z}_n^* \ \xrightarrow{d} \ \mathcal{N}(0, s_{\mathcal{Z}^*}^2).$$

Using similar arguments as for $\widehat{\mathcal{Z}}_n$, we can show that

$$\frac{1}{\sqrt{n}} \widehat{\mathcal{Z}}_n^* = \frac{1}{\sqrt{n}} \mathcal{Z}_n^* + O_p(n^{-1/2}) \ ,$$

where $\widehat{\mathcal{Z}}_n^*$ is constructed as $\mathcal{Z}_n^*$, with β replaced by $\widehat{\beta}$. However, some caution is necessary to derive the variance

$$\mathrm{Var}(\widehat{\mathcal{Z}}_n^*) = \mathrm{Var}(\widehat{u}' W_1^{n\prime} \widehat{u}) + 2E(\widehat{u}' W_1^n \widehat{u}\widehat{u}' W_1^n X\widehat{\beta}) + E(\widehat{\beta}' X' W_1^{n\prime} \widehat{u}\widehat{u}' W_1^n X\widehat{\beta}) \ .$$

Let $\widehat{\mu}_i$ denote the i'th element of the vector $\widehat{\mu} = W_1^n X\widehat{\beta}$. Then,

$$\begin{aligned}
\frac{1}{n}E(\widehat{u}' W_1^{n\prime} \widehat{u}\widehat{u}' W_1^n X\widehat{\beta}) \ &= \ \frac{1}{n}E\left[\left(\sum_{i=2}^{n} \widehat{u}_i\widehat{\xi}_i\right)\left(\sum_{i=1}^{n}\widehat{u}_i\widehat{\mu}_i\right)\right] \\
= E\left\{\left[\left(\frac{1}{\sqrt{n}}\sum_{i=2}^{n} u_i\xi_i\right) + O_p(n^{-1/2})\right]\left[\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} u_i\mu_i - \mu_i x_i'(\widehat{\beta}-\beta)\right) + O_p(n^{-1/2})\right]\right\} \\
= O(n^{-1/2}) \ .
\end{aligned}$$

Finally,

$$\begin{aligned}
\frac{1}{n}E(\widehat{\beta}' X' W_1^{n\prime} \widehat{u}\widehat{u}' W_1^n X\widehat{\beta}) &= \frac{1}{n}\beta' X' W_1^{n\prime} M \, E(uu') \, M W_1^n X\beta + O(n^{-1/2}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\widetilde{\mu}_i^2\sigma_i^2 + O(n^{-1/2}) \ ,
\end{aligned}$$

where $\widetilde{\mu}_i^2$ is the i'th element of the vector $\widetilde{\mu} = M W_1^n X\beta$. It follows that

$$\widehat{s}^2_{\widehat{\mathcal{Z}}^*} = \frac{1}{n}\sum_{i=1}^{n}(\widehat{u}_i\widehat{\zeta}_i^n)^2 \ ,$$

with $\widehat{\zeta}_i^n$ as defined in (1.10), converges to the limiting variance of $n^{-1/2}\widehat{\mathcal{Z}}_n^*$. Hence,

$$\widetilde{LM}^b = \frac{(\widehat{\mathcal{Z}}_n^*)^2}{\sum_{i=1}^{n}(\widehat{u}_i\widehat{\zeta}_i^n)^2} \ \xrightarrow{d} \ \chi_1^2 \ .$$

Finally, using these results, it is easy to verify that

$$\widetilde{LM}^c \ = \ \widehat{\mathcal{Y}}_n' \left(\sum_{i=1}^{n}\widehat{u}_i^2\widehat{\Upsilon}_i^n\widehat{\Upsilon}_i^{n\prime}\right)^{-1}\widehat{\mathcal{Y}}_n \ \xrightarrow{d} \ \chi_2^2 \ ,$$

where $\widehat{\mathcal{Y}}_n = [\widehat{\mathcal{Z}}_n, \widehat{\mathcal{Z}}_n^*]'$ and $\widehat{\Upsilon}_i^n = [\widehat{\xi}_i^n, \widehat{\zeta}_i^n]'$.

(b) If the errors are homoskedastic, the LM statistic suggested by Burridge (1980) has the asymptotic representation

$$LM^a = \frac{(u'W_2^n u)^2}{\sigma^4 \text{tr}\left[(W_2^n)^2 + W_2^{n\prime} W_2^n\right]} + o_p(1) \ .$$

To analyze the asymptotic distribution of the statistic $\widetilde{LM}^a$ under homoskedasticity, we first note that

$$\begin{aligned}
\text{tr}\left[(W_2^n)^2 + W_2^{n\prime} W_2^n\right] &= \text{tr}\left[\left(C_2^n + D_2^{n\prime}\right)\left(C_2^n + D_2^{n\prime}\right) + \left(C_2^{n\prime} + D_2^n\right)\left(C_2^n + D_2^{n\prime}\right)\right] \\
&= \text{tr}\left(C_2^{n\prime} C_2^n\right) + \text{tr}\left(D_2^{n\prime} D_2^n\right) + 2\,\text{tr}\left(D_2^{n\prime} C_2^n\right) \\
&= \text{tr}\left[(C_2^n + D_2^n)'\,(C_2^n + D_2^n)\right] \ .
\end{aligned}$$

It follows that

$$\frac{1}{n}\sum_{i=1}^{n} \widehat{u}_i^2 (\widehat{\zeta}_i^n)^2 \xrightarrow{p} \lim_{n\to\infty} \frac{\sigma^4}{n} \text{tr}\left[(C_2^n + D_2^n)'(C_2^n + D_2^n)\right] \ = \ \lim_{n\to\infty} \frac{\sigma^4}{n} \text{tr}\left[(W_2^n)^2 + W_2^{n\prime} W_2^n\right] \ .$$

Therefore,

$$\widetilde{LM}^a = \frac{(u'(C_2^n + D_2^n)'(C_2^n + D_2^n)u)^2}{\sigma^4 \text{tr}[(W_2^n)^2 + W_2^{n\prime} W_2^n]} + o_p(1)$$

and $\widetilde{LM}^a - LM^a \xrightarrow{p} 0$.

To analyze the asymptotic properties of $\widetilde{LM}^b$ under homoskedastic errors, we first note that from the results in (a) we obtain

$$\frac{1}{\sqrt{n}} \widehat{\mathcal{Z}}_n^* = \frac{1}{\sqrt{n}} \left[u'(C_1^n + D_1^n)u + u'W_1^n X\beta\right] + o_p(1)$$

and

$$\begin{aligned}
\frac{1}{n}\sum_{i=2}^{n} \widehat{u}_i^2 (\widehat{\zeta}_i^n)^2 &\xrightarrow{p} \lim_{n\to\infty} \frac{\sigma^2}{n} E\left[(C_1^n + D_1^n)u + MW_1^n X\beta\right]' \left[(C_1^n + D_1^n)u + MW_1^n X\beta\right] \\
&= \sigma^4 \text{tr}\left[(W_1^n)^2 + W_1^{n\prime} W_1^n\right] + \sigma^2 \beta' X' W_1^{n\prime} M W_1^n X\beta
\end{aligned}$$

It follows that

$$\widetilde{LM}^b = \frac{\left[u'(C_1^n + D_1^n)u + u'W_1^n X\beta\right]^2}{\sigma^4 \text{tr}\left[(W_1^n)^2 + W_1^{n\prime} W_1^n\right] + \sigma^2 \beta' X' W_1^{n\prime} M W_1^n X\beta} + o_p(1)$$

and, therefore, the test statistic $\widetilde{LM}^b$ is asymptotically equivalent to the original LM^b statistic presented in (1.5). The asymptotic equivalence of the joint test statistics LM^c and $\widetilde{LM}^c$ follows directly from these results.

Chapter 2

Testing for Serial Correlation in Fixed-Effects Panel Data Models

2.1 Introduction

Panel data models are increasingly popular in applied work as they have many advantages over cross-sectional approaches (see e.g. Hsiao, 2003, p. 3ff). The classical linear panel data model assumes serially uncorrelated disturbances. As argued by Baltagi (2008, p. 92), this assumption is likely to be violated as the dynamic effect of shocks to the dependent variable is often distributed over several time periods. In such cases, serial correlation leads to inefficient estimates and biased standard errors and, therefore, tests for serial correlation are nowadays routinely applied in regression analysis based on time series.

For panel data a number of tests for serial error correlation have been proposed in the literature. Bhargava et al. (1982) generalize the Durbin-Watson statistic to the fixed-effects panel model. Baltagi and Li (1991, 1995) and Baltagi and Wu (1999) derive LM statistics for first order serial correlation. Drukker (2003), elaborating on an idea originally proposed by Wooldridge (2002), proposes an easily implementable test for serial correlation based on the OLS residuals of the first-differenced model.

However, all these tests have limitations. A serious problem of the Bhargava et al. (1982) statistic is that the distribution depends on N and T and, therefore, the critical values have to be provided in large tables depending on both dimensions.

Baltagi and Li (1995) noted that, for fixed T, their test statistic does not possess the usual χ^2 limiting distribution due to the (Nickell) bias in the estimation of the autocorrelation coefficient. The Wooldridge-Drukker test is not derived from the usual test principles (like LM, LR or Wald) and, therefore, it is not clear whether the test has desirable properties. Furthermore, these tests are not robust against temporal heteroskedasticity and not applicable to unbalanced panels (with the exception of the statistic proposed by Baltagi and Wu (1999)).

In this chapter, we propose new test statistics and modifications of existing test statistics that correct some of these limitations. In Section 2, we first present the model framework and briefly review the existing tests. Our new test procedures are considered in Section 3 and the small sample properties of the tests are studied in Section 4. Section 5 concludes.

2.2 Preliminaries

Consider the usual fixed effects panel data model with serially correlated disturbances

$$y_{it} = x_{it}'\beta + \mu_i + u_{it} \tag{2.1}$$

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it} \ , \tag{2.2}$$

where $i = 1, \ldots, N$ denotes the cross-section dimension and $t = 1, \ldots, T$ is the time dimension. In our benchmark situation, the following assumption is imposed:

Assumption 3. *(i) The error ε_{it} is independently distributed across i and t with $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}^2) = \sigma_i^2$, $\lim_{N\to\infty} \max(\sigma_i^2)/\left(\sum_{i=1}^N \sigma_i^2\right) = 0$ and $E\left|\varepsilon_{it}\right|^{4+\delta} < \infty$ for some $\delta > 0$. (ii) For the vector of regressors it is assumed that*

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)' \ \xrightarrow{p} \ \overline{C}$$

is a finite and positive definite matrix. Furthermore, x_{it} is independent of ε_{is} for all i, t and s.

To test the null hypothesis $\rho = 0$, Bhargava et al. (1982) propose a pooled Durbin-Watson statistic given by

$$p\mathrm{DW} = \frac{\sum\limits_{i=1}^{N}\sum\limits_{t=2}^{T}(\widehat{u}_{it} - \widehat{u}_{i,t-1})^2}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T}\widehat{u}_{it}^2} ,$$

where $\widehat{u}_{it} = y_{it} - x_{it}'\widehat{\beta} - \widehat{\mu}_i$, and $\widehat{\beta}$ and $\widehat{\mu}_i$ denote the least-squares dummy-variable (or within-group) estimator of β and μ_i, respectively. A serious problem of this test is that its null distribution depends on N and T and, therefore, the critical values are provided in large tables depending on both dimensions in Bhargava et al. (1982). Furthermore, no critical values are available for *unbalanced* panels.

Baltagi and Li (1995) derive the LM test statistic for the hypothesis $\rho = 0$ assuming normally distributed errors. The resulting test statistic is equivalent to (the LM version of) the t-statistic of ϱ in the regression

$$\widehat{u}_{it} = \varrho\widehat{u}_{i,t-1} + \nu_{it} \ . \tag{2.3}$$

The LM test statistic results as

$$\mathrm{LM} = \sqrt{\frac{NT^2}{T-1}}\left(\frac{\sum\limits_{i=1}^{N}\sum\limits_{t=2}^{T}\widehat{u}_{it}\widehat{u}_{i,t-1}}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T}\widehat{u}_{it}^2}\right) .$$

Baltagi and Li (1995) show that if $N \to \infty$ and $T \to \infty$, the LM statistic has a standard normal limiting distribution. However, if T is fixed and $N \to \infty$, the application of the respective critical values leads to severe size distortions (see Section 3.2).

2.3 Test Statistics for Fixed T

In this section, we consider test procedures that are valid for fixed T and $N \to \infty$. All statistics have the general form

$$\lambda_{NT} \;=\; \frac{\sum_{i=1}^{N} \widehat{e}_i' A_T \widehat{e}_i}{\sqrt{\sum_{i=1}^{N} (\widehat{e}_i' A_T \widehat{e}_i)^2 - \frac{1}{N}\left(\sum_{i=1}^{N} \widehat{e}_i' A_T \widehat{e}_i\right)^2}}\,, \tag{2.4}$$

where $\widehat{e}_i = [\widehat{e}_{i1}, \ldots, \widehat{e}_{iT}]'$, $\widehat{e}_{it} = y_{it} - x_{it}'\widehat{\beta}$ and A_T is a deterministic $T \times T$ matrix. The matrix A_T eliminates the individual effects from $\widehat{e}_i$ and ensures that $N^{-1}E(\widehat{e}_i' A_T \widehat{e}_i)$ converges in probability to zero, as $N \to \infty$.

The following lemma presents the limiting distribution of the test statistic under a sequence of local alternatives assuming that T is fixed and $N \to \infty$.

Lemma 1. *Let A_T denote a $T \times T$ matrix obeying $A_T\iota = 0$, $A_T'\iota = 0$, and $tr(A_T) = 0$. Under Assumption 1 and the sequence of local alternatives $\rho_N = c/\sqrt{N}$, the asymptotic distribution of the test statistic is given by*

$$\lambda_{NT} \;\xrightarrow{d}\; \mathcal{N}\left(\frac{c\kappa\left(\sum\limits_{t=2}^{T} a_{t,t-1} + a_{t-1,t}\right)}{\sqrt{\sum\limits_{t=1}^{T}\sum\limits_{s=1}^{T} a_{ts}^2 + a_{ts}a_{st}}}, 1\right),$$

where $\kappa = m_2/\sqrt{m_4}$ with $m_k = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N}\sigma_i^k$, and a_{ts} denotes the (t,s) element of the matrix A_T.

Remark 1: Under the null hypothesis ($c = 0$), all test statistics considered in this section are asymptotically standard normal.

Remark 2: The limiting distribution does not depend on the regressor matrix X and, therefore, the estimation error $\widehat{\beta} - \beta$ does not affect the limiting distribution. We therefore treat $e_{it} = y_{it} - x_{it}'\beta$ as being known in what follows.

Remark 3: A straightforward extension to unbalanced panel data with individual specific T_i yields

$$\lambda_{NT} \xrightarrow{d} \mathcal{N}\left(c \frac{tr\left(\lim\limits_{N\to\infty} N^{-1} \sum\limits_{i=1}^{N} \sum\limits_{t=2}^{T_i} \sigma_i^2 (a_{t,t-1} + a_{t-1,t})\right)}{\sqrt{tr\left(\lim\limits_{N\to\infty} N^{-1} \sum\limits_{i=1}^{N} \sum\limits_{t=1}^{T_i} \sum\limits_{s=1}^{T_i} \sigma_i^4 (a_{ts}^2 + a_{ts} a_{st})\right)}}, 1 \right).$$

Hence, to accommodate unbalanced panel data the tests are computed with individual specific transformation matrices A_{T_i} instead of a joint matrix A_T.

2.3.1 The Wooldridge-Drucker Test

To obtain a valid test statistic for fixed T, Wooldridge (2002, p. 282f) suggests to run a least squares regression of the first differences $\Delta e_{it} = e_{it} - e_{i,t-1}$ on the lagged differences $\Delta e_{i,t-1}$. Under the null hypothesis $\rho = 0$, the first order autocorrelation of the first differences converges in probability to -0.5. Since Δe_{it} is serially autocorrelated, Drukker (2003) suggests to employ heteroskedasticity and autocorrelation consistent (HAC) standard errors,[1] yielding the test statistic

$$\text{WD} = \frac{\widehat{\theta} + 0.5}{\widehat{s}_\theta} .$$

Here, $\widehat{\theta}$ denotes the least-squares estimator of θ in the regression

$$\Delta e_{it} = \theta \Delta e_{i,t-1} + \eta_{it} , \qquad (2.5)$$

and $\widehat{s}_\theta^2$ is the HAC estimator of the standard errors given by

$$\widehat{s}_\theta^2 = \frac{\sum\limits_{i=1}^{N} \left(\sum\limits_{t=3}^{T} \Delta e_{i,t-1} \widehat{\eta}_{it} \right)^2}{\left(\sum\limits_{i=1}^{N} \sum\limits_{t=3}^{T} \Delta e_{i,t-1}^2 \right)^2} ,$$

with $\widehat{\eta}_{it}$ as the pooled OLS residual from the autoregression (2.5).

[1] This approach is also known as “robust cluster” or “panel corrected” standard errors.

In the following theorem, we propose a simplified and asymptotically equivalent version of the Wooldridge-Drucker test.

Theorem 2. *Let* $z_{Ti} = \sum_{t=3}^{T}(e_{it} - \frac{1}{2}e_{i,t-1} - \frac{1}{2}e_{i,t-2})(e_{i,t-1} - e_{i,t-2})$ *and*

$$\widetilde{WD} \;=\; \frac{\sum_{i=1}^{N} z_{Ti}}{\sqrt{\sum_{i=1}^{N} z_{Ti}^2 - \frac{1}{N}\left(\sum_{i=1}^{N} z_{Ti}\right)^2}}\;. \tag{2.6}$$

(i) Under Assumption 1, $T \geq 3$*, and* $\rho_N = c/\sqrt{N}$*, it follows that*

$$\widetilde{WD} \;\xrightarrow{d}\; \mathcal{N}\left(c\,\frac{\kappa(T-2)}{\sqrt{2(T-3)+3}}, 1\right),$$

where κ *is defined in Lemma 1.*

(ii) The test statistic $\widetilde{WD}$ *is asymptotically equivalent to* WD *in the sense that* $\widetilde{WD} - WD \xrightarrow{p} 0$.

Remark 4: The form of the test statistic results from

$$\widehat{\theta} + 0.5 \;=\; \frac{\sum_{i=1}^{N}\sum_{t=3}^{T}(\Delta e_{it}\Delta e_{i,t-1} + \frac{1}{2}\Delta e_{i,t-1}^2)}{\sum_{i=1}^{N}\sum_{t=3}^{T}\Delta e_{it}\Delta e_{i,t-1}} \;=\; \frac{\sum_{i=1}^{N} z_{Ti}}{\sum_{i=1}^{N}\sum_{t=3}^{T}\Delta e_{it}\Delta e_{i,t-1}}\;.$$

The $\widetilde{\text{WD}}$ statistic deviates from WD by computing the HAC variance based on the residuals obtained under the null hypothesis, i.e., $\eta_{it}^0 = \Delta e_{it} + 0.5\Delta e_{i,t-1}$ instead of the residuals $\widehat{\eta}_{it}$ used to compute the original WD statistic.

Remark 5: It is interesting to note that under the alternative we have as $N \to \infty$

$$\widehat{\theta} + 0.5 \;\xrightarrow{p}\; \frac{r_1 - r_2}{2(1 - r_1)} + O(T^{-1})\;,$$

where r_j is the j'th autocorrelation of u_{it}. This suggests that WD (and $\widetilde{\text{WD}}$) is a test against the *difference* between the first and second order autocorrelation of the errors. Therefore, the test is expected to have poor power against alternatives with $r_1 \approx r_2$.

Remark 6: The test statistic is robust against cross-sectional heteroskedasticity. However, using $\theta = -0.5$ requires that the variances do not change in time, i.e., $E(u_{it}^2) = E(u_{is}^2) = \sigma_i^2$. Thus, this test rules out time dependent heteroskedasticity.

2.3.2 The LM Test

An important problem with the LM test suggested by Baltagi and Li (1995) is that the limit distribution depends on T. This is due to the fact that the least-squares estimator of ϱ in regression (2.3) is biased and the errors ν_{it} in (2.3) are autocorrelated. Under fairly restrictive assumptions the following asymptotic null distribution is obtained:

Lemma 3. *Under Assumption 1 and $u_{it} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ for all i and t, it holds for fixed T and $N \to \infty$ that*

$$\left(LM + \sqrt{\frac{N}{T-1}} \right) \xrightarrow{d} \mathcal{N}\left(0, \frac{(T+1)(T-2)^2}{(T-1)^3}\right).$$

It follows that applying the usual critical values derived from the χ_1^2 distribution to the (two-sided) test statistic LM^2 yields a test with actual size tending to unity as $N \to \infty$. Note also that the limit distribution of the LM statistic tends to a standard normal distribution if $T \to \infty$ and $\sqrt{N}/T \to 0$. To obtain an asymptotically valid test for fixed T, the transformed statistic

$$\widetilde{\text{LM}} = \sqrt{\frac{(T-1)^3}{(T+1)(T-2)^2}} \left(\text{LM} + \sqrt{\frac{N}{T-1}} \right)$$

may be used, which has a standard normal limiting null distribution. An important limitation of this test statistic is, however, that the result requires the errors to be normally distributed with identical variances for all $i \in \{1, \ldots, N\}$. To obtain a test statistic that is valid in more general (and realistic) situations, we propose a regression based test statistic related to the approach suggested by Wooldridge and Drukker.

Let $\widehat{\varrho}$ denote the least-squares estimator of ϱ in the regression

$$e_{it} - \overline{e}_i \; = \; \varrho(e_{i,t-1} - \overline{e}_i) + \nu_{it} \; , \tag{2.7}$$

where $\overline{e}_i = T^{-1}\sum_{t=1}^{T} e_{it}$. It is convenient to introduce the $T \times 1$ vector of ones ι and the matrices M_1 and M_T that result from $M = I_T - T^{-1}\iota\iota'$ by dropping the first or the last row, respectively. Using this matrix notation we obtain

$$\widehat{\varrho} \;\overset{p}{\longrightarrow}\; \varrho_0 = \frac{tr(M_1'M_T)}{tr(M_T'M_T)} = -\frac{(T-1)/T}{(T-1)^2/T} = -\frac{1}{T-1} \; , \tag{2.8}$$

Thus a regression t-statistic is employed to test the modified null hypothesis H_0' : $\varrho = \varrho_0 = -1/(T-1)$. To account for autocorrelation in ν_{it}, (HAC) robust standard errors are employed, yielding the test statistic

$$\mathrm{LM}^* = \frac{\widehat{\varrho} - \varrho_0}{\widetilde{v}_\rho} \; ,$$

where

$$\widetilde{v}_\rho^2 = \frac{\sum\limits_{i=1}^{N} e_i'M_T'\widetilde{\nu}_i\widetilde{\nu}_i'M_Te_i}{\left(\sum\limits_{i=1}^{N} e_i'M_T'M_Te_i\right)^2} = \frac{\sum\limits_{i=1}^{N} e_i'M_T'(M_1 - \widehat{\varrho}M_T)e_ie_i'(M_1 - \widehat{\varrho}M_T)'M_Te_i}{\left(\sum\limits_{i=1}^{N} e_i'M_T'M_Te_i\right)^2} \; .$$

As for the WD test, we propose a simplified version of this test that is asymptotically equivalent to the statistic LM^*.

Theorem 4. *Let* $w_{Ti} = \sum\limits_{t=2}^{T}\left[(e_{it} - \overline{e}_i)(e_{i,t-1} - \overline{e}_i) + \frac{1}{T-1}(e_{i,t-1} - \overline{e}_i)^2\right]$ *and*

$$\widetilde{LM}^* \; = \; \frac{\sum\limits_{i=1}^{N} w_{Ti}}{\sqrt{\sum\limits_{i=1}^{N} w_{Ti}^2 - \frac{1}{N}\left(\sum\limits_{i=1}^{N} w_{Ti}\right)^2}} \; . \tag{2.9}$$

(i) Under Assumption 1, $T \geq 2$, and the local alternative $\rho_N = c/\sqrt{N}$, the test

statistic $\widetilde{LM}^*$ is asymptotically distributed as

$$\mathcal{N}\left(c\,\kappa\sqrt{T-3+\frac{2}{T^2-T}},1\right),$$

with κ as defined in Lemma 1.

(ii) *Under the the local alternative, the statistic $\widetilde{LM}^*$ is asymptotically equivalent to LM^*.*

2.3.3 A Modified Durbin-Watson Statistic

The pDW statistic suggested by Bhargava et al. (1982) is the ratio of the sum of squared differences and the sum of squared residuals. Instead of the ratio (which complicates the theoretical analysis), our variant of the Durbin-Watson test is based on the linear combination of the numerator and denominator of the Durbin-Watson statistic. Let

$$\delta_{Ti} = e_i' M D' D M e_i - 2\, e_i' M e_i \ ,$$

where D is a $(T-1)\times T$ matrix producing first differences, i.e.

$$D = \begin{pmatrix} -1 & 1 & 0 & \cdots & & 0 \\ 0 & -1 & 1 & \cdots & & 0 \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix},$$

and M as defined above. Using $tr(MD'DM) = 2(T-1)$ and $tr(M) = T-1$, it follows that $E(\delta_{Ti}) = 0$ for all i.

The panel test statistic is based on the normalized mean of the individual statistics:

$$\text{mDW} = \frac{1}{\widehat{s}_\delta\sqrt{N}}\sum_{i=1}^{N}\delta_{Ti} \ ,$$

where

$$\widehat{s}_\delta^2 = \frac{1}{N}\sum_{i=1}^{N}\delta_{Ti}^2 - \left(\frac{1}{N}\sum_{i=1}^{N}\delta_{Ti}\right)^2 .$$

The following theorem presents the asymptotic distribution of the test statistic.

Theorem 5. *Under Assumption 1, $T \geq 3$, and the local alternative $\rho_N = c/\sqrt{N}$, the limiting distribution is*

$$mDW \overset{d}{\longrightarrow} \mathcal{N}\left(-c\kappa\,\frac{T-1}{T}\,\sqrt{T-2}, 1\right).$$

2.3.4 Tests for Higher Order Autocorrelation

In this section, the test for first order autocorrelation is generalized to testing against autocorrelation of order p. Consider the regression

$$e_{it} - \overline{e}_i = \varrho_k(e_{i,t-k} - \overline{e}_i) + \epsilon_{it}\ , \quad t = k+1,\ldots,T; \quad i = 1,\ldots,N\ . \tag{2.10}$$

In the following lemma, we present the probability limit of the OLS estimator of ϱ_k for fixed T and $N \to \infty$.

Lemma 6. *Let $\widehat{\varrho}_k$ be the pooled OLS estimator of ϱ_k in* (2.10). *Under Assumption 1 and $\rho = 0$ we have*

$$\widehat{\varrho}_k \overset{p}{\longrightarrow} -\frac{1}{T-1}\ ,$$

for all $k \in \{1,\ldots,T-2\}$.

Using this lemma, a test for zero autocorrelation at lag k can be constructed based on the (heteroskedasticity and autocorrelation robust) t-statistic of the hypothesis $\varrho_k = -1/(T-1)$. An asymptotically equivalent test statistic is obtained by letting

$$w_{Ti}^{(k)} = \sum_{t=k+1}^{T}\left[(e_{it} - \overline{e}_i)(e_{i,t-k} - \overline{e}_i) + \frac{1}{T-1}(e_{i,t-k} - \overline{e}_i)^2\right]$$

and

$$\widetilde{\mathrm{LM}}_k^* = \frac{\sum_{i=1}^{N} w_{Ti}^{(k)}}{\sqrt{\sum_{i=1}^{N} w_{Ti}^{(k)2} - \frac{1}{N}\left(\sum_{i=1}^{N} w_{Ti}^{(k)}\right)^2}} . \tag{2.11}$$

Along the lines of the proof of Theorem 4, it is not difficult to show that, under the null hypothesis, $\widetilde{\mathrm{LM}}_k^*$ has a standard normal limiting distribution.

In empirical practice it is, however, more interesting to test the hypothesis that the errors are not autocorrelated *up to order p*. Unfortunately, the probability limit of the autoregressive coefficients $\phi_1, \ldots, \phi_p$ in the autoregression

$$e_{it} - \overline{e}_i = \phi_1(e_{i,t-1} - \overline{e}_i) + \cdots + \phi_p(e_{i,t-p} - \overline{e}_i) + \upsilon_{it}$$

is a much more complicated function of T, since the regressors are mutually correlated. Instead of presenting the probability limit, we therefore suggest a simple transformation of the regressors such that the autoregressive coefficients can be estimated unbiasedly under the null hypothesis. Specifically, we propose an implicit correction that eliminates the nonzero mean. In matrix notation, the (transformed) autoregression is given by

$$Me_i = \phi_1 L_1 Me_i + \phi_2 L_2 Me_i + \cdots + \phi_p L_p Me_i + \upsilon_i , \tag{2.12}$$

where

$$L_k = \begin{pmatrix} \psi_k/T & 0 & & \cdots & 0 \\ 0 & \psi_k/T & & \cdots & 0 \\ \vdots & & & & \vdots \\ 1 & \cdots & 0 & \ddots & \\ \vdots & \ddots & \vdots & & \\ 0 & \cdots & 1 & & \psi_k/T \end{pmatrix} ,$$

$k \in \{1, \ldots, p\}$. Note that the matrix L_k (i) produces lagged values of the (mean-adjusted) residuals and (ii) includes further terms to yield $E(e_i' M L_k M e_i) = 0$. Let

$$L_k M e_i \ = \ \begin{pmatrix} 0 \\ \vdots \\ 0 \\ e_{i1} - \overline{e}_i \\ \vdots \\ e_{i,T-k} - \overline{e}_i \end{pmatrix} + (\psi_k / T)\, e_i \ ,$$

where the first vector on the right hand side represents the lagged values of the demeaned residuals, whereas the second vector is introduced to remove the bias. It is not difficult to show that

$$E(e_i' M L_k M e_i) \ = \ 0 \quad \text{for} \ \ \psi_k = \frac{T-k}{T-1} \ ,$$

and all $k = 1, \ldots, p$ and, therefore, under the null hypothesis, the least-squares estimators of $\phi_1, \ldots, \phi_p$ converge to zero in probability for all T and $N \to \infty$. The null hypothesis $\phi_1 = \cdots = \phi_p = 0$ can therefore be tested using the associated Wald statistic

$$Q(p) \ = \ \widehat{\phi}' \widehat{S}^{-1} \widehat{\phi} \ , \qquad (2.13)$$

where

$$\widehat{S} = \left(\sum_{i=1}^{N} Z_i' Z_i \right)^{-1} \left(\sum_{i=1}^{N} Z_i' \widehat{v}_i \widehat{v}_i' Z_i \right) \left(\sum_{i=1}^{N} Z_i' Z_i \right)^{-1}$$

is the HAC estimator for the covariance matrix of the vector of coefficients with $\widehat{\phi} = \left[\widehat{\phi}_1, \ldots, \widehat{\phi}_p \right]'$, $Z_i = [L_1 M e_i, \ldots, L_p M e_i]$ and $\widehat{v}_i = M e_i - Z \widehat{\phi}$. The $Q(p)$ statistic is asymptotically equivalent to the statistic

$$\widetilde{Q}(p) \ = \ \sum_{i=1}^{N} e_i' M Z_i \left[\sum_{i=1}^{N} Z_i' M e_i e_i' M Z_i - \frac{1}{N} \left(\sum_{i=1}^{N} Z_i' M e_i \right) \left(\sum_{i=1}^{N} e_i' M Z_i \right) \right]^{-1} \sum_{i=1}^{N} Z_i' M e_i \ .$$

Along the lines of Theorem 4, it is straightforward to show that the statistics $Q(p)$ and $\widetilde{Q}(p)$ are asymptotically equivalent and have an asymptotic χ^2 distribution with p degrees of freedom.

2.3.5 A Heteroskedasticity Robust Test Statistic

An important drawback of all test statistics considered so far is that they are not robust against time dependent heteroskedasticity. This is due to the fact that the implicit or explicit bias correction of the autocovariances depends on the error variances. To overcome this drawback of the previous test statistics, we construct an unbiased estimator of the autocorrelation coefficient. The idea is to apply backward and forward transformations such that the products of the transformed series are uncorrelated under the null hypothesis. Specifically, we employ the following transformations for eliminating the individual effects:

$$
\begin{aligned}
z_{it}^f &= e_{it} - \frac{1}{T-t+1}\left(e_{it} + \dots + e_{iT}\right) \\
z_{it}^b &= e_{it} - \frac{1}{t}\left(e_{i1} + \dots + e_{it}\right) \ .
\end{aligned}
$$

The hypothesis can be tested based on the regression

$$
z_{it}^f = \psi z_{i,t-1}^b + \omega_{it} \ , \quad t = 3, \dots, T-1 \ , \tag{2.14}
$$

or, in matrix notation,

$$
V_0 e_i = \psi V_1 e_i + \eta_i \ ,
$$

where the $(T-3) \times T$ matrices V_0 and V_1 are defined as

$$
\begin{bmatrix}
-\frac{1}{2} & \frac{1}{2} & 0 & 0 & \dots & 0 & 0 & 0 \\
-\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & 0 & \dots & 0 & 0 & 0 \\
\vdots & & & & & & & \vdots \\
-\frac{1}{T-2} & -\frac{1}{T-2} & -\frac{1}{T-2} & -\frac{1}{T-2} & \dots & \frac{T-3}{T-2} & 0 & 0
\end{bmatrix}
$$

and

$$
\begin{bmatrix}
0 & 0 & \frac{T-3}{T-2} & -\frac{1}{T-2} & -\frac{1}{T-2} & \dots & -\frac{1}{T-2} & -\frac{1}{T-2} \\
0 & 0 & 0 & \frac{T-4}{T-3} & -\frac{1}{T-3} & \dots & -\frac{1}{T-3} & -\frac{1}{T-3} \\
\vdots & & & & & & & \vdots \\
0 & 0 & 0 & 0 & 0 & \dots & \frac{1}{2} & -\frac{1}{2}
\end{bmatrix} .
$$

Under the null hypothesis of no (first order) autocorrelation, we have $\psi = 0$. Since the error term η_i is heteroskedastic and autocorrelated, we again employ robust (HAC) standard errors. The resulting test statistic is equivalent to the test statistic

$$\widetilde{\mathrm{HR}} = \frac{\sum\limits_{i=1}^{N} \zeta_{Ti}}{\sqrt{\sum\limits_{i=1}^{N} \zeta_{Ti}^2 - \frac{1}{N}\left(\sum\limits_{i=1}^{N} \zeta_{Ti}\right)^2}},$$

where $\zeta_{Ti} = \sum_{t=3}^{T-1} z_{it}^f z_{i,t-1}^b$. Since ζ_{Ti} is independent of ζ_{Tj} for all $i \neq j$, it follows from the central limit theorem for independent random variables that $\widetilde{\mathrm{HR}}$ has a standard normal limit distribution.

2.4 Monte Carlo Study

This section presents the finite sample performance of the test statistics for serial correlation and assesses the reliability of the asymptotic results derived in Section 3. We also conduct experiments with different forms of serial correlation and heteroskedasticity.

The benchmark data generating process for all simulations is a linear panel data model of the form

$$y_{it} = x_{it}\beta + \mu_i + u_{it},$$

where $i = 1, \ldots, N$ and $t = 1, \ldots, T$. We set β to 1 in all simulations and draw the individual effects μ_i from a $\mathcal{N}(0, 2.5^2)$ distribution. To create correlation between the regressor and the individual effect, we follow Drukker (2003) by drawing x_{it}^0 from a $\mathcal{N}(0, 1.8^2)$ distribution and computing $x_{it} = x_{it}^0 + 0.5\mu_i$. The regressor is drawn once and then held constant for all experiments. In our benchmark model, the disturbance term follows an autoregressive process of order 1,

$$u_{it} = \alpha u_{i,t-1} + \varepsilon_{it} \ ,$$

where $\varepsilon_{it} \sim \mathcal{N}(0, 1)$ and we discard the first 100 observations to eliminate the influence of the initial value.

Table 2.1 presents the results of the local power analysis. The empirical values are computed using 10,000 Monte Carlo simulations, while the asymptotic values (given in parentheses) are computed according to the results given in Theorems 2, 4, and 5. All tests exhibit good size control (i.e. for $c = 0$).[2] They reject the null hypothesis of no serial correlation slightly more often than the nominal size of 5% but are never above 6%. In all cases, the LM statistic has superior power compared to the competing statistics. While the modified Durbin-Watson statistic is close to the LM statistic, the Wooldridge-Drukker test has considerably less power.

To evaluate the small-sample properties of the test for higher-order autocorrelation $\widetilde{Q}(p)$, we simulate a model with disturbances following an AR(2) process:

$$u_{it} = \alpha_1 u_{i,t-1} + \alpha_2 u_{i,t-2} + \varepsilon_{it} \ ,$$

where again $\varepsilon_{it} \sim \mathcal{N}(0,1)$ and we discard the first 100 observations to eliminate the influence of the initial value. The results for three different AR(2) processes are presented in Table 2.2. The $\widetilde{Q}(2)$ statistic has good power properties in all configurations considered. Not surprisingly, the Wooldridge-Drukker test shows considerable power when $\alpha_1 = 0.1$ and $\alpha_2 = -0.1$ or $\alpha_1 = 0$ and $\alpha_2 = 0.2$. As mentioned in Remark 5, the WD (and $\widetilde{\mathrm{WD}}$) statistic is a test of the *difference* between the first and second order autocorrelation of the errors. As a consequence, it also has power against most AR(2) alternatives. However, choosing $\alpha_1 = \alpha_2 = 0.1$, which implies equal first- and second-order autocorrelations of u_{it}, leads to a complete loss of power of the Wooldridge-Drukker test, confirming our assertion in Remark 5. The LM statistic has less power than the $\widetilde{Q}(2)$ statistic, in particular for the setup with $\alpha_1 = 0$ and $\alpha_2 = 0.2$.

We consider four different types of time-dependent heteroskedasticity by using the following error process:

$$u_{it} = \sqrt{h_t}\varepsilon_{it} \ ,$$

where $\varepsilon_{it} \sim \mathcal{N}(0,1)$. The first type is a break in the variance function, i.e. $h_t = 10$ for $t = 1, \ldots, T/5$ and $h_t = 1$ otherwise. The second specification is a U-shaped

[2] The cross-sectional dimension is held constant at $N = 100$. Values of $c = 0$, $c = 0.5$, and $c = 1$ therefore imply $\alpha = 0$, $\alpha = 0.05$, and $\alpha = 0.1$, respectively.

variance function of the form $h_t = (t - T/2)^2 + 1$. The third and fourth types of heteroskedasticity are positive and negative exponential variance functions, i.e. $h_t = \exp(-0.5t)$ and $h_t = \exp(0.5t)$, respectively. Table 2.3 shows that the modified Durbin-Watson test has massive size distortions under all four types of heteroskedasticity. While the Wooldridge-Drukker test has the proper size in case of a U-shaped variance function, the break in the variance function and the exponential variance functions lead to large size distortions. The modified LM statistic does surprisingly well under heteroskedasticity, but there are some size distortions when T is small, especially for the negative exponential variance function. On the other hand our proposed heteroskedasticity-robust test statistic retains its correct size under all types of heteroskedasticity considered here.

2.5 Conclusion

This chapter proposes various tests for serial correlation in fixed-effects panel data regression models with a small number of time periods. First, a simplified version of the test suggested by Wooldridge (2002) and Drukker (2003) is considered. The second test is based on the LM statistic suggested by Baltagi and Li (1995), and the third test is a modification of the classical Durbin-Watson statistic. Under the null hypothesis of no serial correlation, all tests possess a standard normal limiting distribution as $N \to \infty$ and T is fixed. Analyzing the local power of the tests, we find that the LM statistic has superior power properties. We also propose a generalization to test for autocorrelation up to some given lag order and a test statistic that is robust against time dependent heteroskedasticity. Monte Carlo simulations show that the proposed tests have favorable size and power properties compared to the popular Wooldridge-Drukker test.

Appendix to Chapter 2

Proof of Lemma 1:

Let $y_i = [y_{i1}, \ldots, y_{iT}]'$ and $X_i = [x_{i1}, \ldots, x_{iT}]'$. The residual vector can be written as

$$\widehat{e}_i = e_i - X_i(\widehat{\beta} - \beta) \ .$$

It follows that

$$\sum_{i=1}^{N} \widehat{e}_i' A_T \widehat{e}_i = \sum_{i=1}^{N} e_i' A_T e_i - 2\sum_{i=1}^{N} e_i' A_T X_i(\widehat{\beta} - \beta) + \sum_{i=1}^{N} (\widehat{\beta} - \beta)' X_i' A_T X_i (\widehat{\beta} - \beta) \ .$$

For the estimation error of $\widehat{\beta}$, we obtain

$$\begin{aligned}
\widehat{\beta} - \beta &= \left(\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)' \right)^{-1} \sum_{t=1}^{T} \left(\frac{1}{N} \sum_{i=1}^{N} (x_{it} - \overline{x}_i) u_{it} \right) \\
&= \left(\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)' \right)^{-1} \sum_{t=1}^{T} \xi_{Nt} \ ,
\end{aligned}$$

where $\xi_{Nt} = N^{-1} \sum_{i=1}^{N} (x_{it} - \overline{x}_i) u_{it}$. Assumption 1 implies

$$\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)' \ \xrightarrow{p} \ \overline{C}$$

and $\xi_{Nt} = O_p(N^{-1/2})$. It follows that $\widehat{\beta} - \beta = O_p(N^{-1/2})$. Furthermore,

$$\begin{aligned}
\| \frac{1}{N} \sum_{i=1}^{N} X_i' A_T X_i \| &\leq \| A_T \| \frac{1}{N} \sum_{i=1}^{N} \| X_i' X_i \| = O_p(1) \\
\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i' A_T e_i \| &\leq \| A_T \| \cdot \| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i' e_i \| = O_p(1) \ .
\end{aligned}$$

Using $e_i' A_T e_i = (\mu_i \iota' + u_i') A_T (\mu_i \iota + u_i) = u_i' A_T u_i$, it follows that

$$\sum_{i=1}^{N} \widehat{e}_i' A_T \widehat{e}_i = \sum_{i=1}^{N} u_i' A_T u_i + O_p(N^{-1/2}) \ .$$

Under the sequence of local alternatives, we have

$$\Omega = E(u_i u_i') = \frac{\sigma_i^2}{1-\rho_N^2} \begin{pmatrix} 1 & \rho_N & \rho_N^2 & \cdots & \rho_N^{T-1} \\ \rho_N & 1 & \rho_N & \cdots & \rho_N^{T-2} \\ \vdots & & & & \vdots \\ \rho_N^{T-1} & \rho_N^{T-2} & \rho_N^{T-3} & \cdots & 1 \end{pmatrix} = I_T + \frac{c\,\sigma_i^2}{\sqrt{N}} H_T + O(N^{-1}) \ ,$$

where H_T is a band matrix with ones on the first off-diagonals and zeros elsewhere. Since $tr(A_T) = 0$ it follows that

$$E(u_i' A_T u_i) = E(A_T \Omega) = \frac{c\sigma_i^2}{\sqrt{N}} tr(A_T H_T) + O(N^{-1}) \ .$$

Using standard results for the variance of quadratic forms (e.g. Searle, 1971), we have

$$var(u_i' A_T u_i) = \sigma_i^4 \, tr(A_T^2 + A_T' A_T) + O(N^{-1/2}) \ .$$

It follows from the Lindeberg-Feller central limit theorem

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N u_i' A_T u_i \ \xrightarrow{p} \ \mathcal{N}\Big(c\, m_2 tr(A_T H_T), m_4 \, tr(A_T^2 + A_T' A_T)\Big),$$

where m_2 and m_4 are defined in Lemma 1.

For the second moment we obtain from the law of large numbers

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N (\widehat{e}_i' A_T \widehat{e}_i)^2 \quad &= \quad \frac{1}{N} \sum_{i=1}^N \Big(u_i' A_T u_i + O_p(N^{-1/2})\Big)^2 \\ &= \quad \frac{1}{N} \sum_{i=1}^N (u_i' A_T u_i)^2 + O_p(N^{-1/2}) \\ &\xrightarrow{p} \quad m_4 \, tr(A_T^2 + A_T' A_T) \ . \end{aligned}$$

Furthermore, under the sequence of local alternatives

$$\frac{1}{N} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \widehat{e}_i' A_T \widehat{e}_i \right)^2 \ = \ O_p(N^{-1})$$

and, therefore, the second term of the denominator in (2.4) does not affect the limiting distribution.

It follows that the limiting distribution of the test statistic is given by

$$\lambda_{NT} \;=\; \frac{\frac{1}{\sqrt{N}}\sum\limits_{i=1}^{N} u_i' A_T u_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(u_i' A_T u_i\right)^2}} + o_p(1) \;\stackrel{d}{\longrightarrow}\; \mathcal{N}\left(c\,\frac{\kappa tr(A_T H_T)}{\sqrt{tr(A_T^2 + A_T' A_T)}}, 1\right) \;.$$

Finally, we have

$$\begin{aligned}
tr(A_T H_T) &= \sum_{t=2}^{T} a_{t,t-1} + a_{t-1,t} \\
tr(A_T' A_T) &= \sum_{t=2}^{T}\sum_{s=1}^{T} a_{ts}^2 \\
tr(A_T^2) &= \sum_{t=2}^{T}\sum_{s=1}^{T} a_{ts} a_{st} \;,
\end{aligned}$$

which gives the limiting distribution as stated in Lemma 1.

Proof of Theorem 2:

(ii) Define the $(T-1)\times T$ first-difference matrix

$$D = \begin{pmatrix} -1 & 1 & 0 & \cdots & & 0 \\ 0 & -1 & 1 & \cdots & & 0 \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

and let D_k denote a $(T-2)\times T$ matrix that results from dropping the k'th row of D so that

$$z_{Ti} \;=\; e_i'(D_T' D_1 + 0.5 D_T' D_T) e_i \;\equiv\; e_i' A_T e_i \;,$$

where

$$A_T = \begin{bmatrix} 0.5 & -0.5 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & -0.5 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1.5 & 0 & -0.5 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1.5 & 0 & -0.5 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1.5 & 0 & -0.5 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & & & & & \vdots & \cdots & \vdots & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 1.5 & 0 & -0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1.5 & 0 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 1.5 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & -1 & 1 & 0 \end{bmatrix}.$$

If $u_{it} \overset{iid}{\sim} N(0, \sigma_i^2)$, the variance of the quadratic form $e_i' A_T e_i = u_i' A_T u_i$ is $\sigma_i^4 tr(A_T^2 + A_T' A_T)$, where

$$\begin{aligned} tr(A_T^2) &= -\frac{3}{2}(T-3) \\ tr(A_T' A_T) &= 3 + \frac{7(T-3)}{2} \; . \end{aligned}$$

It follows that $var(e_i' A_T e_i) = \sigma_i^4 [2(T-3)+3]$. Furthermore,

$$tr(A_T H_T) = T - 2$$

and, by using Lemma 1, we obtain

$$\widetilde{\mathrm{WD}} \xrightarrow{d} \mathcal{N}\left(c \frac{\kappa(T-2)}{\sqrt{2(T-3)+3}}, 1 \right).$$

(ii) The original version of the WD test can be written as

$$
\begin{aligned}
\text{WD} &= \frac{\sum_{i=1}^{N} z_{Ti}}{\sqrt{\sum_{i=1}^{N}\left(\sum_{t=3}^{T}\Delta e_{i,t-1}(\Delta e_{it}-\widehat{\theta}\Delta e_{i,t-1})\right)^2}} \\
&= \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_{Ti}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=3}^{T}\Delta e_{i,t-1}(\Delta e_{it}+0.5\Delta e_{i,t-1})+(\widehat{\theta}+0.5)\Delta e_{t-1}^2\right)^2}} \\
&= \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_{Ti}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=3}^{T}\Delta e_{i,t-1}(\Delta e_{it}+0.5\Delta e_{i,t-1})\right)^2}} + O_p(N^{-1/2}) \\
&= \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_{Ti}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} z_{Ti}^2}} + O_p(N^{-1/2}) ,
\end{aligned}
$$

where $\widehat{\theta}+0.5 = O_p(N^{-1/2})$ under the null hypothesis and local alternative. Furthermore, by using the results in the proof of Lemma 1:

$$
\begin{aligned}
\widetilde{\text{WD}} &= \frac{\sum_{i=1}^{N} z_{Ti}}{\sqrt{\sum_{i=1}^{N} z_{Ti}^2 - \frac{1}{N}\left(\sum_{i=1}^{N} z_{Ti}\right)^2}} \\
&= \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_{Ti}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} z_{Ti}^2}} + O_p(N^{-1/2})
\end{aligned}
$$

It follows that $\text{WD} - \widetilde{\text{WD}} = O_p(N^{-1/2})$.

Proof of Lemma 3:

For the first order autocorrelation, Cox and Solomon (1988) derive the following asymptotic approximation:

$$\sqrt{N}\left(\frac{\sum_{i=1}^{N}\sum_{t=2}^{T}(u_{it}-\overline{u}_i)(u_{i,t-1}-\overline{u}_i)}{\sum_{i=1}^{N}\sum_{t=1}^{T}(u_{it}-\overline{u}_i)^2}\right) \xrightarrow{d} \mathcal{N}\left(-\frac{1}{T}, \frac{(T+1)(T-2)^2}{T^2(T-1)^2}\right) .$$

Using this result, the limiting distribution of the LM statistic is easily derived.

Proof of Theorem 4:

(i) Using the matrix notation introduced in section 2, we have

$$w_{Ti} \;=\; u_i'\left(M_1'M_T + \frac{1}{T-1}M_T'M_T\right)u_i \;\equiv\; u_i'A_Tu_i \; .$$

The variance of w_{Ti} is obtained as $var(w_{Ti}) = \sigma_i^4 tr(A^2 + AA')$ by using the following results:

$$\begin{aligned}
tr(M_1'M_TM_1'M_T) &= -(T^2-2T-1)/T^2\\
tr(M_1'M_TM_T'M_1) &= tr(M_T'M_TM_T'M_T) = (T^3-2T^2+1)/T^2\\
tr(M_1'M_TM_T'M_T) &= tr(M_T'M_TM_1'M_T) = tr(M_T'M_TM_T'M_1) = -(T^2-T-1)/T^2 \; .
\end{aligned}$$

It follows that

$$var(w_{Ti}) = \sigma_i^4\left(T-3+\frac{2}{T(T-1)}\right).$$

Furthermore,

$$tr(M_1'M_TH_T) = tr(S_T'MH_TMS_1) = tr(S_T'H_TS_1) - 2T^{-1}tr(S_T'\iota\iota'H_TS_1) + T^{-2}tr(S_T'\iota\iota'H_T\iota\iota'S_1) \; ,$$

where S_k, is a $T \times (T-1)$ selection matrix that eliminates the k'th row. Using

$$\begin{aligned}
&tr(S_T' H_T S_1) = T-1 \\
&tr(S_T' H_T S_T) = 0 \\
&tr(S_T' \iota\iota' H_T S_1) = tr(S_T' \iota\iota' H_T S_T) = 2(T-1)-1 \\
&tr(S_T' \iota\iota' H_T \iota\iota' S_1) = tr(S_T' \iota\iota' H_T \iota\iota' S_T) = 2(T-1)^2
\end{aligned}$$

yields

$$tr(A_T H_T) \;=\; tr\left(M_1' M_T H_T - \frac{1}{T-1} M_1' M_T H_T\right) = T - 3 + \frac{2}{T(T-1)} \;,$$

which is identical to $tr(A^2 + A'A)$. With these results the limiting distribution follows.

(ii) Using $\widehat{\varrho} + \frac{1}{T} = O_p(N^{-1/2})$, the LM* statistic can be written as

$$\begin{aligned}
\text{LM}^* \;&=\; \frac{\sum\limits_{i=1}^{N} w_{Ti}}{\sqrt{\sum\limits_{i=1}^{N}\left[\sum\limits_{t=2}^{T}(e_{i,t-1}-\overline{e}_i)\Big(e_{it}-\widehat{\varrho}e_{i,t-1}-(1-\widehat{\varrho})\overline{e}_i\Big)\right]^2}} \\
&=\; \frac{\frac{1}{\sqrt{N}}\sum\limits_{i=1}^{N} w_{Ti}}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N}\left[\sum\limits_{t=2}^{T}(e_{i,t-1}-\overline{e}_i)(e_{it}+\frac{1}{T}e_{i,t-1}-\frac{T+1}{T}\overline{e}_i)+(\widehat{\varrho}+\frac{1}{T})(e_{i,t-1}-\overline{e}_i)^2\right]^2}} \\
&=\; \frac{\frac{1}{\sqrt{N}}\sum\limits_{i=1}^{N} w_{Ti}}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N} w_{Ti}^2}} + O_p(N^{-1/2})
\end{aligned}$$

Furthermore, under the null hypothesis and local alternative

$$\widetilde{\text{LM}}^* \;=\; \frac{\frac{1}{\sqrt{N}}\sum\limits_{i=1}^{N} w_{Ti}}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N} w_{Ti}^2}} + O_p(N^{-1/2}) \;.$$

Proof of Theorem 5:

Let

$$\delta_{Ti} \;=\; e_i' M(D'D - 2I_T)Me_i \;=\; u_i' A_T u_i \;,$$

where

$$A_T = M \begin{bmatrix} -1 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & -1 \end{bmatrix} M$$

is a symmetric $T \times T$ matrix. The variance is obtained as

$$\begin{aligned} var(\delta_{Ti}) \;&=\; 2\sigma_i^4 tr(A_T^2) \\ &=\; 4\sigma_i^4 (T-2) \;. \end{aligned}$$

Furthermore,

$$tr(A_T H_T) \;=\; -\frac{T-1}{T} 2(T-2) \;.$$

It follows from Lemma 1 that under the local alternative

$$\text{mDW} \;\xrightarrow{d}\; \mathcal{N}\left(c\,\kappa \frac{T-1}{T}\,\sqrt{T-2}, 1\right) .$$

Proof of Lemma 6:

Rewrite equation (2.10) in matrix terms as

$$\widetilde{L}_0 M e_i = \varrho_k \widetilde{L}_k M e_i + \epsilon_i \;,$$

where

$$\widetilde{L}_0 = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \quad \widetilde{L}_k = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots & 0 & \cdots & 0 \\ \vdots & & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$$

are $(T-k) \times T$ matrices. Under Assumption (1) and using the following results:

$$\begin{aligned} tr(M\widetilde{L}_k'\widetilde{L}_0 M) &= tr\left[\left(\widetilde{L}_k' - \frac{1}{T}\iota\iota'\widetilde{L}_k'\right)\widetilde{L}_0\right] \\ &= tr\left(\widetilde{L}_k'\widetilde{L}_0\right) - \frac{1}{T}tr\left(\iota'\widetilde{L}_k'\widetilde{L}_0\iota\right) \\ &= -\frac{T-k}{T} \ , \\ tr(M\widetilde{L}_k'\widetilde{L}_k M) &= tr\left[\left(\widetilde{L}_k' - \frac{1}{T}\iota\iota'\widetilde{L}_k'\right)\widetilde{L}_k\right] \\ &= tr\left(\widetilde{L}_k'\widetilde{L}_k\right) - \frac{1}{T}tr\left(\iota'\widetilde{L}_k'\widetilde{L}_k\iota\right) \\ &= (T-1)\frac{T-k}{T} \ , \end{aligned}$$

it follows as $N \to \infty$:

$$\widehat{\varrho}_k \ \xrightarrow{p} \ \frac{tr(M\widetilde{L}_k'\widetilde{L}_0 M)}{tr(M\widetilde{L}_k'\widetilde{L}_k M)} \ = \ = \ -\frac{1}{T-1} \ ,$$

for all $k \in \{1, \ldots, T-2\}$.

Table 2.1: Size and power of alternative tests

T	WD	$\widetilde{\mathrm{WD}}$	LM*	$\widetilde{\mathrm{LM}}^*$	mDW
			$c = 0$		
10	0.059	0.057	0.056	0.053	0.055
20	0.058	0.057	0.059	0.056	0.057
50	0.059	0.059	0.057	0.056	0.055
100	0.054	0.054	0.053	0.053	0.054
			$c = 0.5$		
10	0.172	0.187	0.272	0.264	0.253
	(0.163)	(0.163)	(0.263)	(0.263)	(0.247)
20	0.321	0.338	0.545	0.541	0.527
	(0.316)	(0.316)	(0.541)	(0.541)	(0.522)
50	0.675	0.687	0.925	0.924	0.921
	(0.683)	(0.683)	(0.929)	(0.929)	(0.924)
100	0.933	0.936	0.998	0.998	0.998
	(0.937)	(0.937)	(0.998)	(0.998)	(0.998)
			$c = 1$		
10	0.487	0.508	0.746	0.739	0.713
	(0.492)	(0.492)	(0.755)	(0.755)	(0.721)
20	0.820	0.832	0.983	0.982	0.979
	(0.841)	(0.841)	(0.985)	(0.985)	(0.981)
50	0.997	0.997	1.000	1.000	1.000
	(0.998)	(0.998)	(1.000)	(1.000)	(1.000)
100	1.000	1.000	1.000	1.000	1.000
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Note: $N = 100$. Empirical values are computed using 10,000 Monte Carlo simulations. Asymptotic values (given in parentheses) are computed according to the results in Theorems 2 – 5.

Table 2.2: AR(2) processes

T	WD	$\widetilde{\text{WD}}$	LM*	$\widetilde{\text{LM}}^*$	mDW	$\widetilde{Q}(2)$
			$\alpha_1 = 0.1$, $\alpha_2 = -0.1$			
10	0.781	0.798	0.654	0.635	0.589	0.741
20	0.984	0.986	0.902	0.897	0.882	0.978
50	1.000	1.000	0.999	0.999	0.998	1.000
100	1.000	1.000	1.000	1.000	1.000	1.000
			$\alpha_1 = \alpha_2 = 0.1$			
10	0.063	0.059	0.293	0.276	0.288	0.438
20	0.057	0.056	0.732	0.722	0.722	0.922
50	0.061	0.060	0.996	0.996	0.995	1.000
100	0.061	0.060	1.000	1.000	1.000	1.000
			$\alpha_1 = 0$, $\alpha_2 = 0.2$			
10	0.776	0.714	0.188	0.173	0.144	0.870
20	0.983	0.974	0.116	0.109	0.101	0.999
50	1.000	1.000	0.082	0.080	0.077	1.000
100	1.000	1.000	0.072	0.071	0.070	1.000

Note: $N = 50$. Rejection frequencies are computed using 10,000 Monte Carlo replications.

Table 2.3: Size under temporal heteroskedasticity

T	$\widetilde{\text{WD}}$	$\widetilde{\text{LM}}^*$	mDW	HR	$\widetilde{\text{WD}}$	$\widetilde{\text{LM}}^*$	mDW	HR
	break in variance				u-shaped variance			
10	0.946	0.084	0.438	0.059	0.056	0.064	0.798	0.057
20	0.340	0.065	0.211	0.058	0.061	0.059	0.601	0.059
50	0.145	0.055	0.102	0.057	0.060	0.057	0.319	0.057
100	0.101	0.057	0.081	0.057	0.059	0.058	0.197	0.057
	negative exponential trend				positive exponential trend			
10	0.855	0.101	0.445	0.059	0.343	0.058	0.454	0.054
20	0.868	0.072	0.719	0.057	0.361	0.059	0.723	0.058
50	0.868	0.055	0.834	0.058	0.372	0.058	0.836	0.058
100	0.871	0.055	0.867	0.056	0.377	0.060	0.862	0.058

Note: $N = 50$. Rejection frequencies are computed using 10,000 Monte Carlo replications.

Chapter 3

Policy Risk and the Business Cycle

3.1 Introduction

The supposedly negative influence of "policy risk", i.e. uncertainty about fiscal and monetary policy, has become a recurring theme in the political discourse. The popular argument espoused in speeches and newspaper articles by politicians and economists alike is that the uncertainty surrounding future policy stuns economic activity by inducing a "wait-and-see approach".[1] In the following, we think of uncertainty as the dispersion of the economic shock distribution. Rational consumers and firms will react to the fact that future shocks will be drawn from a wider distribution. This reaction is distinct from the ex-post effect of higher uncertainty resulting from on average more extreme shock realizations.[2] The goal of the present study is to isolate the first effect and answer the question: Are uncertainty shocks to policy variables quantitatively important?

[1] See e.g. The Wall Street Journal, October 29th, 2009: "*For these small businesses, and many others [...], there's an additional dark cloud: uncertainty created by Washington's bid to reorganize a wide swath of the U.S. economy.*" (Fields, 2009). For other proponents of this view, see Boehner (2010); Cantor (2010); Imrohoroglu (2010); Lowrie (2010); McKinnon (2010); see Klein (2010); Reeve (2010); Wingfield (2010) for dissenting opinions.

[2] Uncertainty shocks are mean preserving spreads to the shock distribution. They are not associated with the expectation of shocks going into a specific direction, like expecting an expansionary stimulus package. Hence, they are also distinct from news shocks (Beaudry and Portier, 2006; Schmitt-Grohé and Uribe, 2010), which are future level shocks of which both the sign and the magnitude are already perfectly known today.

Clearly, during the so-called Great Recession U.S. citizens were facing a period of extraordinary uncertainty regarding economic policy. On the one hand, both the output decline due to the financial crisis and the fiscal stimuli designed to counteract this decline had led to a considerable deterioration of the U.S. fiscal situation. Given this unsustainable fiscal path, many commentators and politicians were arguing for a quick consolidation of government finances, possibly by raising taxes. On the other hand, the U.S. unemployment rate stood at 9.6% at the end of 2010, its highest value since 1983. Hence, there were considerable calls for more fiscal stimulus, preferably in the form of reduced taxes due to supposedly higher multipliers (see e.g. Romer and Romer, 2010). At the same time, Republicans and Democrats were fighting over the continuation of the Bush tax cuts. On the monetary side, the amount of policy risk was equally high. Hawks and doves at the Federal Reserve System fought over the extent of quantitative easing and the correct monetary stance given conflicting signals from core and headline inflation measures.

Scientific evidence on the aggregate effects of uncertainty is still inconclusive and mostly confined to TFP uncertainty. Empirical studies using different proxies and identification schemes to uncover the effects of uncertainty have produced a variety of results. One group of studies reports an important impact of uncertainty about productivity on real aggregate variables like GDP and employment (Alexopoulos and Cohen, 2009; Bloom, 2009; Bloom et al., 2010). A one-standard deviation shock to uncertainty in these studies typically leads to a 1%-2% drop in GDP, followed by a recovery with a considerable overshooting. In contrast, a second group of studies reports little to no impact at all (Bachmann and Bayer, 2011; Bachmann et al., 2010; Bekaert et al., 2010; Chugh, 2011; Popescu and Smets, 2010). In the theoretical literature, while most studies have emphasized the contractionary effects of uncertainty on economic activity, it is generally acknowledged that there are different effects working in opposite directions, thereby making the overall effect ambiguous. For example, while an increase in uncertainty may depress investment due to a "wait-and-see approach", economic agents may want to self-insure by working more to build up a buffer capital stock, which ceteris paribus leads to an increase in investment.

We answer the question of whether policy risk shocks are quantitatively important in an estimated DSGE-model. We focus on aggregate uncertainty as it has been shown to have potentially important output effects (Fernández-Villaverde et al., forthcoming). We add to the previous literature in the following ways. First, we are to our knowledge the first to study the effect of policy risk on business cycles.[3] Second, we directly measure aggregate uncertainty from the respective time series without the need to resort to proxies. Third, we jointly consider level shocks and uncertainty shocks. Regarding uncertainty shocks, we focus on policy risk, i.e. uncertainty about future tax liabilities, government spending, and monetary policy, to test the hypothesis that policy risk may be an important factor in explaining the prolonged Great Recession. We also include uncertainty with respect to total factor productivity (TFP) and investment-specific technology in order to have a benchmark against which we can judge our findings. Fourth, we integrate these processes into a medium-scale New Keynesian DSGE-model of the type typically used for policy analysis (see e.g. Christiano et al., 2005; Smets and Wouters, 2007) and solve this model using third-order perturbation methods. We then estimate the model using the Simulated Method of Moments. This approach allows us to control for the effects of level shocks to TFP, investment-specific technology, government spending, monetary policy, and taxes when estimating the importance of policy risk.

We find that the role of policy risk in explaining the prolonged slump is largely overstated. Although the output effects of policy risk are an order of magnitude larger than the effects of TFP uncertainty, even a large (two-standard deviation) shock to policy risk decreases output by a mere 0.025%. The reason for this result is the existence of strong general equilibrium effects that dampen the effects of aggregate uncertainty and imply a low shock amplification. Most notably, monetary policy reacts fast and decisively to current economic conditions, implying an interest rate response that dampens aggregate fluctuations arising from uncertainty shocks. If we

[3] We have recently become aware of independently conducted work by Fernández-Villaverde et al. (2011), studying a similar issue in a calibrated model. The studies differ in the set of shocks considered and in the details of the model specification. However, the results are quite similar, with even large uncertainty shocks generating only a contained output decline. In their baseline calibration, a two-standard deviation policy risk shock decreases output by 0.06% compared to 0.025% in our estimated baseline specification. The advantage of our approach is that we estimate the parameters of our model. Moreover, we allow for time-varying volatility in technology, allowing us to relate our findings to the literature on TFP uncertainty and to "good luck" explanations of the Great Moderation.

allow for a stronger amplification, uncertainty shocks generate considerably larger output effects, but at the same time imply counterfactually volatile business cycles.

From a methodological viewpoint, the paper most closely related to our work is Fernández-Villaverde et al. (forthcoming). Their study also employs Sequential Monte Carlo Methods combined with third-order perturbation to estimate the effect of interest risk on the Argentinean economy. In terms of results, our paper is most closely related to Bachmann and Bayer (2011), who show for the case of idiosyncratic uncertainty about technology that general equilibrium effects may considerably reduce the effect of uncertainty shocks typically found in partial equilibrium models (e.g. Bloom, 2009). Our paper is also related to the work of Primiceri (2005) and Justiniano and Primiceri (2008). Using a time-varying Bayesian VAR and an estimated DSGE-model, respectively, the authors document the importance of time-varying volatility for explaining the time series behavior of output and inflation and the Great Moderation in particular. We differ from their work in two major points: first, we allow for a non-linear transmission of volatility shocks into the economy. Second, by using a third-order approximation instead of a first-order approximation, we are able to distinguish uncertainty-effects from the ex-post effect of uncertainty in the form of more extreme level shocks. We show that their result is mainly due to the differing size of the realized level shocks when the dispersion of the distribution from which they are drawn changes. In contrast, the pure uncertainty-effect is only of secondary importance.

The outline of the paper is as follows. Section 3.2 presents a short literature review on the transmission channels of uncertainty. In Section 3.3, we build a quantitative business cycle model featuring several channels identified in the theoretical literature through which aggregate uncertainty may impact economic activity. We measure policy risk and technological uncertainty directly from aggregate time series using Sequential Monte Carlo methods in Section 3.4. In Section 3.5, we feed the uncertainty processes estimated in Section 3.4 as driving processes into the model and fit it to U.S. data using a Simulated Method of Moments approach. With the estimated model at hand, we then study the effects of policy risk in Section 3.6. Section 3.7 concludes.

3.2 Uncertainty: Potential Transmission Channels

Three different mechanisms through which aggregate uncertainty may affect economic activity have been identified in the microeconomic literature: Hartman-Abel effects, real option effects, and precautionary savings. While these categories are helpful in shaping our thinking about the effects of uncertainty, they are partial equilibrium effects. In general equilibrium, each of these effects necessarily induces equilibrating price and quantity changes that may significantly dampen the aggregate effects. While in a partial equilibrium model uncertainty may have ceteris paribus largely contractionary effects on investment and output (e.g. Bloom, 2009), in general equilibrium wages and interest rates may adjust, thereby significantly reducing the resulting net effect (Bachmann and Bayer, 2011).

The first category are the so called Hartman-Abel-effects (Abel, 1983; Hartman, 1972). Under certain conditions,[4] it follows from the firms's FOC that the expected marginal revenue product of capital is convex in output prices and TFP.[5] Hence, due to Jensen's Inequality larger uncertainty about these variables increases the demand for capital and thus investment. In our model, while capital is predetermined, both the utilization of capital and labor input can be adjusted, opening up the possibility of expansionary Hartman-Abel effects.

Second, there may be real option effects at work (Bernanke, 1983), e.g. through investment being (partially) irreversible and/or partially expandable. For example, if the resale (purchase) price of capital in the future differs from the current acquisition price, a firm installing capital that it may sell later, effectively acquires a put option. Moreover, investment today destroys a call option, namely the opportunity to buy capital later at a possibly lower price. Hence, in the investment decision these option values have to be taken into account (Abel et al., 1996). Higher uncertainty decreases investment as the call option to purchase the capital later, which is "killed" by investing today, becomes more valuable. However, in the presence of partial

[4] Constant-returns-to-scale production function with i) a predetermined capital stock, ii) perfect competition, iii) risk neutrality, and iv) symmetric convex adjustment costs.

[5] The reason is that labor can flexibly react to shocks and hence the marginal revenue product reacts stronger than one for one to the movement in the respective variable. To see this, assume a fixed capital stock of capital and that the output price rises. There is a direct positive effect of this price increase on profits via quantity times price change. Additionally, there is a positive indirect effect through the increase in optimal output that is achieved by increasing labor.

Table 3.1: Overview: Potential transmission mechanism

	Hartman-Abel eff.	Real option effects			Precaution. sav.
		Call	Put	Interest rate	
Investment	+	−	+	+/−	+
Consumption	?	?	?	?	−

Notes: + indicates a positive effect of uncertainty, − a negative effect, and +/− an ambiguous effect on the respective variable. ? denotes that the respective effect makes no prediction for this variable due to its partial equilibrium nature.

reversibility, the value of the put option that is obtained by investing today increases with higher uncertainty. Hence, the total effect of uncertainty on investment in such a framework is generally ambiguous.

In our model, several features give rise to option effects. First, capital is predetermined for one period. Second, the relative price of investment and consumption is stochastic, thereby giving rise to potentially costly irreversibility and expandability. Third, through the presence of depreciation allowances investment generates a tax shield at historical costs of investment so that investment effectively "kills" the option to purchase this tax shield later. Fourth, the interest rate in our model is stochastic, giving rise to additional countervailing option effects as discussed in Ingersoll and Ross (1992).

The third effect is the precautionary saving motive (Leland, 1968), defined as the "additional saving that results from the knowledge that the future is uncertain" (Carroll and Kimball, 2008). Faced with higher uncertainty, agents may both consume less and work more in order to self-insure against future shocks, i.e. they build a buffer stock.[6] As the preferences of the agents in our model feature prudence (Garcia et al., 2007; Kimball, 1990) uncertainty should increase precautionary savings in our model.

In the end, due to these three effects acting on different variables and potentially working in opposite directions as well as the presence of general equilibrium effects, only a rigorous quantitative evaluation can answer the question what the net effect

[6]Real option effects and the precautionary saving motive are not disjunct effects. Consumption is completely irreversible as the consumed good is not available for consumption in later periods when the marginal utility of consumption may be high.

of uncertainty on aggregate activity is. We pursue this question by estimating a structural model featuring time-varying volatility, which we present in the next section.

3.3 A DSGE-Model with Policy Risk

We use a standard quantitative New Keynesian business cycle model (Smets and Wouters, 2007). The model economy is populated by a large representative family, a continuum of unions $j \in [0, 1]$ selling differentiated labor services to intermediate firms, a continuum of intermediate firms producing differentiated intermediate goods using bundled labor services and capital, and a final good firm bundling intermediate goods to a final good. In addition, the model features a government sector that finances government spending with distortionary taxation and transfers, and a monetary authority which sets the nominal interest rate according to an interest rate rule.

3.3.1 Household Sector

The economy is populated by a large representative family with a continuum of members, each consuming the same amount and working the same number of hours. Preferences are defined over per capita consumption C_t and per capita labor effort L_t. Following the framework in Schmitt-Grohé and Uribe (2006), labor is supplied to a continuum of unions $j \in [0, 1]$, which are monopolistically competitive and supply differentiated labor services $l_t\,(j)$. Household members supply their labor uniformly to all unions. Hence, total labor supply of the representative family is given by the integral over all labor markets j, i.e. $L_t = \int_0^1 l_t\,(j)\,dj$. The labor market structure will be discussed in more detail below. We assume the preference specification of Jaimovich and Rebelo (2009), but allow for habits in consumption:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{\left(C_t - \phi_c C_{t-1} - \gamma \frac{L_t^{1+\sigma_l}}{1+\sigma_l} S_t \right)^{1-\sigma_c} - 1}{1 - \sigma_c} \right\}, \tag{3.1}$$

where $\phi_c \in [0,1]$ measures the degree of internal habit persistence, $\sigma_c \geq 0$ governs the intertemporal elasticity of substitution, $\sigma_l \geq 0$ is related to the Frisch elasticity of labor supply, and $\gamma \geq 0$ measures the relative disutility of labor effort. The term

$$S_t = (C_t - \phi_c C_{t-1})^{\sigma_G} S_{t-1}^{1-\sigma_G} \tag{3.2}$$

makes the preferences non-separable in both consumption and work effort, where $\sigma_G \in [0,1]$ parameterizes the strength of the wealth effect on the labor supply. When $\sigma_G = 1$, the preference specification is equal to the one discussed in King et al. (1988), while with $\sigma_G = 0$ the preference specification of Greenwood et al. (1988) with no wealth effect on the labor supply is obtained.

The household faces the budget constraint

$$\begin{aligned} C_t + z_t^I I_t + \frac{B_{t+1}}{P_t} =& (1-\tau_t^n) \int_0^1 W_t(j)\, l_t(j)\, dj + \left(1-\tau_t^k\right) r_t^k u_t K_t \\ &+ \left(1-\tau_t^k\right)(R_{t-1}-1)\frac{B_t}{P_t} + \frac{B_t}{P_t} + \Phi_t + T_t + \left(1-\tau_t^k\right)\Xi_t\ , \end{aligned} \tag{3.3}$$

where the household earns income from supplying differentiated labor services $l_t(j)$ at the real wage $W_t(j)$ to union j, and from renting out capital services $u_t K_t$ at the rental rate r_t^k. In addition, it receives lump sum transfers T_t from the government and profits Ξ_t from owning the firms in the economy. All forms of income are taxed at their respective tax rates τ_t^n and τ_t^k. The term $\left(1-\tau_t^k\right)(R_{t-1}-1)\frac{B_t}{P_t} + \frac{B_t}{P_t}$ is the after-tax return on savings in bonds, where the net returns are taxed at the capital tax rate. Bonds are in zero net supply. The household spends its income on consumption C_t and investment $z_t^I I_t$, where I_t is gross investment and z_t^I denotes a shock to the relative price of investment in terms of the consumption good. This price is equal to the technical rate of transformation between investment and consumption goods. Due to the presence of a temporary shock, it is exogenous and stochastic. Changes in z_t^I do not affect the productivity of already installed capital, but do affect newly installed capital and become embodied in it. We assume the shock to follow an

AR(2)-process[7]

$$\log z_t^I \;=\; \rho_1^{zI} \log z_{t-1}^I + \rho_2^{zI} \log z_{t-2}^I + e^{\sigma_t^{zI}} \nu_t^{zI} \;, \tag{3.4}$$

where σ_t^{zI} allows for time-varying volatility and is discussed in detail in Section 3.4. Apart from the fact that this form of investment-specific technology may be an important source of economic fluctuations (Greenwood et al., 1997, 2000), a stochastic relative price of investment introduces costly reversibility and expandability of investment into the model as the future purchase/resale price is stochastic.

The term Φ_t captures depreciation allowances, which are an important feature of the U.S. tax code. We assume depreciation allowances of the form

$$\Phi_t = \tau_t^k \sum_{s=1}^{\infty} \delta_\tau \left(1 - \delta_\tau\right)^{s-1} z_{t-s}^I I_{t-s} \;, \tag{3.5}$$

where δ_τ is the depreciation rate for tax purposes.[8] By providing new investment with a tax shield, depreciation allowances may be important in capturing the dynamics of investment following shocks (Christiano et al., 2007; Yang, 2005). Through this tax shield at historical investment prices, combined with a stochastic relative price of investment z^I, depreciation allowances contribute to costly reversibility and expandability of investment.

The household owns the capital stock K_t, whose law of motion is given by

$$K_{t+1} = \left[1 - \left(\delta_0 + \delta_1 \left(u_t - 1\right) + \frac{\delta_2}{2}\left(u_t - 1\right)^2\right)\right] K_t + I_t - \frac{\kappa}{2}\left(\frac{I_t}{K_t} - \delta_0\right)^2 K_t \;, \tag{3.6}$$

where I_t is gross investment. Household members do not simply rent out capital, but capital services $u_t K_t$, where u_t denotes the capital utilization, i.e. the intensity with which the existing capital stock is used. Without loss of generality, capital utilization in steady state is normalized to 1. Using capital with an intensity higher than normal incurs costs to the household in the form of a higher depreciation

[7]The lag lengths for the individual exogenous driving processes is chosen to provide a good empirical fit. Details are provided in Section 3.4.

[8]Following Auerbach (1989), we allow the depreciation rate for tax purposes to differ from the physical rate.

$\delta(u_t) = \delta_0 + \delta_1(u_t - 1) + \delta_2/2(u_t - 1)^2$, which, assuming $\delta_0, \delta_1, \delta_2 > 0$, is an increasing and convex function of the capital utilization. The last term in equation (3.6) captures capital adjustment costs at the household level of the form introduced by Hayashi (1982), where $\kappa \geq 0$ is a parameter governing the curvature of the cost function. This functional form implies that the capital adjustment costs are minimized and equal to 0 in steady state. We choose this type of adjustment costs for three reasons. First, while this functional form clearly is unable to explain some micro-level phenomena like lumpy investment, it has nevertheless been shown to provide a good fit of firm level investment data and performs better than the Christiano et al. (2005)-formulation with quadratic adjustment costs in investment changes (Eberly et al., 2008). Second, with the flow specification of Christiano et al. (2005), Tobin's marginal q would be independent of the capital stock, which would essentially shut off intertemporal linkages and thereby the option effects (Wu, 2009).

Thus, the household maximizes its utility (3.1) by choosing $C_t, B_{t+1}, u_t, K_{t+1}, I_t, S_t, L_t$, subject to the constraints (3.2) - (3.6) and the resource constraint for aggregate labor.

3.3.2 Labor Market

The household supplies labor $l_t(j)$ equally to a continuum of unions j, $j \in [0,1]$. This labor market structure allows to introduce differentiated labor services and staggered wage setting without letting idiosyncratic wage risk affect the household members, which would make aggregation intractable. Monopolistically competitive unions supply differentiated labor $l_t(j)$ to intermediate firms at wage $W_t(j)$. Every period, each union may re-optimize its wage with probability $(1-\theta_w), 0 < \theta_w < 1$. If a union j cannot re-optimize, its nominal wage is indexed to the price level according to $W_t(j) P_t = \Pi_{t-1}^{\chi_w} W_{t-1}(j) P_{t-1}$, where $\chi_w \in [0,1]$ measures the degree of indexing. Hence, when the union has not been able to re-optimize for τ periods, its real wage τ periods ahead is given by:

$$W_{t+\tau}(j) = \begin{cases} W_{t+\tau}^{opt}(j), & \text{if able to re-optimize in } t+\tau, \\ \prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} W_t(j), & \text{otherwise.} \end{cases} \tag{3.7}$$

Household members supply the amount of labor services that is demanded at the current wage. The objective of each union able to reset its wage is to choose the real wage that maximizes the expected utility of its members, given the demand for its labor services $l_t(j) = (W_t(j)/W_t)^{-\eta_w} L_t^{comp}$, where L_t^{comp} is the aggregate demand for composite labor services and η_w is the substitution elasticity, the respective resource constraint $L_t = L_t^{comp} \int_0^1 (W_t(j)/W_t)^{-\eta_w} dj$, and the aggregate wage level $W_t = \left(\int_0^1 W_t(j)^{1-\eta_w} dj\right)^{\frac{1}{1-\eta_w}}$.

3.3.3 Firm Side

There is a continuum of monopolistically competitive intermediate goods firms i, $i \in [0,1]$, which produce differentiated intermediate goods Y_{it} using capital services $K_{it}^{serv} = u_{it} K_{it-1}$ and a composite labor bundle L_{it}^{comp} according to a Cobb-Douglas production function with capital share α

$$Y_{it} = z_t \left(K_{it}^{serv}\right)^{\alpha} \left(L_{it}^{comp}\right)^{1-\alpha} - \phi \ , \quad \text{if} \ \ z_t \left(K_{it}^{serv}\right)^{\alpha} \left(L_{it}^{comp}\right)^{1-\alpha} - \phi > 0 \tag{3.8}$$

and $Y_{it} = 0$ otherwise. The fixed cost of production ϕ is set to reduce economic profits to 0 in steady state, thereby ruling out entry or exit (Christiano et al., 2005). The stationary TFP shock z_t follows an $AR(2)$-process

$$\log z_t = \rho_1^z \log z_{t-1} + \rho_2^z \log z_{t-2} + e^{\sigma_t^z} \nu_t^z \ . \tag{3.9}$$

The composite labor bundle is built from differentiated labor inputs $L_{it}(j)$ according to a Dixit-Stiglitz aggregator $L_{it}^{comp} = \left(\int_0^1 L_{it}(j)^{\frac{\eta_w - 1}{\eta_w}} dj\right)^{\frac{\eta_w}{\eta_w - 1}}$.

We assume staggered price setting a la Calvo (1983) and Yun (1996). Each period, intermediate firms can re-optimize their prices with probability $(1-\theta_p)$, $0 < \theta_p < 1$. In between two periods of re-optimization, the prices are indexed to the aggregate price index P_t according to $P_{it+1} = \left(\frac{P_t}{P_{t-1}}\right)^{\chi_p} P_{it} = (\Pi_t)^{\chi_p} P_{it}$, where $\chi_p \in [0,1]$ governs the degree of indexation. Intermediate goods producers maximize their discounted stream of profits subject to the demand from composite goods producers, equation (3.11).

There is a competitive final goods firm which bundles a final good Y_t from a continuum of intermediate goods using a Dixit-Stiglitz aggregation technology with substitution elasticity η_p

$$Y_t = \left(\int_0^1 Y_{it}^{\frac{\eta_p - 1}{\eta_p}} \, di \right)^{\frac{\eta_p}{\eta_p - 1}} . \tag{3.10}$$

Expenditure minimization yields the optimal demand for intermediate good i as

$$Y_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\eta_p} Y_t \quad \forall \, i \, . \tag{3.11}$$

3.3.4 Government Sector

Government spending, which may be thought of as entering the utility function additively separable, follows the process

$$\log \left(\frac{G_t}{\bar{G}} \right) = \rho_1^g \log \left(\frac{G_{t-1}}{\bar{G}} \right) + \rho_2^g \log \left(\frac{G_{t-2}}{\bar{G}} \right) + e^{\sigma_t^g} \nu_t^g \ , \tag{3.12}$$

where $\bar{G}$ is government spending in steady state. The government finances its expenditures by distortionary taxation of labor at the rate τ_t^n and capital and interest income at rate τ_t^k. We assume $AR(2)$-processes for the tax rates as this has been found to be a good empirical description for the U.S. (McGrattan, 1994; Mertens and Ravn, 2011)

$$\tau_t^k = (1 - \rho_1^{\tau k} - \rho_2^{\tau k}) \bar{\tau}^k + \rho_1^{\tau k} \tau_{t-1}^k + \rho_2^{\tau k} \tau_{t-2}^k + e^{\sigma_t^{\tau k}} \nu_t^{\tau k} \tag{3.13}$$

$$\tau_t^n = (1 - \rho_1^{\tau n} - \rho_2^{\tau n}) \bar{\tau}^n + \rho_1^{\tau n} \tau_{t-1}^n + \rho_2^{\tau n} \tau_{t-2}^n + e^{\sigma_t^{\tau n}} \nu_t^{\tau n} \ , \tag{3.14}$$

where $\bar{\tau}^n$ and $\bar{\tau}^k$ are the unconditional means of the labor and capital tax rates, respectively. The government also sets lump-sum transfers T_t to balance the budget. This assumed structure yields the government budget constraint

$$T_t + G_t + \Phi_t = \tau_t^n W_t L_t^{comp} + \tau_t^k \left(r_t^k u_t K_t + \Xi_t \right) \ . \tag{3.15}$$

Transfers plus government spending plus depreciation allowances equal tax revenues from taxing labor, capital income, and profits.

We close the model by assuming that the central bank follows a Taylor rule that reacts to inflation and output growth.

$$\frac{R_t}{\bar{R}} = \left(\frac{R_{t-1}}{\bar{R}}\right)^{\rho_R} \left(\left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_{t-1}}\right)^{\phi_y}\right)^{1-\rho_R} \exp\left(m_t\right). \tag{3.16}$$

Here, ρ_R is a smoothing parameter introduced to capture the empirical evidence of gradual movements in interest rates (Clarida et al., 2000; Rudebusch, 1995), $\bar{\Pi}$ is the target interest rate set by the central bank, and the parameters ϕ_y and ϕ_Π capture the responsiveness of the nominal interest rate to deviations of inflation and output growth from their steady state values. We assume that the central bank responds to changes in output rather than its level as this specification conforms better with empirical evidence and avoids the need to define a measure of trend growth that the central bank can observe (see Lubik and Schorfheide, 2007). Finally, m_t is a shock to the nominal interest rate that follows an $AR(1)$-process

$$\log m_t = \rho^m \log m_{t-1} + e^{\sigma_t^m} \nu_t^m \ . \tag{3.17}$$

The definition of equilibrium and the market aggregation are standard and omitted for brevity.

3.4 Policy Risk: Time Series Evidence

In this section, we present empirical evidence on the importance of time-varying volatility in modeling macroeconomic time series. We demonstrate that the data tend to reject the homoskedasticity of macroeconomic driving processes and show that a stochastic volatility (SV) model is able to capture the salient features of the data. Using a particle smoother, we are able to recover the historical series of uncertainty shocks and show that both “good luck” and “good policy” contributed to the Great Moderation.

3.4.1 Estimation Methodology

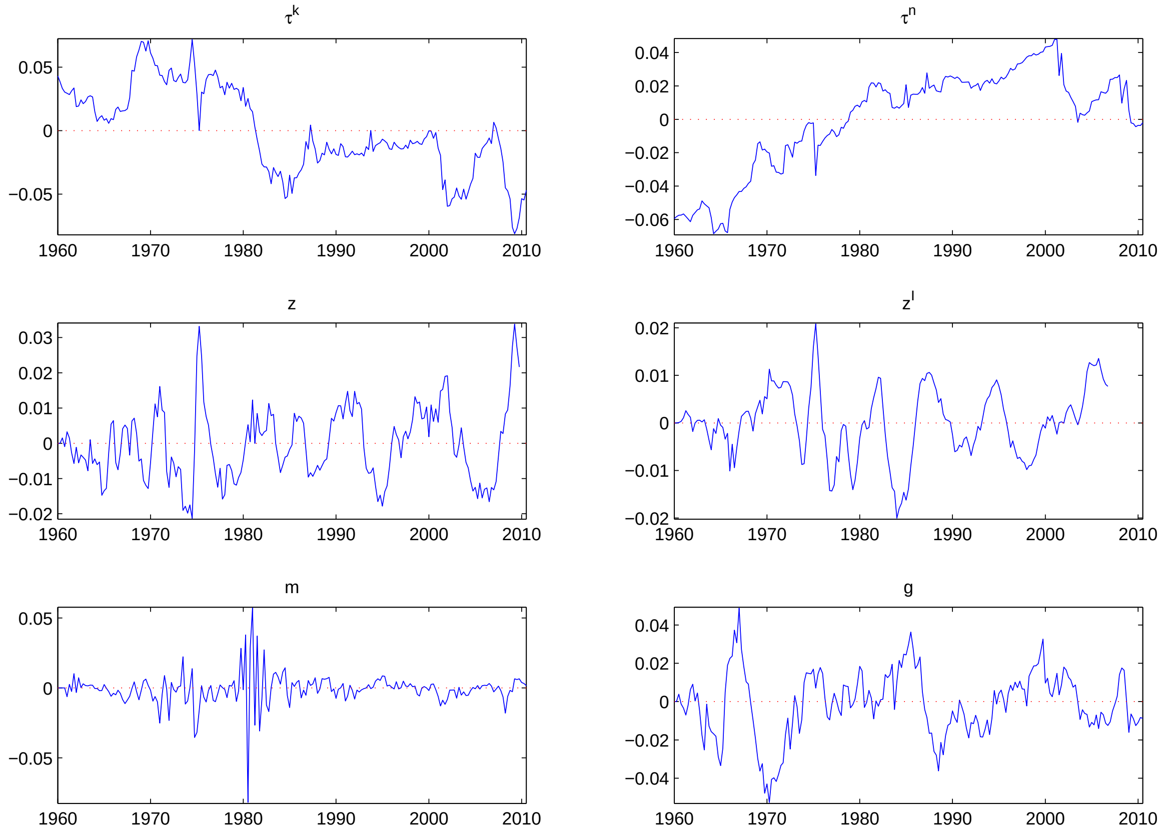
We perform a two-step estimation procedure. Due to the non-linear solution of the model required to capture uncertainty effects and the high-dimensional state space, it is computationally infeasible to jointly estimate all model parameters. Hence, we first estimate the exogenous stochastic driving processes of the model using Sequential Monte Carlo (SMC) methods. In the next section we feed these processes into the model presented in Section 3.3 and estimate the parameters of the remaining model equations with a Simulated Method of Moments (SMM) approach.

The model includes 6 exogenous stochastic driving processes with time-varying volatility, i.e. capital and labor tax rates, government spending, a monetary policy shock, total factor productivity, and investment-specific technology. We estimate these processes on quarterly U.S. time series, starting in 1960Q1 and using the longest available sample for each series. Details about the data sources can be found in Appendix A. Because we use a stationary model, we need to extract the deviations of the non-stationary time series from their respective trend. Hence, we apply a one-sided HP-filter to the logarithms of government spending and the two technology processes. Using a one-sided, i.e. "causal" filter (Stock and Watson, 1999) assures that the time ordering of the data remains undisturbed and the autoregressive structure is preserved. We allow for AR(2)-processes in all variables, except for the monetary policy shocks,[9] as the partial autocorrelations generally indicate the presence of a second root different from zero. Figure 3.1 shows the time series of the exogenous driving processes on which we estimate our laws of motion. In particular for monetary policy, the presence of time-varying volatility is immediately evident. In Appendix C, we provide further evidence for the presence of time-varying volatility.

There are two major competing approaches to model time-varying standard deviations: GARCH models and stochastic volatility (SV) models (Fernández-Villaverde and Rubio-Ramírez, 2010). In the standard GARCH model, σ_t^2 is a function of the squared scaled lagged innovation in the level equation ν_{t-1}^2 and its own lagged value: $\sigma_t^2 = \omega + \alpha(\sigma_{t-1}\nu_{t-1})^2 + \beta\sigma_{t-1}^2$. The GARCH model has one important drawback: there are no distinct volatility shocks. The only innovations to the volatility equation

[9] Although theory suggests that monetary policy shocks in the Taylor rule should be unpredictable and thus i.i.d., we find a moderate degree of first-order autocorrelation.

Figure 3.1: Time series of exogenous driving processes



Notes: From left to right and top to bottom: capital taxes, labor taxes, TFP, investment-specific technology, monetary policy shocks, and government spending. Tax rates are demeaned; government spending and technology processes are detrended using one-sided HP-filter.

are past level shocks, meaning that they cannot be separated from volatility shocks. As we are especially interested in the effects of shocks to the volatility, we cannot use a GARCH model but instead employ a stochastic volatility model. Specifically, we model the standard deviations σ_t^i as an AR(1) stochastic volatility process (see e.g. Fernández-Villaverde et al., forthcoming; Shephard, 2008)

$$\sigma_t^i = (1 - \rho^{\sigma_i})\,\bar{\sigma}^i + \rho^{\sigma_i}\sigma_{t-1}^i + \eta_i \varepsilon_t^i \;, \quad \varepsilon_t^i \sim \mathcal{N}(0,1), \tag{3.18}$$

where $\bar{\sigma}^i$ is is the unconditional mean of σ_t^i, $i \in \{\tau k, \tau n, g, m, z, zI\}$. The shock to the volatility ε_t^i is assumed to be independent from the level shock ν_t^i.

Due to the nonlinearity embedded in the stochastic volatility setup of the shocks, we cannot simply employ the Kalman filter as in the case of linearity and normally distributed shocks. For this case, Fernández-Villaverde and Rubio-Ramírez (2007) propose to use the *Sequential Importance Resampling* (SIR) particle filter, a special application of the more general class of *SMC* methods, to evaluate the likelihood.[10]

After obtaining the likelihood of the observables given the parameters, we use a *Tailored Randomized Block Metropolis-Hastings (TaRB-MH)* algorithm (Chib and Ramamurthy, 2010) to maximize the posterior likelihood. The prior distributions of the parameters, which are relatively weak, are given in Table 3.2.[11]

We are also interested in backing out the historical values of the latent state σ_t, given the whole set of observations. After filtering, it is straightforward to employ the *backward-smoothing routine* (Godsill et al., 2004) to obtain a historical distribution of the volatilities. The smoothed values were computed at the mean of the posterior distribution using 10,000 particles.

3.4.2 Estimation Results

The estimation results are presented in Table 3.2. Detailed convergence diagnostics are shown in Appendix C. In general, all parameters are quite precisely estimated as evidenced by the percentiles. All shocks, except for the monetary policy shock, exhibit a high degree of persistence in their levels, with less persistence in their volatilities. Moreover, the estimated processes show considerable evidence of uncertainty, with η^i ranging between 0.3 and 0.6. As a one-standard deviation uncertainty shock increases the volatility of the respective process by $(exp(\eta^i) - 1) \times 100$ percent, such a shock increases the variance of capital taxes, labor taxes, TFP, investment specific technology, monetary policy, and government spending by 46%, 92%, 38%, 39%, 34%,

[10]Technical details of the algorithms used in this subsection can be found in Appendices B-B.

[11]For the autoregressive parameters of the level equation ρ_1^i and ρ_2^i, we impose a uniform prior for each of the corresponding autoregressive roots over the stability region $(-1, +1)$. Let ξ_1 and ξ_2 be the roots of such an $AR(2)$-process. The autoregressive parameters corresponding to these roots can be recovered from: $\rho_1 = \xi_1 + \xi_2$ and $\rho_2 = -\xi_1\xi_2$. The posterior distribution was computed from a 20,500 draw Monte Carlo Markov Chain using 3,000 particles, where the first 2,500 draws were discarded as burn-in draws. Acceptance rates were generally between 20% and 45%. We also checked identifiability of the SV-process by simulating data from the process and trying to recover the true parameters from this artificial data.

Table 3.2: Prior and posterior distributions of the shock processes

Parameter	Prior distribution			Posterior distribution		
	Distribution	Mean	Std. Dev.	Mean	5 Percent	95 Percent
			Capital Tax Rates			
ρ_1	Uniform*	0.00	0.577	0.856	0.819	0.893
ρ_2	Uniform*	0.00	0.577	0.103	0.070	0.137
ρ_σ	Beta*	0.90	0.100	0.795	0.745	0.860
η_σ	Gamma	0.50	0.100	0.379	0.333	0.426
$\bar{\sigma}$	Uniform	-7.00	5.333	-5.071	-5.361	-4.786
			Labor Tax Rates			
ρ_1	Uniform*	0.00	0.577	1.051	1.018	1.084
ρ_2	Uniform*	0.00	0.577	-0.052	-0.085	-0.019
ρ_σ	Beta*	0.90	0.100	0.581	0.514	0.670
η_σ	Gamma	0.50	0.100	0.651	0.587	0.718
$\bar{\sigma}$	Uniform	-7.00	5.333	-5.901	-6.253	-5.531
			Total Factor Productivity			
ρ_1	Uniform*	0.00	0.577	1.021	0.965	1.080
ρ_2	Uniform*	0.00	0.577	-0.175	-0.230	-0.125
ρ_σ	Beta*	0.90	0.100	0.679	0.611	0.781
η_σ	Gamma	0.50	0.100	0.320	0.272	0.369
$\bar{\sigma}$	Uniform	-7.00	5.333	-5.349	-5.555	-5.138
			Investment-Specific Technology			
ρ_1	Uniform*	0.00	0.577	1.420	1.369	1.468
ρ_2	Uniform*	0.00	0.577	-0.501	-0.536	-0.461
ρ_σ	Beta*	0.90	0.100	0.807	0.765	0.861
η_σ	Gamma	0.50	0.100	0.332	0.295	0.368
$\bar{\sigma}$	Uniform	-7.00	5.333	-6.206	-6.427	-5.983
			Government Spending			
ρ_1	Uniform*	0.00	0.577	0.919	0.866	0.972
ρ_2	Uniform*	0.00	0.577	-0.028	-0.079	0.018
ρ_σ	Beta*	0.90	0.100	0.719	0.623	0.865
η_σ	Gamma	0.50	0.100	0.295	0.227	0.368
$\bar{\sigma}$	Uniform	-7.00	5.333	-4.887	-5.193	-4.585
			Monetary Policy Shock			
ρ_1	Uniform*	0.00	0.577	0.427	0.385	0.469
ρ_σ	Uniform*	0.90	0.100	0.921	0.895	0.947
η_σ	Beta*	0.50	0.100	0.364	0.330	0.400
$\bar{\sigma}$	Gamma	-7.00	5.333	-5.188	-5.512	-4.849

Notes: Beta* indicates that the parameter divided by 0.999 follows a beta distribution. Uniform* indicates that the roots of the autoregressive process are estimated instead of the autoregressive coefficients and follow the specified prior distribution.

and 45%, respectively.[12] Appendix C shows the results of model misspecification tests applied to the SV model. In general, the model fits the data well and cannot be rejected.

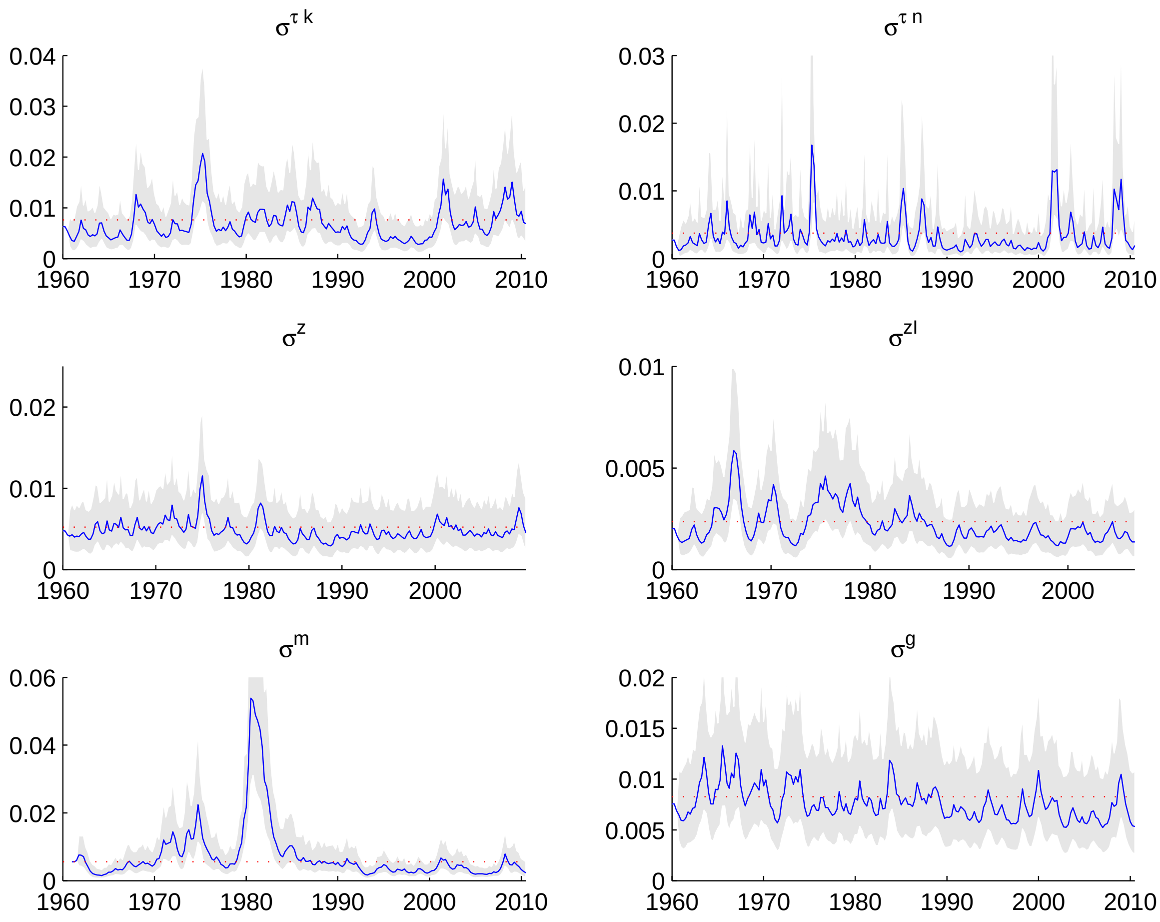
The relevance of stochastic volatility in modeling the behavior of the exogenous driving processes can be seen in the smoothed estimates of the historical variances of the shocks in Figure 3.2. The end of the 1960s and particularly the 1970s were plagued by high shock volatilities, both in the technology and the policy shocks. Particularly during the 1970s, the volatilities increased and reached their sample maxima for both tax rates and technology shocks. In contrast, the decade from 1985 to 2000 was characterized by shock volatilities to the technology variables well below their unconditional mean, indicating the role of "good luck" in explaining the Great Moderation. However, from about 1990 on "good policy" also contributed to this phenomenon as is evidenced by the low volatilities of the tax and government spending shocks, although the change in volatility is not as pronounced for the latter. For monetary policy shocks, there is clear evidence of a lower shock volatility following the Volcker disinflation from 1979-1983, a trend that also continued under Greenspan. In contrast, the early tenure of Volcker experienced a volatility of monetary shocks considerably larger than during the first oil price shock. With the height of the dot-com bubble the volatility of TFP shocks somewhat increased again, while the investment-specific technology growth remained tranquil over the whole 2000s. The largest changes in volatility in the 2000s came under George W. Bush who considerably changed the tax law, resulting in a pronounced increase in the volatility of tax rates. At the end of our sample, the Great Recession again results in an increase in policy risk with a rise in the volatility of government spending, tax rates, and monetary policy to comparable levels as after 9/11. For government spending and taxes, this mostly reflects the provisions in the *American Recovery and Reinvestment Act* that contained $288 billion in tax relief to companies and individuals, e.g. in the form of $116 billion in payroll tax relief.

Note that the SV-framework used in the present study does not imply a mechanical link between the level shocks and the volatility shocks as a GARCH-model would do. Of course, as a comparison of Figures 3.1 and 3.2 shows, a large level shock tends to

[12] Thus, e.g. a one-standard deviation monetary policy risk shock increases the volatility of the monetary policy shocks from $exp(-5.19) = 0.56\%$ to $exp(-5.19 + 0.364) = 0.8\%$.

coincide with an increase in the conditional variance. However, the reason for this increase in the estimated conditional variance is not a mechanical effect of this level shock subsequently entering the volatility equation. Rather, the Bayesian estimation of the SV-model weighs the likelihood of observing such a large shock being drawn from a narrow distribution, i.e. without observing a simultaneous/previous volatility shock, against the likelihood of observing a shock of this size that is drawn from a wider distribution due to the occurrence of a variance shock.

Figure 3.2: Smoothed standard deviations



Notes: From left to right and top to bottom: capital taxes, labor taxes, TFP, investment-specific technology, monetary policy shocks, and government spending. Red dotted line: unconditional mean; shaded area: two standard deviation bands.

3.5 Fitting the Model to the Data

Using the parameter estimates of the stochastic driving processes obtained in the previous section, we are now in a position to estimate the deep parameters of the model presented in Section 3.3.

3.5.1 Simulated Method of Moments Estimation

We use the Simulated Method of Moments (SMM) approach as proposed in Ruge-Murcia (2010). Intuitively, this method minimizes the weighted distance between the empirical moments and the moments resulting from artificial data simulated from the model (details can be found in Appendix B).

In order to simulate data, we first need to solve the model non-linearly. Due to the high-dimensional state space of our model, we employ perturbation methods to obtain an approximation of the policy function around the deterministic steady state (see e.g. Judd, 1998). Specifically, we need to obtain a third-order approximation, because we are interested in the pure effects of volatility shocks, i.e. when holding the level shocks constant. Loosely speaking, a first-order approximation yields no effects of uncertainty; a second-order approximation yields both a constant effect and an effect mediated through the corresponding level shock. Only in the third-order approximation does time-varying uncertainty play a separate role (for a more detailed explanation, see Appendix B).

Table 3.3 presents the values of parameters we fix prior to the estimation. We set gross steady state inflation $\bar{\Pi}$ to 1 and the discount factor β to 0.99. Regarding the depreciation parameters, $\delta_0 = 0.05$ is chosen to imply a 10% annual depreciation rate, $\delta_1 = 0.0351$ sets the steady state capital utilization to 1, and the depreciation rate for tax purposes δ_τ is set to twice the rate of physical depreciation (Auerbach, 1989). The fixed-cost parameter $\phi = 0.038$ implies that firms make zero profit in steady state and the labor disutility parameter $\gamma = 19.1$ sets the steady state share of hours worked to total time to 20%. Regarding the preference parameters, we set the parameter governing the intertemporal elasticity of substitution σ_c to 2 and set $\sigma_G = 0.001$, the value chosen in Jaimovich and Rebelo (2009).[13] Hence, preferences

[13]When attempting to estimate this parameter, it hit the lower bound of 0 as in Schmitt-Grohé and

Table 3.3: Parameters fixed prior to estimation

Param.	Value	Target/Motivation	Param.	Value	Target/Motiv.
$\bar{\Pi}$	1	Zero infl. steady state	σ_c	2	Standard value
β	0.99	Standard value	η_p	10	11% Markup
δ_0	0.025	10% annual deprec.	η_w	10	11% Markup
δ_1	0.0351	$\bar{u} = 1$	α	0.295	Sample mean
δ_τ	0.05	Auerbach (1989)	τ^n	0.1984	Sample mean
ϕ	0.038	0 profits in SS	τ^k	0.388	Sample mean
γ	19.1	SS labor of 0.2	G/Y	0.2031	Sample mean
σ_G	0.001	Jaimovich-Rebelo (2009)			

are close to the GHH-specification and imply a small wealth effect on the labor supply, which is consistent with evidence from studies focusing on the effects of news (Schmitt-Grohé and Uribe, 2010) and government spending (Monacelli and Perotti, 2008). The elasticity of substitution parameters for differentiated labor services and intermediate goods are set to 10, resulting in a steady state markup of 11%. The capital share α, the steady state tax rates τ^k and τ^n, and the steady state share of government spending to output are set to their respective sample means.

The empirical moments to be matched are the standard deviations and first- and second-order autocovariances of output, consumption, investment, inflation, the real wage, and the nominal interest rate. Moreover, we target the covariance of output with the other variables. All variables are logged and detrended using a one-sided HP-filter with smoothing parameter $\lambda = 1600$. The second and fourth columns of Table 3.5 display the respective sample moments.[14]

3.5.2 Parameter Estimates

The parameter estimates are shown in Table 3.4. All parameters except for the capital adjustment cost parameter κ are precisely estimated as seen in columns 4

Uribe (2010). Hence, we fix the parameter to a small value that still assures a balanced growth path.

[14]Some of the target moments are transformed to correlations for better interpretation. The relative standard deviations with respect to the standard deviation of output are only implicitly targeted through the standard deviations of the respective series.

Table 3.4: Parameters estimated by SMM

Parameter	Description	Mean	-1 std.-dev.	+1 std.-dev.
ϕ_c	Consumption habits	0.9665	0.9660	0.9671
δ_2/δ_1	Capital utilization costs	0.0414	0.0314	0.0546
κ	Capital adjustment costs	10.0857	0.8007	127.0438
θ_p	Calvo parameter prices	0.9644	0.9641	0.9646
θ_w	Calvo parameter wages	0.7785	0.7615	0.7947
χ_p	Price indexation	0.4170	0.3809	0.4539
χ_w	Wage indexation	0.9751	0.9725	0.9774
σ_l	Frisch elasticity parameter	0.0683	0.0652	0.0716
ρ_R	Interest smoothing	0.4889	0.4541	0.5238
ϕ_π	Taylor rule inflation	1.9691	1.9058	2.0422
ϕ_y	Taylor rule output growth	1.2195	0.8416	1.7671

and 5.[15] Consumers have strong habits in consumption with $\phi_c = 0.97$, which is at the upper end of values generally considered plausible. Capital utilization costs show little convexity with $\delta_2/\delta_1 = 0.04$, while capital adjustment is costly as indicated by $\kappa = 10.09$, ensuring that investment is not excessively volatile. Prices are estimated to be quite sticky with $\theta_p = 0.96$, while the degree of wage stickiness is moderate with an average duration of 4.3 quarters. The high degree of price stickiness compared to e.g. Smets and Wouters (2007) reflects the absence of real rigidities like a non-constant elasticity of substitution in our setup. The degree of indexation to past inflation is considerably higher for wages than for prices, with the former being almost perfectly indexed to past inflation. An estimated value of $\sigma_l = 0.07$ indicates almost linear disutility of labor. In the Taylor rule, there is a moderate degree of interest smoothing. The reaction coefficients of monetary policy are in line with values found in the literature.

The first and third column of Table 3.5 show the fit of the model. Output is 92% as volatile in the simulated as model as in the data, while investment is 108% as volatile. The volatility of consumption is well-matched, while its correlation with output is

[15] The confidence bands rely on the asymptotic normality of the estimator as shown in equation (3.36). However, this is only a rough approximation as most parameters, e.g. the Calvo parameters, have bounded support. Unfortunately, SMM is computationally too intensive to rely on bootstrapping the standard errors.

Table 3.5: Simulated and empirical moments

	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data
	$\sigma(x_t)$		$\rho(x_t, y_t)$		$\sigma_{x_t}/\sigma_{y_t}$		$\rho(x_t, x_{t-1})$		$\rho(x_t, x_{t-2})$	
Y	1.44%	1.57%	1.00	1.00	1.00	1.00	0.93	0.90	0.84	0.75
C	0.93%	0.95%	0.71	0.85	0.65	0.60	0.99	0.90	0.95	0.74
I	5.74%	5.30%	0.91	0.85	3.98	3.37	0.88	0.93	0.74	0.80
Π	0.22%	0.27%	0.23	0.17	0.16	0.17	0.91	0.50	0.75	0.32
W	0.82%	0.90%	0.23	0.10	0.57	0.57	0.97	0.84	0.91	0.69
R	0.40%	0.39%	0.28	0.34	0.28	0.25	0.73	0.86	0.49	0.67

Notes: Time Series X_t are output (Y_t), consumption (C_t), investment (I_t), inflation (Π_t), the real wage (W_t), and the nominal interest rate (R_t). Small letters denote variables that are logged and detrended using a one-sided HP-filter with smoothing parameter $\lambda = 1600$.

too low. The volatilities of the real wage, inflation, and the nominal interest rate are on target. Their correlation with output is also well matched. Only the real wage is somewhat too procyclical. The autocorrelations are also in general well-matched. Only consumption exhibits a slightly too high autocorrelation.

3.5.3 The Effects of Time-Varying Volatility

With the estimated model at hand, we can perform a simple counterfactual experiment to demonstrate the importance of time-varying volatility for explaining U.S. macroeconomic time series. However, the effects of time-varying volatility reflect both the ex-ante uncertainty effect of knowing that the shocks are drawn from a wider distribution and the ex-post effect of more extreme shock realizations. In the next section, we will therefore separate these two by using the model to keep the level shocks constant.

In Figure 3.2, we found clear evidence of a decrease in the variance of both the technological shocks and the policy shocks since the mid 1980s, which contributed to the lower volatility of output and inflation during the Great Moderation. Using our estimated DSGE-model, we can ask what a counterfactual economy without time-varying volatility would have looked like. For this purpose, we completely shut off time-varying volatility by setting uncertainty shocks to zero. We then simulate the model again using the new set of driving forces where both the uncertainty effect and the effects of the corresponding more extreme level shocks are absent due to $\sigma_t^i = \bar{\sigma}^i$ for all $i \in \{\tau k, \tau n, g, m, z, zI\}$. This unconditional sample mean of the log-volatility of the level shocks $\bar{\sigma}^i$ lies between the high volatility pre-Great Moderation period's value and the value in the subsequent low volatility Great Moderation phase. The corresponding simulated moments are presented in Table 3.6. The co-movement of the model variables still fits the data quite well. However, compared to the actual data, such an economy fails to generate sufficient volatility: output, consumption, and investment are only about 65%, 73%, and 75% as volatile as the data, respectively.[16] In contrast, as seen in Table 3.5, the model with time-varying volatility captures the data moments well. These results clearly indicate the importance of time-varying volatility in explaining U.S. macroeconomic time series (see e.g. Justiniano and Primiceri, 2008; Primiceri, 2005).

Table 3.6: Simulated and empirical moments for the model without time-varying volatility

	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data
	$\sigma(x_t)$		$\rho(x_t, y_t)$		$\sigma_{x_t}/\sigma_{y_t}$		$\rho(x_t, x_{t-1})$		$\rho(x_t, x_{t-2})$	
Y	0.99%	1.57%	1.00	1.00	1.00	1.00	0.94	0.90	0.85	0.75
C	0.71%	0.95%	0.67	0.85	0.72	0.60	0.99	0.90	0.95	0.74
I	3.91%	5.30%	0.89	0.85	3.97	3.37	0.92	0.93	0.79	0.80
Π	0.18%	0.27%	-0.19	0.17	0.18	0.17	0.91	0.50	0.76	0.32
W	0.53%	0.90%	0.56	0.10	0.54	0.57	0.97	0.84	0.91	0.69
R	0.30%	0.39%	-0.11	0.34	0.30	0.25	0.78	0.86	0.61	0.67

Notes: Time Series X_t are output (Y_t), consumption (C_t), investment (I_t), inflation (Π_t), the real wage (W_t), and the nominal interest rate (R_t). Small letters denote variables that are logged and detrended using a one-sided HP-filter with smoothing parameter $\lambda = 1600$.

[16] If we had used a linearized version of the model, this effect would not have been observed, as periods of high volatility would offset periods of low volatility. However, due to the non-linearity of our model, this is not the case here.

3.6 The Aggregate Effects of Policy Risk

We now turn to analyzing the effects of aggregate uncertainty on business cycle fluctuations. First, having estimated the deep parameters of the model, we conduct policy experiments to trace out the effects of uncertainty shocks. We then study their transmission into the economy and analyze the underlying amplification mechanisms. We find that the model is in principle able to generate large effects of uncertainty, but that the estimated parameterization implies that the aggregate effects of uncertainty are quantitatively small. The reason for the small aggregate response to uncertainty shocks is the presence of general equilibrium effects that imply only a weak amplification.

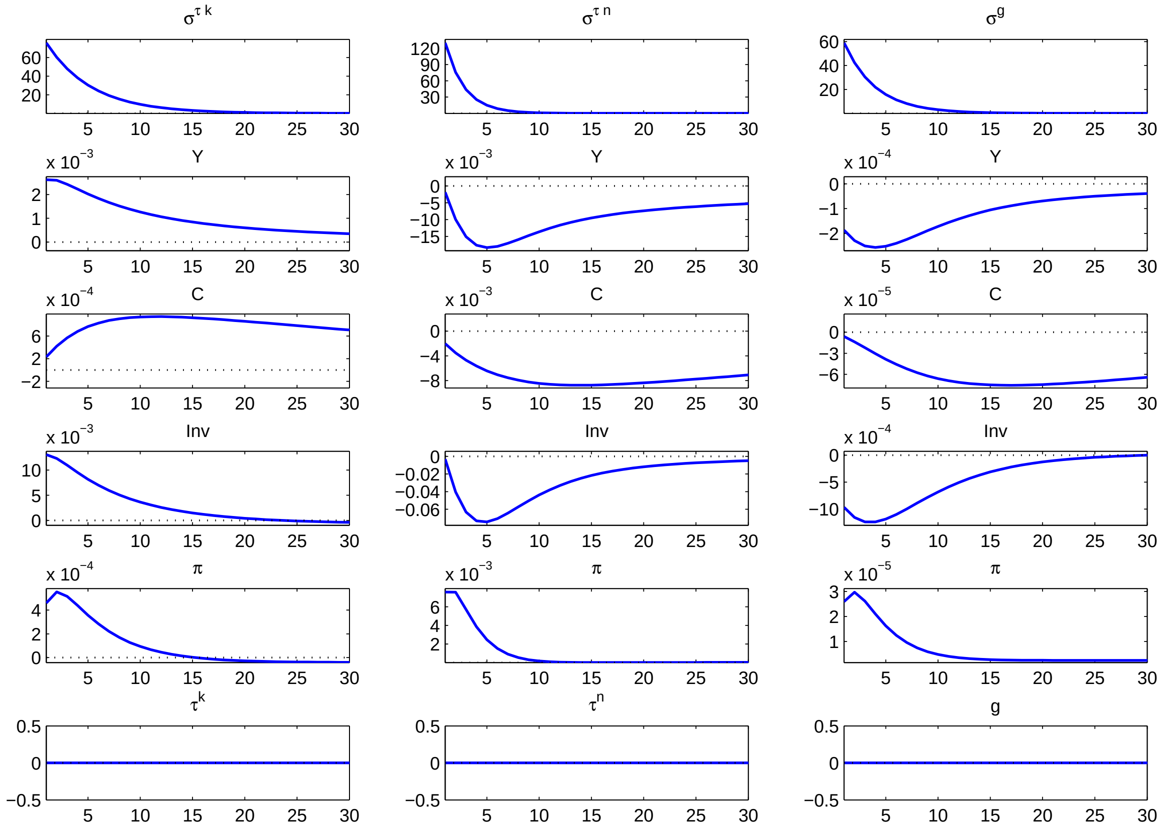
3.6.1 Impulse Response Analysis

We first analyze the pure uncertainty effect resulting from time-varying volatility by separating it from the ex-post effect of more extreme shock realizations. We do so by computing impulse response functions to uncertainty shocks while keeping constant the realizations of the level shocks.

Figures 3.3 and 3.4 show the impulse response functions to two-standard deviation policy risk and technology risk shocks with each column representing the impulse responses to a different shock. The ex-post level effect has been shut off, which is reflected in the flat impulse response for τ^k, τ^n, g, m, z, and z^I depicted in the bottom row.[17] The left column of Figure 3.3 shows that a capital tax risk shock acts like a positive demand shock. Output and inflation both increase on impact and slowly return to zero. Initially the output response is mostly driven by the positive response of investment, which has a peak response on impact of 0.014%. Consumption increases less strongly and follows a hump-shape, peaking after 12 quarters. Due to the estimated strong degree of habit persistence in consumption, the consumption response decays only slowly and drives the output response after about four years, when investment is already almost back to its initial level. The middle and right columns show the impulse responses to labor tax risk and government spending risk, respectively. Both emulate the characteristics of a negative supply shock, with output,

[17] In the subsequent graphs, we generally omit the flat level impulse responses.

Figure 3.3: Impulse responses to a two-standard deviation uncertainty shock to capital taxes, labor taxes, and government spending (from left to right column)



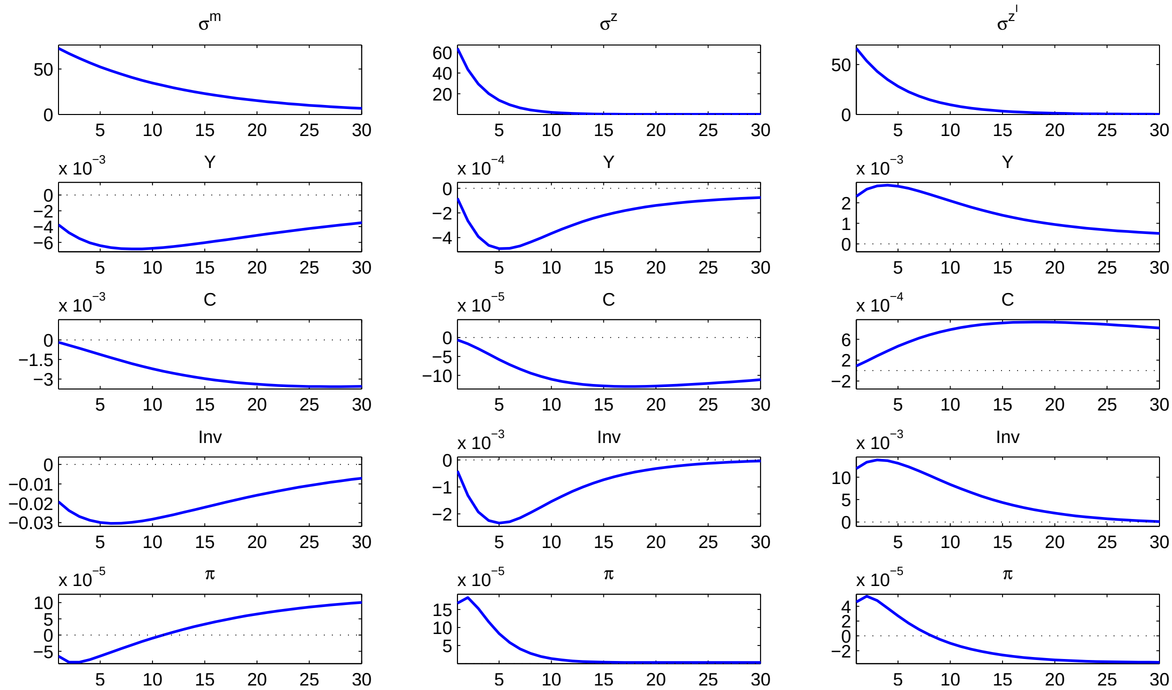
Notes: Level shocks are held constant. All responses are in percent, except for π which is in percentage points.

consumption, and investment exhibiting a hump-shaped decline, while inflation rises.

Labor tax risk induces the strongest output response of all uncertainty shocks considered, with output showing a peak decline of 0.02% and investment dropping by four times as much. The reason for this relatively strong response, compared to e.g. the government spending risk shock, is that a two-standard deviation labor tax risk shock increases uncertainty about labor taxes by about 120%, compared to around 60% for the other uncertainty shocks. Due to the relatively low persistence of the underlying shock process for labor tax risk, the effect on inflation subsides after 10 quarters, while the effect on consumption is again considerable more drawn out.

The left column of Figure 3.4 displays the response to a two-standard deviation

Figure 3.4: Impulse responses to a two-standard deviation uncertainty shock to monetary policy, TFP, and investment-specific technology (from left to right column)



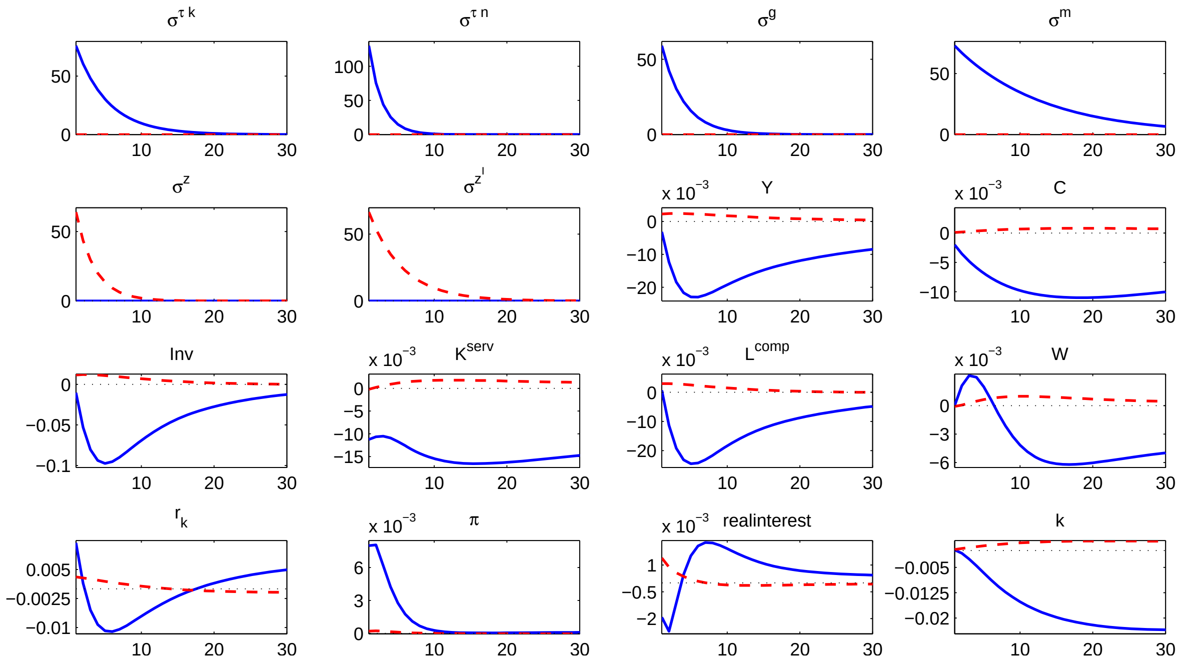
Notes: Level shocks are held constant. All responses are in percent, except for π which is in percentage points.

monetary policy risk shock. This shock has a contractionary effect on output, mostly driven by a decline in investment that peaks at -0.03% after 7 quarters. In contrast, consumption reacts sluggishly, peaking only after 30 quarters. Inflation initially drops, overshoots after 10 quarters and then slowly returns, driven by a large persistence in the underlying risk shock process.

The historical volatility estimates shown in Figure 3.2 indicated that uncertainty about the future path of economic policy increased for all policy instruments during the Great Recession. We simulate such a situation in the form of a simultaneous two-standard deviation increase in policy risk.[18] Results are shown in Figure 3.5. A

[18]Due to the nonlinearity inherent in our model and the solution method that preserves this nonlinearity up to third order, the resulting impulse responses are not necessarily identical to the sum of the impulse responses to the individual uncertainty shocks.

Figure 3.5: Impulse responses to a joint two-standard deviation policy risk shock (solid blue line) and to a joint technology risk shock (dashed red line)



Notes: Level shocks are held constant. All responses are in percent, except for π and *realinterest* which are in percentage points.

simultaneous two-standard deviation policy risk shock (solid lines) acts like a negative supply shock. It leads to an immediate decrease in output of 0.025%, before output slowly returns to its initial level as the shock subsides. This decrease in output is driven by both consumption and investment, with investment dropping initially by 0.1%. While the capital stock reacts sluggishly due to the presence of relatively high capital adjustment costs, capital services decline immediately due to an accompanying decline in capital utilization. At the same time inflation rises. As a consequence, the real wage rises for a few periods, reflecting the indexation to the rising inflation, and then starts to decrease, reaching its minimum after 15 quarters. Due to monopolistic competition in the labor market and the non-separability of the utility function, the initial increase in the real wage does not induce an increase in labor supplied by the household. Rather, household members decrease their labor supply and consume more leisure. The real interest rate, computed as the difference between the policy rate and inflation, declines initially and then follows a hump-shaped pattern, reaching its peak after 7 quarters. The initial decline in the real interest rate reflects both the interest smoothing present in the estimated Taylor rule as well as the response of the central bank to the initial decline in output. Only when output starts to recover does the real interest rate rise to bring down inflation. The similarity in both the size and the shape of the impulse response functions of a policy risk shock and the labor tax risk shock indicates that the latter dominates the effects of the other policy risk shocks.[19]

It is instructive to compare the policy risk results to the benchmark of uncertainty about technology. The middle and right columns of Figure 3.4 show the impulse responses to a two-standard deviation risk shock to total factor productivity and investment-specific technology, respectively. The response to TFP risk is qualitatively similar to what could have been expected from the previous literature: it triggers an investment driven decline in output while inflation increases. In contrast, investment-specific technology risk triggers exactly the opposite effect: output increases initially and peaks after 4 quarters, with the response again being mainly driven by the investment response. It is noteworthy that the response to TFP uncertainty is an order of magnitude smaller than the effects of uncertainty about the investment-specific technology shocks. This result suggests that the role of investment-specific technology risk might be underappreciated in the uncertainty literature.[20] Figure 3.5 also shows the impulse responses to a joint technology risk shock of the type occurring in the middle of the 1970s. The comparison of technology risk (dashed lines) with policy risk (solid lines) shows that policy risk generates responses that are one order of magnitude larger.

[19] While strictly speaking the impulse responses to single shocks are not additive, the opposite signs of the output response for some sources of uncertainty have important consequences for periods of generally heightened uncertainty. The simultaneous increase in uncertainty from different sources does not necessarily translate into a large output response. In times like the Great Recession, where policy risk jointly increased, different sources of uncertainty may partially offset each other, resulting in a low overall effect. For example, Figure 3.3 documents that capital taxation risk acts expansionary and could more than offset the negative effect of government spending risk on output and investment.

[20] While the effects of level shocks to investment-specific technology have received considerable attention in recent years (Fisher, 2006; Justiniano et al., 2010; Schmitt-Grohé and Uribe, 2011), we are to our knowledge the first to study the effects of uncertainty about investment-specific technology.

Summarizing, our results show that the finding of relatively minor effects of uncertainty on aggregate activity for the case of TFP (Bachmann and Bayer, 2011; Bachmann et al., 2010; Bekaert et al., 2010; Chugh, 2011; Popescu and Smets, 2010) also holds true for policy risk and investment-specific technology risk.

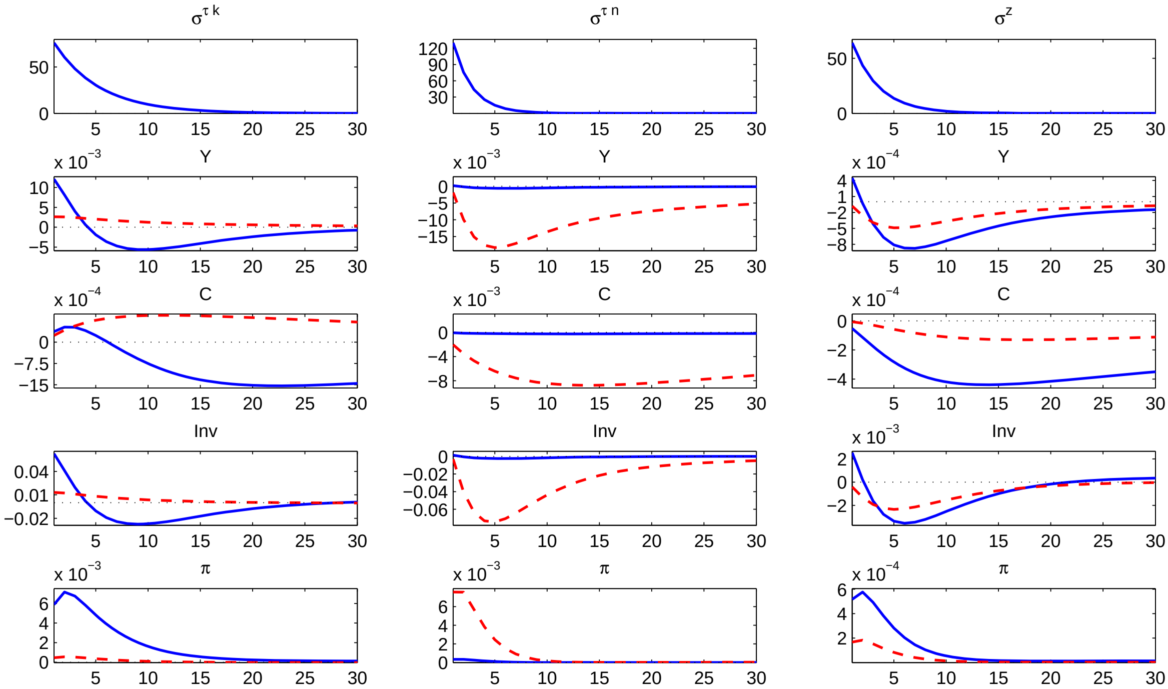
3.6.2 What Drives the Response to Policy Risk?

Of the transmission channels discussed in Section 3.2, the precautionary savings motive does not play a dominant role. In all sets of impulse responses, consumption and investment move in the same direction, while in the case of a dominant precautionary savings motive we would expect agents to decrease their consumption in order to self-insure against aggregate uncertainty by investing in a buffer-stock. Of course, it is conceivable that the precautionary savings motive counteracts the observed effects, which then would have been larger in its absence.

While it is virtually impossible to disentangle the different real option, Hartman-Abel, and general equilibrium effects, we can gain some insight into the transmission of uncertainty by shutting off various features of the model. First, as can be seen by fixing the relative price of investment to consumption at 1, the real option effect embedded in the depreciation allowances via the stochastic resale price of capital hardly plays a role. However, while their role in providing current investment with a tax shield at historical investment prices does not seem to create strong real option effects in our model, this does not mean that depreciation allowances do not play an important role. With their effect on Tobin's marginal q and the capital utilization decision, they have an important amplifying effect on the investment response and hence on output. When shutting them off completely, i.e. setting $\delta_\tau = 0$, capital drops less and the negative consumption response is cut in half (figures omitted for brevity).

Second, the low wealth effect on the labor supply implied by the preferences being close to the GHH-form ($\sigma_G \approx 0$) has a considerable effect on the responses to uncertainty, amplifying the response to some shocks and dampening the one to others. As shown in Figure 3.6, when setting the preferences to the standard King-Plosser-Rebelo specification ($\sigma_G = 1$), the negative response to labor tax risk declines by two orders of magnitude. At the same time, the effect of uncertainty shocks that

Figure 3.6: Impulse responses to a two-standard deviation uncertainty shock to capital taxes, labor taxes, and TFP (from left to right column)



Notes: solid blue line: KPR-preferences ($\sigma_G = 1$); red dashed line: preferences close to GHH ($\sigma_G \approx 0$). Level shocks are held constant. All responses are in percent, except for π which is in percentage points.

mainly affect the capital margin, i.e. capital tax and TFP risk, substantially increases, with the former now being the dominant policy risk factor. The output response to government spending, monetary policy and investment-specific technology risk stays largely unaltered (figures omitted for brevity).[21]

As noted in Section 3.2, the theoretical literature predicts an ambiguous effect of

[21]This finding of an important role of the preference specification for the transmission of uncertainty shocks suggests that adopting a certain form of utility function may already predetermine the sign of the output response to an uncertainty shock. Hence, future studies dealing with the effects of uncertainty should devote more attention to tracing out which preference specification may be the most suitable one. Our estimation results hint at a utility function featuring a low wealth effect on the labor supply. This is in line with an increasing number of studies from the fiscal policy (Monacelli and Perotti, 2008), open economy (Chang and Fernández, 2010; Garcia-Cicco et al., 2010), and news literature (Jaimovich and Rebelo, 2009; Schmitt-Grohé and Uribe, 2010), which also suggest the presence of a low wealth effect on the labor supply.

Table 3.7: Counterfactual calibration implying large uncertainty effects

Parameter	Description	Estimated mean	Counterfactual
ϕ_c	Consumption habits	0.96	0.9
κ	Capital adjustment costs	10.1	5
θ_p	Calvo parameter prices	0.96	0.9
σ_l	Frisch elasticity parameter	0.07	4
ρ_R	Interest smoothing	0.49	0.9
ϕ_y	Taylor rule output growth	1.22	0

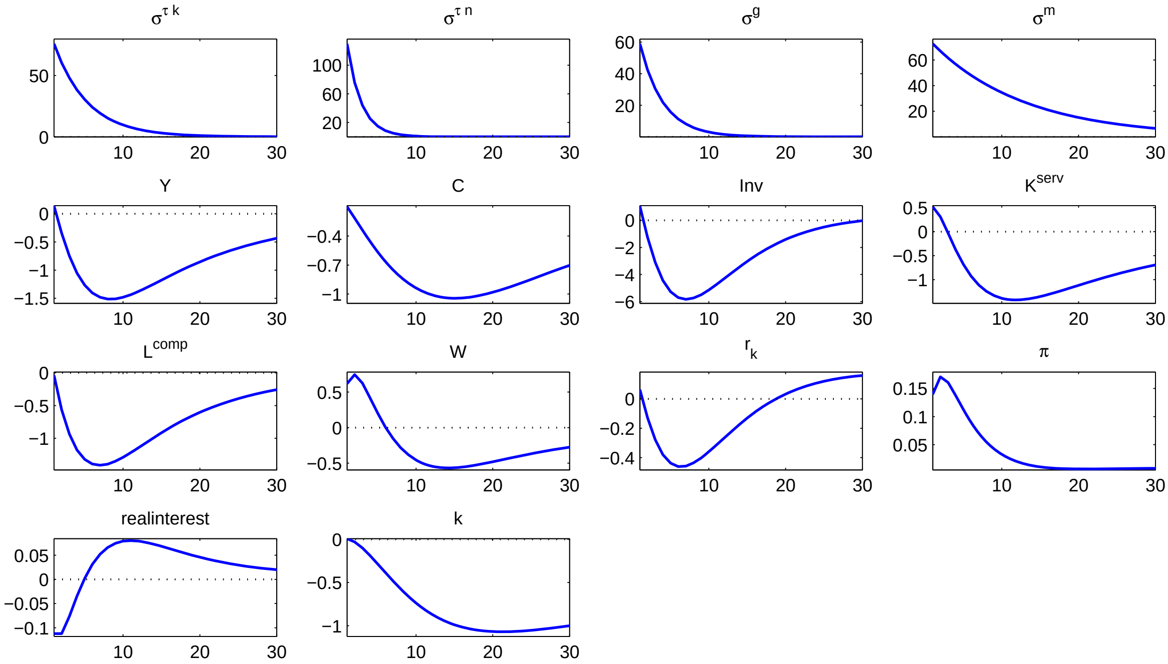
uncertainty as real option, Hartman-Abel, and general equilibrium effects drive the dynamics and may work in opposite directions. That this is actually the case for the specific types of uncertainty considered can be seen from, e.g., the impulse response of consumption to a capital tax shock depicted in the middle left panel of Figure 3.6. The consumption response is mostly negative for the case of $\sigma_G \approx 0$ but unambiguously positive for $\sigma_G = 1$. This suggests that different partial effects are dominating the respective responses for the different parameterizations. While a contractionary effect dominates in the GHH-case, an expansive effect prevails in the KPR-case. The strong dependence of uncertainty effects on the specific parameterization underscores the need for model estimation as opposed to calibration in order to trace out the aggregate effects of uncertainty.

3.6.3 Why Are the Effects of Uncertainty Small?

We identify strong general equilibrium effects – constraining the amplification of uncertainty shocks – as the main reason for the small effect of uncertainty on economic activity. While the model is in principle capable of generating large real effects of uncertainty, strong stabilizing effects are required to match the data moments. Therefore, SMM estimates the model parameters to imply strong equilibrating effects.

Consider the simple counterfactual experiment displayed in Table 3.7. Here, we decrease habit persistence, capital adjustment costs, price rigidities, and the Frisch elasticity of labor supply. To dampen the general equilibrium response of the nominal interest rate, we shut off the reaction to output growth and considerably increase the interest smoothing. In this case, as shown Figure 3.7, policy risk leads to a drop in output of 1.5%, which is mostly driven by a large decline in investment. While this

Figure 3.7: Impulse responses to a two-standard deviation policy risk shock under counterfactually volatile calibration



Notes: Level shocks are held constant. All responses are in percent, except for π and *realinterest* which are in percentage points.

calibration allows for larger effects of uncertainty, it comes at a cost: the model with this calibration implies unrealistically large business cycles. As shown in Table 3.8, output would be almost three times as volatile as found in the data, investment five times, and wages almost four times as volatile.

Hence, given the estimated exogenous driving processes, SMM estimates the parameters to imply a shock amplification more in line with the actually observed data. First, consumption habits, capital adjustment costs, and price rigidities are estimated to be quite high, generating a high persistence and thereby limiting the reaction of consumption, investment, and inflation to shocks and thus the deviations from the ergodic mean that are realized over time. Second, the parameter governing the Frisch elasticity of labor supply is estimated to be low so household's labor supply reacts quite flexibly to shocks. Third and most importantly, monetary policy reacts

Table 3.8: Simulated and empirical moments: counterfactual with stronger amplification

	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data
	$\sigma(x_t)$		$\rho(x_t, y_t)$		$\sigma_{x_t}/\sigma_{y_t}$		$\rho(x_t, x_{t-1})$		$\rho(x_t, x_{t-2})$	
Y	4.47%	1.57%	1.00	1.00	1.00	1.00	0.66	0.90	0.38	0.75
C	0.65%	0.95%	0.45	0.85	0.15	0.60	0.95	0.90	0.85	0.74
I	24.90%	5.30%	0.99	0.85	5.58	3.37	0.63	0.93	0.34	0.80
Π	0.29%	0.27%	0.76	0.17	0.07	0.17	0.80	0.50	0.53	0.32
W	3.55%	0.90%	0.85	0.10	0.79	0.57	0.83	0.84	0.57	0.69
R	0.31%	0.39%	-0.90	0.34	0.07	0.25	0.82	0.86	0.57	0.67

Notes: Time Series X_t are output (Y_t), consumption (C_t), investment (I_t), inflation (Π_t), the real wage (W_t), and the nominal interest rate (R_t). Small letters denote variables that are logged and detrended using a one-sided HP-filter with smoothing parameter $\lambda = 1600$.

fast and decisively to current economic conditions and in particular to output. The resulting transmission of both uncertainty and level shocks into the economy then implies less pronounced business cycles.

The decisive reaction to output growth is evident from the large coefficient estimate in the Taylor rule. The monetary authority's aggressive reaction to changes in output has a considerable dampening effect on the business cycle as it prevents output from deviating too far from steady state. When keeping all parameters at their baseline values but setting $\phi_y = 0$, thus shutting off the response of interest rates to output growth, triples the negative output response following a policy risk shock (figures omitted for brevity). The main reason for this behavior is the response of the real interest rate. The uncertainty shock acts like a negative supply shock, agents reduce their labor and capital input, and inflation rises. The monetary authority responds to this increase in inflation by raising the nominal interest rate without considering the negative impact on output. As a result, the real interest rate now has a positive impact response, amplifying the original shock's contractionary effect on output. In contrast, if the monetary authority also reacts to changes in output, the interest rate hike is more muted and the negative output response lower. The real interest initially declines to counteract the contractionary effect on output and only rises after several quarters.

The fast reaction of nominal interest rates to exogenous shocks can be seen from the relatively low degree of interest smoothing, meaning that current economic conditions affect nominal interests more than past interest rates. This low amount of interest smoothing exerts a considerable influence on the economy's response to uncertainty shocks, allowing a stronger counteracting reaction of the nominal interest rate, which dampens the uncertainty effects in a similar way as the output feedback of monetary policy. When giving more weight to past interest rates compared to the currently desired nominal interest, the nominal interest rate responds more sluggishly to shocks to the system, thereby temporarily allowing for larger deviations from steady state.

Hence, our result lend support to the findings of Bachmann and Bayer (2011). Their study showed for the case of idiosyncratic uncertainty about technology that general equilibrium effects, most importantly the endogenous feedback to wages and interest rates may considerably dampen the output effects of uncertainty shocks. Our results indicate that this also holds true for the case of aggregate uncertainty in an estimated DSGE-model.

These results suggest a potential issue for studies using a "proof-of-concept"-approach. Such studies typically show that uncertainty may matter by putting one source of uncertainty along one level shock into a model and then designing a transmission mechanism that enables this source to explain the whole business cycle. Our findings indicate that more attention needs to be devoted to what happens if other shocks, both uncertainty and level are present. As soon as other competing sources of aggregate fluctuations documented in the literature are added to these models, the effects of uncertainty are bound to decrease. Moreover, the approach of considering only one source of uncertainty and designing a particular amplification mechanism to generate an output drop in response may neglect that specially designed amplification mechanisms may interact with other types of shocks in undesired ways.[22]

[22] For example, expansionary output effects of uncertainty, which in our model e.g. arise with capital tax risk, might be amplified in the same way.

3.7 Conclusion

The current paper analyzes the effects of policy risk, i.e. aggregate uncertainty about labor and capital tax rates, monetary policy, and government spending on aggregate activity. We find that aggregate policy risk has only minor effects on the business cycle. Although its effects are an order of magnitude larger than the ones of technological uncertainty, a two standard-deviation policy risk shock still only generates a 0.025% drop in output. The reason for this small effect is that our parameter estimates imply strong general equilibrium effects that dampen the aggregate effects of uncertainty on economic activity. Most notably, the monetary authority's estimated strong and rapid response to current conditions implies a nominal interest rate reaction that considerably reduces aggregate fluctuations. While our model is capable of generating strong uncertainty effects, such a calibration would imply unrealistically large business cycle fluctuations. Thus, SMM estimates the amplification of uncertainty shocks to be rather low.

The small effect of uncertainty on output does not imply that time-varying volatility is unimportant. In accordance with the previous literature (e.g. Justiniano and Primiceri, 2008; Primiceri, 2005), our findings suggest that the Great Moderation can be explained through a combination of "good luck" and "good policy". The historical variance estimates indicate that the standard deviation of both technology and policy shocks significantly decreased since the mid-1980s. However, most of the effect of this time-varying volatility comes in the form of a different size of the realized level shocks instead of through the uncertainty-effect.

As our analysis focuses on aggregate uncertainty, it does not necessarily contradict studies finding large effects of idiosyncratic uncertainty. However, these studies clearly require different transmission mechanisms that do not give rise to large general equilibrium effects (see also Bachmann and Bayer, 2011).

Appendix to Chapter 3

A Data construction

Unless otherwise noted, all data are from the Bureau of Economic Analysis (BEA)'s NIPA Tables and available in quarterly frequency from 1960Q1 until 2010Q3.

Data for the exogenous processes

Capital and labor tax rates. Our approach to calculate average tax rates closely follows Mendoza et al. (1994), Jones (2002), and Leeper et al. (2010). We first compute the average personal income tax rate

$$\tau^p = \frac{IT}{W + PRI/2 + CI} ,$$

where IT is personal current tax revenues (Table 3.1 line 3), W is wage and salary accruals (Table 1.12 line 3), PRI is proprietor's income (Table 1.12 line 9), and $CI \equiv PRI/2 + RI + CP + NI$ is capital income. Here, RI is rental income (Table 1.12 line 12), CP is corporate profits (Table 1.12 line 13), and NI denotes the net interest income (Table 1.12 line 18).

The average labor and capital income tax rates can then be computed as

$$\tau^n = \frac{\tau^p(W + PRI/2) + CSI}{EC + PRI/2} ,$$

where CSI denotes contributions for government social insurance (Table 3.1 line 7), and EC is compensation of employees (Table 1.12 line 2), and

$$\tau^k = \frac{\tau^p CI + CT + PT}{CI + PT} ,$$

where CT is taxes on corporate income (Table 3.1 line 5), and PT is property taxes (Table 3.3 line 8).

Government spending. Government spending is the sum of government consumption (Table 3.1 line 16) and government investment (Table 3.1 line 35) divided by the GDP deflator (Table 1.1.4 line 1) and the civilian noninstitutional population (BLS, Series LNU00000000Q).

Monetary policy shock. Computed as the residual from a Taylor rule as in Clarida et al. (2000) (see Appendix B). The sample only starts in 1961Q1 as we lose the first year of data due to the use of four time lags as instruments in the GMM estimation.

Total factor productivity (TFP). The construction of TFP closely follows Beaudry and Lucke (2010), i.e.

$$TFP_t = \frac{Y_t}{K^{\alpha} H^{1-\alpha}} \ .$$

To construct K, we use data on capital services for the private non-farm business sector (Bureau of Labor Statistics (BLS), Historical Multifactor Productivity Tables),[23] multiply it by the total capacity utilization rate (Federal Reserve System, Statistical Release G.17 - Industrial Production and Capacity Utilization), and divide it by the civilian noninstitutional population above 16 years of age (BLS, Series LNU00000000Q). Real GDP per capita Y is nominal GDP (Table 1.1.5 line 1) divided by the GDP deflator (line 1 in Table 1.1.4) and the population, and per capita hours H are non-farm business hours worked (BLS, Series PRS85006033) divided by the population. The capital share α is set at 0.295, the mean over the sample compiled by the BLS (Bureau of Labor Statistics (BLS), Historical Multifactor Productivity Tables). The TFP-series ends in 2009Q4.

Relative price of investment. The relative price of investment is taken from Schmitt-Grohé and Uribe (2011) and only available until 2006Q4. They base their calculations on Fisher (2006).

The different sample lengths are not an issue as we estimate each exogenous process separately. Using the longest available sample assures that we make optimal use of the available information for each series.

[23] Quarterly data is interpolated from the annual series using cubic spline interpolation.

Data for SMM

Output. Nominal GDP (Table 1.1.5 line 1) divided by the GDP deflator (Table 1.1.4 line 1) and the civilian noninstitutional population (BLS, Series LNU00000000Q).

Investment. Sum of Residential fixed investment (Table 1.1.5 line 12) and nonresidential fixed investment (Table 1.1.5 line 9) divided by the GDP deflator (Table 1.1.4 line 1) and the civilian noninstitutional population (BLS, Series LNU00000000Q).

Consumption. Sum of personal consumption expenditures for nondurable goods (Table 1.1.5 line 5) and services (Table 1.1.5 line 6) divided by the GDP deflator (Table 1.1.4 line 1) and the civilian noninstitutional population (BLS, Series LNU00000000Q).

Real wage. Hourly compensation in the nonfarm business sector (BLS, Series PRS85006103) divided by the GDP deflator (Table 1.1.4 line 1).

Inflation. Computed as the log-difference of the GDP deflator (Table 1.1.4 line 1).

Nominal interest rate. Geometric mean of the effective Federal Funds Rate (St.Louis FED - FRED Database, Series FEDFUNDS).

Additional data for GMM

Interest term spread. We use the difference of the quarterly geometric mean of the 10-Year Treasury Constant Maturity Rate (FRED Database, Series GS10) and the quarterly geometric mean of the 3-Month Treasury Bill: Secondary Market Rate (FRED Database, Series TB3MS).

Money growth rate. Growth rate of the M2 Money Stock (FRED Database, Series M2SL).

Commodity inflation. Commodity inflation is computed as the growth rate of the X12-seasonally adjusted Producer Price Index: All Commodities (FRED Database, Series PPIACO).

Output gap. The output gap is constructed as the percentage difference between real GDP (FRED Database, Series GDPC96) and Real Potential Gross Domestic Product (FRED Database, Series GDPPOT).

B Econometric Methods

The Particle Filter

For ease of exposition, let x_t be a generic observable AR(1) process

$$x_t = \rho x_{t-1} + e^{\sigma_t}\nu_t \ , \quad \nu_t \sim \mathcal{N}(0,1) \tag{3.19}$$

where the unobserved/latent state σ_t follows a stochastic volatility process

$$\sigma_t = (1-\rho^\sigma)\,\bar{\sigma} + \rho^\sigma \sigma_{t-1} + \eta\varepsilon_t \ , \quad \varepsilon_t \sim \mathcal{N}(0,1), \tag{3.20}$$

where $\bar{\sigma}$ is is the unconditional mean of σ_t. The shock to the volatility ε_t is assumed to be independent from the level shock ν_t.

Hence, a filter is required to obtain the so-called filtering density $p\left(\sigma_t|\, x^t;\Theta\right)$. Due to the nonlinearity embedded in the stochastic volatility setup of the shocks, we cannot simply employ the Kalman filter as in the case of linearity and normally distributed shocks. Instead, we employ the *Sequential Importance Resampling* (*SIR*) particle filter, a special application of the more general class of *Sequential Monte Carlo* methods, to evaluate the likelihood (Fernández-Villaverde and Rubio-Ramírez, 2007; Fernández-Villaverde et al., forthcoming). Given the structure in (3.19) and (3.20) and some initial value x_0, the factorized likelihood of observing x^T can be written as

$$\begin{aligned}
p\left(x^T;\Theta\right) &= \prod_{t=1}^{T} p\left(x_t|x^{t-1};\Theta\right) \\
&= \int p\left(x_1|x_0,\sigma_0;\Theta\right) d\sigma_0 \ \prod_{t=2}^{T}\int p\left(x_t|x_{t-1},\sigma_t;\Theta\right) p\left(\sigma_t|x^{t-1};\Theta\right) d\sigma_t \\
&= \int \frac{1}{e^{\sigma_0}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_1-\rho x_0}{e^{\sigma_0}}\right)^2\right] d\sigma_0 \\
&\times \prod_{t=2}^{T}\int \frac{1}{e^{\sigma_t}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_t-\rho x_{t-1}}{e^{\sigma_t}}\right)^2\right] p\left(\sigma_t|x^{t-1};\Theta\right) d\sigma_t \ ,
\end{aligned} \tag{3.21}$$

where x^t is a $(t\times 1)$ vector that stacks the observations on x up to time t, Θ stacks the parameters, and the last equality follows from the assumption of normally distributed shocks. Although we do not have an analytical expression for $p\left(\sigma_t|x^{t-1};\Theta\right), t = 1,\ldots,T$, and can therefore not compute it directly, we can employ the particle filter to estimate the likelihood by iteratively drawing from $p\left(\sigma_t|x^{t-1};\Theta\right)$.

The underlying idea of the particle filter is to use an approximation of the filtering density $p\left(\sigma_t|\,x^t;\Theta\right)$ with a simulated distribution generated from empirical data. This distribution can be formed from mass points, or particles,

$$p\left(\sigma_t|\,x^t;\Theta\right) \simeq \sum_{i=0}^{N} \omega_t^i \delta_{\sigma_t^i}\left(\sigma_t\right), \quad \sum_{i=0}^{N} \omega_t^i = 1, \quad \omega_t^i \geqslant 0 \tag{3.22}$$

where δ is the Dirac delta function and ω_t^i is the weight attached to the respective draw/particle σ_t^i (Godsill et al., 2004). We can then use a *Sequential Importance Resampling* (SIR)-approach to update particles from time t to $t+1$ and obtain the new filtering distribution at $t+1$ (see e.g. Fernández-Villaverde et al., forthcoming). A convenient by-product of this filtering approach is that we also approximate $p\left(\sigma_t|x^{t-1};\Theta\right)$, the distribution we need to build the likelihood.

The SIR is a two-step procedure that, by using a prediction and a resampling/filtering step for each time period, ultimately allows to iteratively draw from $p\left(\sigma_t|x^{t-1};\Theta\right)$. Starting with $p\left(\sigma_0|\,x^0;\Theta\right) = p\left(\sigma_0;\Theta\right)$, the prediction step uses the law of motion for the states $f(\sigma_{t+1}|\sigma_t)$, equation (3.20), to obtain the conditional density $p\left(\sigma_1|\,x^0;\Theta\right) = p\left(\varepsilon_1\right) p\left(\sigma_0|\,x^0;\Theta\right)$. That is, given N draws $\left\{\sigma_{t|t}^i\right\}_{i=1}^N$ from $p\left(\sigma_t|\,x^t;\Theta\right)$, (here $p\left(\sigma_0|\,x^0;\Theta\right)$) and a draw of exogenous shocks $\varepsilon_t^i \sim \mathcal{N}(0,1)$, we can use equation (3.20) to compute $\left\{\sigma_{t+1|t}^i\right\}_{i=1}^N$.[24]

Next, the resampling/filtering step uses importance resampling to update the conditional probability from $p\left(\sigma_t|\,x^{t-1};\Theta\right)$ to $p\left(\sigma_t|\,x^t;\Theta\right)$. The crucial idea is that if $\left\{\sigma_{t|t-1}^i\right\}_{i=1}^N$ is a draw from $p\left(\sigma_t|x^{t-1};\Theta\right)$ and $\{\widetilde{\sigma}_t^i\}_{i=1}^N$ is a draw with replacement from $\left\{\sigma_{t|t-1}^i\right\}_{i=1}^N$ using the resampling probabilities

$$\omega_t^i = \frac{p\left(x_t|x^{t-1},\sigma_{t|t-1}^i;\Theta\right)}{\sum_{i=1}^N p\left(x_t|x^{t-1},\sigma_{t|t-1}^i;\Theta\right)}\ , \tag{3.23}$$

[24] The notation $t+1|t$ indicates a draw at time $t+1$ conditioned on the information available at time t.

then $\left\{\sigma_{t|t}\right\}_{i=1}^{N} = \{\widetilde{\sigma}_t^i\}_{i=1}^{N}$ is a draw from $p\left(\sigma_t|x^t;\Theta\right)$. The resampling with probabilities given in (3.23) serves two purposes. First, the reweighting implements an importance sampling approach, i.e. draws are obtained from a proposal density that is easy to draw from and are then subsequently reweighted to reflect the density to be approximated (see Arulampalam et al., 2002, for a derivation).[25] Second, without the resampling step, there would be an increase in the unconditional variance of ω_t over time, yielding only one particle with non-zero weight (known as degeneracy or sample impoverishment, see Arulampalam et al. (2002)). By resampling, we keep only those particles with high ω_t^i (i.e. those that are closer to the true state vector). Having now obtained draws from $p\left(\sigma_t|\, x^t;\Theta\right)$, we can again start with the prediction step to obtain draws for time period $t+1$.

After T iterations, we get an estimate of our likelihood as[26]

$$\begin{aligned} p\left(x^T;\Theta\right) \simeq\; & \frac{1}{N}\sum_{i=1}^{N}\frac{1}{e^{\sigma_{0|0}}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x_1-\rho x_0}{e^{\sigma_{0|0}}}\right)^2\right] \\ & \times\prod_{t=2}^{T}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{e^{\sigma_{t|t-1}}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x_t-\rho x_{t-1}}{e^{\sigma_{t|t-1}}}\right)^2\right] \;. \end{aligned} \tag{3.24}$$

Particle Smoother

We employ the *backward-smoothing routine* suggested by Godsill et al. (2004) to draw from the smoothing density $p(\sigma^T|x^T;\Theta)$ to get a historical distribution of the volatilities. Specifically, we start with the factorization

$$p(\sigma^T|x^T;\Theta) = p(\sigma_T|x^T;\Theta)\prod_{t=1}^{T-1}p(\sigma_t|\sigma_{t+1:T},x^T;\Theta) \;. \tag{3.25}$$

[25]In our case, we use the prior density $p\left(\sigma_t|\,\sigma^{t-1};\Theta\right)$ as the importance density.

[26]See Fernández-Villaverde and Rubio-Ramírez (2007) and Doucet and Johansen (2009) and the references contained therein for the conditions required for a central limit theorem to apply, yielding a consistent estimator of $p\left(x^T;\Theta\right)$.

The second factor can be further simplified

$$\begin{aligned} p(\sigma_t|\sigma_{t+1:T}, x^T; \Theta) &= p(\sigma_t|\sigma_{t+1}, x^t; \Theta) \\ &= \frac{p(\sigma_t|x^t; \Theta) f(\sigma_{t+1}|\sigma_t)}{p(\sigma_{t+1}|x^t)} \\ &\propto p(\sigma_t|x^t; \Theta) f(\sigma_{t+1}|\sigma_t) \ , \end{aligned} \tag{3.26}$$

where the first equality results from the Markovian properties of the model and f denotes the state transition density following from equation (3.20). Equation (3.22) describes how to approximate $p(\sigma_t|x^t; \Theta)$ by forward filtering. Therefore, we can approximate $p(\sigma_t|\sigma_{t+1:T}, x^T; \Theta) \propto p(\sigma_t|x^t; \Theta) f(\sigma_{t+1}|\sigma_t)$ by

$$p(\sigma_t|\sigma_{t+1}, x^T; \Theta) \simeq \sum_{i=1}^{N} \omega^i_{t|t+1} \delta_{\sigma^i_t}(\sigma_t) \ , \tag{3.27}$$

where the new weights $\omega^i_{t|t+1}$ are given by

$$\omega^i_{t|t+1} = \frac{\omega^i_t f(\sigma_{t+1}|\sigma^i_t)}{\sum_{j=1}^{N} \omega^j_t f(\sigma_{t+1}|\sigma^j_t)} \ . \tag{3.28}$$

and the ω^i_t are the weights obtained in the filtering step. Denote with $\tilde{\sigma}^i_t$ the i^{th} draw from the smoothing density at time t. At time T, we can obtain draws $\tilde{\sigma}^i_T$ by drawing from $p(\sigma_T|x^T)$ with the weights ω^i_T. Then, going backwards in time, we can use the above recursions to iteratively obtain draws $\tilde{\sigma}^i_t$ by resampling using the weights given in (3.28).

Tailored Randomized Block Metropolis Hastings Algorithm

Let Θ, $p\left(x^T|\Theta\right)$, and $\pi(\Theta)$ denote the vector of parameters to be estimated, the likelihood function, and the prior distribution of the parameters, respectively. The posterior distribution $\pi(\Theta|x^T)$ can be computed as

$$\pi\left(\Theta|x^T\right) \propto p\left(x^T|\Theta\right) \pi\left(\Theta\right) \ . \tag{3.29}$$

Given this usually analytically intractable posterior, most macroeconomic applications employ a Random Walk Metropolis-Hastings (RW-MH) algorithm to generate draws from the posterior distribution. However, the standard RW-MH algorithm often has poor mixing properties, leading to highly autocorrelated draws, and is therefore often very inefficient. Hence, to increase the efficiency, we use the Tailored Randomized Block Metropolis Hastings (TaRB-MH) algorithm proposed by Chib and Ramamurthy (2010).[27] Instead of in each iteration step simultaneously drawing an entire new parameter vector from a proposal density, the parameter vector is randomly split up into several blocks. Each block is then subsequently updated by a separate MH run, conditional on the previous step's values of the parameters in the other blocks. Ideally, the blocks should be formed according to the correlation between parameters, with highly correlated parameters belonging to the same block. However, we have no a priori knowledge about the correlation between parameters and resort to a blocking scheme where both the number of blocks and its composition are randomized in each step. This algorithm provides a good compromise between the standard RW-MH and tailored multiple block MH algorithms that use multiple blocks, which are particularly designed for the problem at hand. The second feature that improves on the standard RW-MH is that in each step the proposal density is "tailored" to the location and the curvature of the posterior density in that block by using a non-derivative based global optimizer. We deviate from Chib and Ramamurthy (2010) by using the CMAES algorithm (Hansen et al., 2003) instead of a simulated annealing as the former has been shown to be more efficient (Andreasen, 2010).[28] Moreover, it requires considerably less tuning than a simulated annealing. The TaRB-MH algorithm proceeds as follows.

1. At each iteration step n, $n = 1, \ldots, N$, the elements of the parameter vector θ are separated into random blocks $(\theta_{n,1}, \theta_{n,2}, \ldots, \theta_{n,p_n})$ by perturbing their initial ordering and assigning the first parameter in the perturbed vector to the first block and each following parameter with probability $p = 0.5$ to a new block, leaving us with 2.5 blocks on average as we estimate 5 parameters.

2. At each iteration step n, each block $\theta_{n,l}, l = 1, \ldots, p_n$ is sampled by a Metropolis-

[27] Using the TaRB-MH decreased the inefficiency factors from values around 10 to below 2.

[28] For an intuitive introduction to the working of the CMAES algorithm, see Binsbergen et al. (2010).

Hastings step using a proposal density adapted to the posterior in the following way. Denote with $\theta_{n,-l}$ the most current value of all blocks except for the lth one, i.e. their value at the end of step $n-1$. To generate a new draw for $\theta_{n,l}$, the CMAES-algorithm is used to find

$$\hat{\theta}_{n,l} = \underset{\theta_{n,l}}{\arg\max} \log \left[p\left(x^T | \theta_{n,l}, \theta_{n,-l}\right) \pi\left(\Theta\right) \right]. \tag{3.30}$$

That is, we use a global optimizer to maximize the posterior over the current block l, given the value of all other parameters at the end of step $n-1$. Having found the "conditional mode" $\hat{\theta}_{n,l}$, we compute the curvature of the target posterior distribution in the standard way as the negative inverse of the Hessian at the "conditional mode"

$$V_{n,l} = \left(- \frac{\partial \log \left[p\left(x^T | \theta_{n,l}, \theta_{n,-l}\right) \pi\left(\Theta\right) \right]}{\partial \theta_{n,l} \theta'_{n,l}} \right)^{-1} \Bigg|_{\theta_{n,l} = \hat{\theta}_{n,l}}. \tag{3.31}$$

Following Chib and Ramamurthy (2010), we use a multivariate t-distribution with ν degrees of freedom as proposal density for $\theta_{n,l}$, $q_l\left(\theta_{n,l} | \, \theta_{n,-l}, x^T\right)$. Mean and variance are set to the "conditional mode" and the negative inverse of the Hessian at this point:

$$q_l\left(\theta_{n,l} | \, \theta_{n,-l}, x^T\right) = t\left(\theta_{n,l} | \, \hat{\theta}_{n,l}, V_{n,l}, \nu\right). \tag{3.32}$$

In the Metropolis-Hastings-step, a proposed value $\theta^*_{n,l}$ is accepted as the new value of the block with probability

$$\alpha_l\left(\theta_{n,l}, \theta^*_{n,l} \middle| \, \theta_{n,-l}, x^T\right) = \min \left[\frac{p\left(x^T | \theta^*_{n,l}, \theta_{n,-l}\right) \pi\left(\theta^*_{n,l}\right)}{p\left(x^T | \theta_{n,l}, \theta_{n,-l}\right) \pi\left(\theta_{n,l}\right)} \frac{t\left(\theta_{n,l} | \, \hat{\theta}_{n,l}, V_{n,l}, \nu\right)}{t\left(\theta^*_{n,l} \middle| \, \hat{\theta}_{n,l}, V_{n,l}, \nu\right)}, 1 \right]. \tag{3.33}$$

If the proposed value $\theta^*_{n,l}$ is rejected, we set $\theta_{n+1,l} = \theta_{n,l}$. This step is repeated for all p_n blocks before the algorithm starts over with step 1.

Setting $\nu = 5$ and iterating over steps 1 and 2, we can - after a suitable burn-in-period - obtain samples from the desired posterior distribution, which is the invariant distribution of the resulting Markov Chain. In our case, a burn-in of 2500 proved sufficient.

Model Solution

Let s_t denote the $n_s \times 1$ vector of state variables in deviations from steady state, including the exogenous shocks and the perturbation parameter Λ, and let s_t^i denote its ith entry. The policy function/law of motion for an arbitrary model variable $\widehat{X}_t$ then has the form

$$\widehat{X}_t = \sum_{i=1}^{n_s} \xi_i^X s_t^i + \frac{1}{2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \xi_{i,j}^X s_t^i s_t^j + \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \sum_{l=1}^{n_s} \xi_{i,j,l}^X s_t^i s_t^j s_t^l \ , \tag{3.34}$$

where the ξ's are scalars that depend on the deep parameters of the model and hats denote percentage deviations from steady state. Equation (3.34) shows why lower-order approximations would not be sufficient for our purpose.

As is well known, a first-order approximation exhibits certainty equivalence. This implies $\xi_v^X = 0$, where v denotes the position of a volatility shock in the state vector s. That is, up to first order, uncertainty shocks do not enter the policy function at all.

For a second-order approximation, it is well known from Schmitt-Grohé and Uribe (2004) for the homoskedastic case that uncertainty only enters the policy function through a constant term via the second derivative with respect to the perturbation parameter, i.e. through $\xi_{\Lambda,\Lambda} \neq 0$. However, things are more complicated in the heteroscedastic case where shocks to the variance occur, leading to an additional effect. Fernández-Villaverde et al. (2010) prove that in this case, the volatility shocks additionally only enter the policy function with non-zero coefficients in their interaction term with the respective level shock. Algebraically, only the cross-product of $\hat{\sigma}^i \times \hat{\nu}^i$ is different from 0. In contrast, all other cross-terms with the uncertainty shocks are zero, i.e $\xi_{v,j\neq u}^X = 0$, where v and u denote the positions of a volatility and its corresponding level shock in the state vector s, respectively. Hence, the effect of uncertainty is always mediated through level shocks. It is not possible to shock the variance of the level shocks independently from the level shock as its effect would be 0 by construction.

Only in the third-order approximation do the volatility shocks enter the policy function separately from the level shocks in a non-constant form. Most importantly, the term $\xi_{i,\Lambda,\Lambda}$ is in general different from 0 for all volatility shocks.

Simulated Method of Moments

The idea of the Simulated Method of Moments (SMM) is the following. Let x_t be a time t vector of observables from a stationary and ergodic distribution and let $\{x_t\}_{t=1}^{T}$ be the corresponding sequence. Furthermore, let $m(x_t)$ denote a $k \times 1$ vector of empirical moments computed from this data. Denote with $\{x_t^{sim}(\theta)\}_{t=1}^{aT}$ the corresponding time series of length aT generated from simulating the model using the $p \times 1$ parameter vector $\theta \in \Theta$, with $\Theta \subset R^p$. Let $m(x_t^{sim}(\theta))$ be the vector of simulated moments computed from the artificial data. The SMM estimator is the value of θ that satisfies

$$\widehat{\theta} = \underset{\theta \in \Theta}{\arg\min} \left[m(x_t) - m\left(x_t^{sim}(\theta)\right)\right]' W \left[m(x_t) - m\left(x_t^{sim}(\theta)\right)\right] \ , \tag{3.35}$$

where W is a $p \times p$ positive definite weighting matrix. Under the assumption that the model with $\theta = \theta_0$ is a correct representation of the true process that generated $m(x_t)$ and the regularity conditions spelled out in Duffie and Singleton (1993), $\hat{\theta}$ is a consistent estimator of θ_0 with asymptotic distribution

$$\sqrt{T}\left(\widehat{\theta} - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, (1 + 1/\tau)\left(J'WJ\right)^{-1} J'WSWJ\left(J'WJ\right)^{-1}\right), \tag{3.36}$$

where

$$S = \lim_{T \to \infty} Var\left((1/\sqrt{T})\sum_{t=1}^{T} m(x_t)\right), \tag{3.37}$$

and $J = E(\partial m(x_t^{sim})/\partial\theta)$ (see Ruge-Murcia, 2010).

This estimator is asymptotically efficient when using the weighting matrix

$$W = \left(V^{longrun}\right)^{-1} = \left[\lim_{T \to \infty} Var\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} m(x_t)\right)\right]^{-1} . \tag{3.38}$$

The ideal weighting matrix places the most weight on the linear combination of moments that are the most precisely measured in the data. However, for two reasons, we use only the diagonal of the optimal weighting matrix:

$$W^{diag} = \text{diag}\left(V^{longrun}\right)^{-1}. \tag{3.39}$$

First, we would like to put more weight on moments that are actually observed in the data and that are economically meaningful, rather than on a linear combination of moments (see also Cochrane, 2005). Second, in practice, fully specified weighting matrices often lead to diverging parameter estimates. As shown in Ruge-Murcia (2010), using only the main diagonal of the optimal weighting matrix leads to a loss in efficiency but nevertheless delivers good results in most cases.

The simulation proceeds as follows. Starting at the deterministic steady state, we simulate the model for 3015 quarters using shocks drawn from the estimated shock distributions. Shocks larger than two standard deviations are trimmed. To assure non-explosive behavior of the simulations, we use the pruning algorithm of Kim et al. (2008). We discard the first 2000 quarters as a burn-in in order to reach the ergodic distribution. We then use the remaining 1015 quarters to compute the respective moments. The results are robust to using a longer burn-in period. The choice of using five times the length of the original data sample (i.e. $a = 5$) to compute the moments is motivated by the simulations in Ruge-Murcia (2010), who finds this choice to deliver a good balance between the precision of the estimates and computation time.

Impulse Responses

The nonlinearity of our model complicates the computation of impulse responses compared to linear models. We follow Fernández-Villaverde et al. (forthcoming) and generate impulse responses as the response to a two standard deviation shock to uncertainty at the ergodic mean. First, we simulate the model for 2,000 quarters by drawing shocks from the respective estimated distributions. Shocks larger than two standard deviations are trimmed to assure convergence, which technically depends on the shocks being bounded. To assure non-explosive behavior of the simulations, we use the pruning algorithm of Kim et al. (2008). We discard the first 2,000 quarters as a burn-in in order to reach the ergodic distribution and use the next 675 quarters to compute the ergodic mean. Starting at the ergodic mean, we compute the IRFs as the percentage difference of the respective variables between the system shocked with the respective shock and the baseline model response, i.e. the model response without shocks. To account for sampling uncertainty, we generate 50 different IRFs with different starting values of the random number generator and take the cross-sectional average as our impulse response.

GMM

We construct the monetary policy shocks by specifying the Federal Reserve's policy reaction function and estimating it by the generalized method of moments (GMM). Our approach is similar to the one used in Clarida et al. (2000), with the difference that Clarida et al. (2000) use a forward-looking policy reaction function, while we use a rule that reacts to contemporaneous variables to stay consistent with our DSGE-model. Specifically, the policy reaction function to be estimated is given by

$$r_t = \rho r_{t-1} + (1-\rho)\left[\bar{r} + \phi_\pi \left(\pi_t - \bar{\pi}\right) + \phi_y y_t^{gap}\right] + \varepsilon_t \ , \tag{3.40}$$

where π_t is inflation with target rate $\bar{\pi}$, y_t^{gap} is the output gap, r_{t-1} allows for interest smoothing, $\bar{r}$ ist the target nominal interest rate, and ε_t is an error term. Using the vector of instruments $\mathbf{z}_t$, the set of moment conditions for our GMM estimation procedure can be written as

$$E\left[\left\{r_t - \rho r_{t-1} - \alpha - \beta\pi_t - \gamma y_t^{gap}\right\} z_t\right] = 0 \tag{3.41}$$

where $\alpha = (1-\rho)\left(\bar{r} + \phi_\pi\bar{\pi}\right)$ collects all constant terms, $\beta = (1-\rho)\phi_\pi$, and $\gamma = (1-\rho)\phi_y$.

Hence, we regress the average effective Federal Funds Rate in the first month of the quarter on the lagged FFR, the inflation rate, and the output gap, where all rates are annualized. The set of instruments includes four lags of the FFR, the inflation rate, the output gap, commodity price inflation, money growth, and the interest term spread. Because we are only interested in the residuals of the policy reaction function

Table 3.9: GMM estimation of Taylor rule

Coefficient	Mean	Std. Error	t-Statistic	Prob.
ρ	0.898	0.018	48.926	0.000
α	0.001	0.001	0.874	0.383
β	0.1741	0.027	6.361	0.000
γ	0.102	0.017	5.950	0.000
R-squared	0.890	Mean dependent var		0.058
Adjusted R-squared	0.888	Sum squared resid		0.027
S.E. of regression	0.012	J-statistic		18.545
Durbin-Watson stat	2.314	pval(J-statistic)		0.552

Note: Kernel: Bartlett, Bandwidth: Fixed (4), No prewhitening; Simultaneous weighting matrix & coefficient iteration; Convergence achieved after: 28 weight matrices, 29 total coef iterations.

$\widehat{\varepsilon}_t$, we do not need to separately identify the target nominal rate $\bar{r}$ and target inflation $\bar{\pi}$.

Table 3.9 presents the estimation results, which are all in the range typically reported in the literature. There is strong evidence of interest smoothing with $\rho = 0.898$. The point estimates of the feedback parameters are $\phi_\pi = 1.718$ and $\phi_y = 1.003$. The test of overidentifying restrictions shows that the model cannot be rejected at conventional significance levels.

C Diagnostics

Testing for Heteroskedasticity

Table 3.10 presents evidence of the need to model time-varying volatility. Despite our relatively short sample size and the low power of tests for heteroskedasticity, the null hypothesis of homoskedastic shocks can be rejected at the 10% level for all series except labor taxes. This result is consistent with evidence that the standard deviation of structural shocks has changed over time (see e.g. Justiniano and Primiceri, 2008; Primiceri, 2005).

Table 3.10: Tests for heteroskedasticity

	τ^k	τ^n	z	z_I	g	m
White	0.000*	0.932	0.001*	0.042*	0.360	0.068*
WW	0.169	0.523	0.265	0.005*	0.076*	0.068*
BPK	0.004*	0.890	0.126	0.770	0.511	0.298

Notes: Asterisks indicate significance at the 10% level. White refers to the standard White (1980)-test, WW refers to the Wooldridge (1990)-version of this test, and BPK refers to the Breusch and Pagan (1979)/Koenker (1981)-test.

Convergence Diagnostics

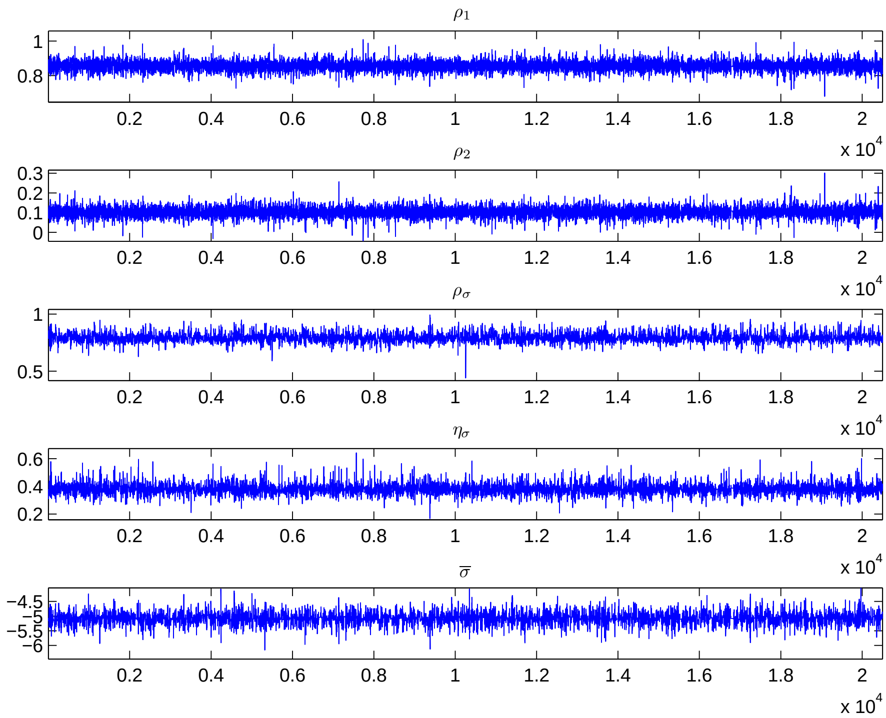
Table (3.11) shows the results from the Geweke (1992)-convergence diagnostics that compares the means of the first 20% of draws with that of the last 50% of the draws. In general, all MCMC chains have converged to their stationary distribution as indicated by the p-values of the χ^2-test for equal means. Figures 3.8 to 3.13 show the corresponding mean plots.

Table 3.11: Geweke (1992) convergence diagnostics

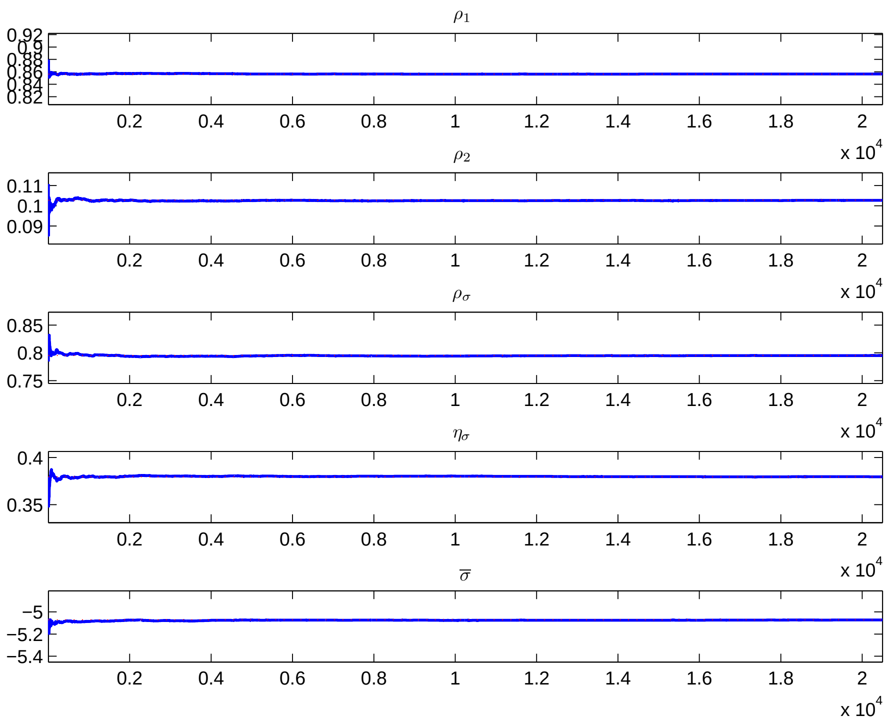
Parameter	4% taper	8% taper	15% taper	4% taper	8% taper	15% taper
	Capital Tax Rates			Labor Tax Rates		
ρ_1	0.160	0.165	0.145	0.909	0.890	0.887
ρ_2	0.947	0.941	0.937	0.926	0.913	0.904
ρ_σ	0.623	0.596	0.566	0.648	0.652	0.653
η_σ	0.929	0.927	0.919	0.327	0.319	0.271
$\bar{\sigma}$	0.760	0.744	0.738	0.922	0.921	0.917
	Total Factor Productivity			Investment Specific Technology		
ρ_1	0.891	0.887	0.879	0.199	0.174	0.124
ρ_2	0.679	0.681	0.665	0.353	0.340	0.297
ρ_σ	0.643	0.615	0.583	0.546	0.534	0.520
η_σ	0.456	0.453	0.391	0.638	0.649	0.638
$\bar{\sigma}$	0.772	0.765	0.706	0.304	0.260	0.187
	Government Spending			Monetary Policy Shock		
ρ_1	0.608	0.598	0.572	0.192	0.200	0.181
ρ_2	0.605	0.606	0.558			
ρ_σ	0.550	0.561	0.562	0.231	0.227	0.155
η_σ	0.293	0.267	0.232	0.885	0.870	0.860
$\bar{\sigma}$	0.412	0.402	0.369	0.066	0.078	0.071

Notes: Numbers are p-values of the χ^2-test for equal means of the first 20% of draws and the last 50% of the draws (after the first 2500 draws are discarded as burn-in).

Figure 3.8: Evolution of MCMC sampler over time for τ^k

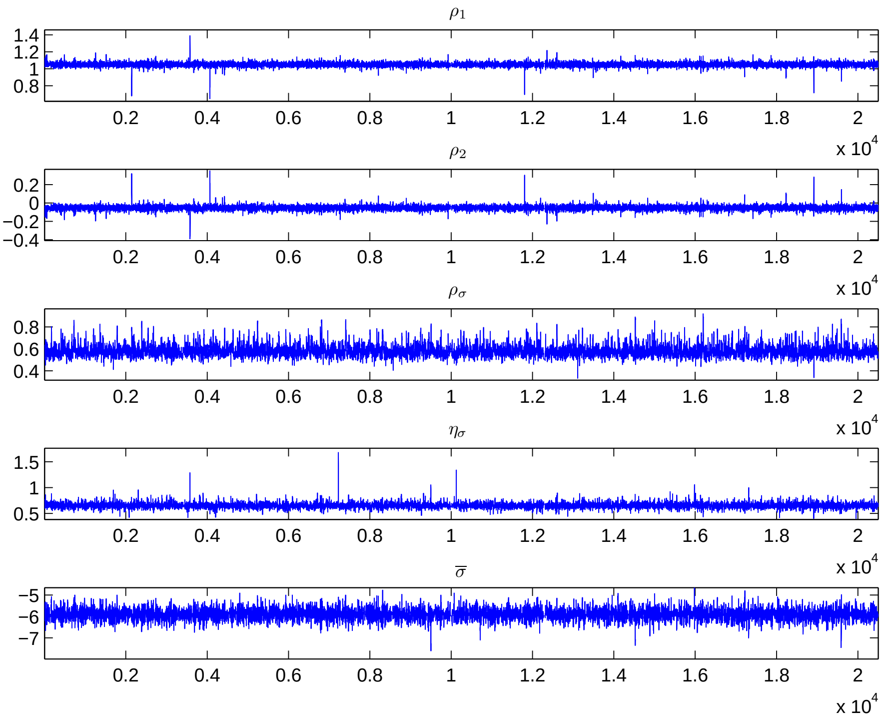


(a) MCMC draws

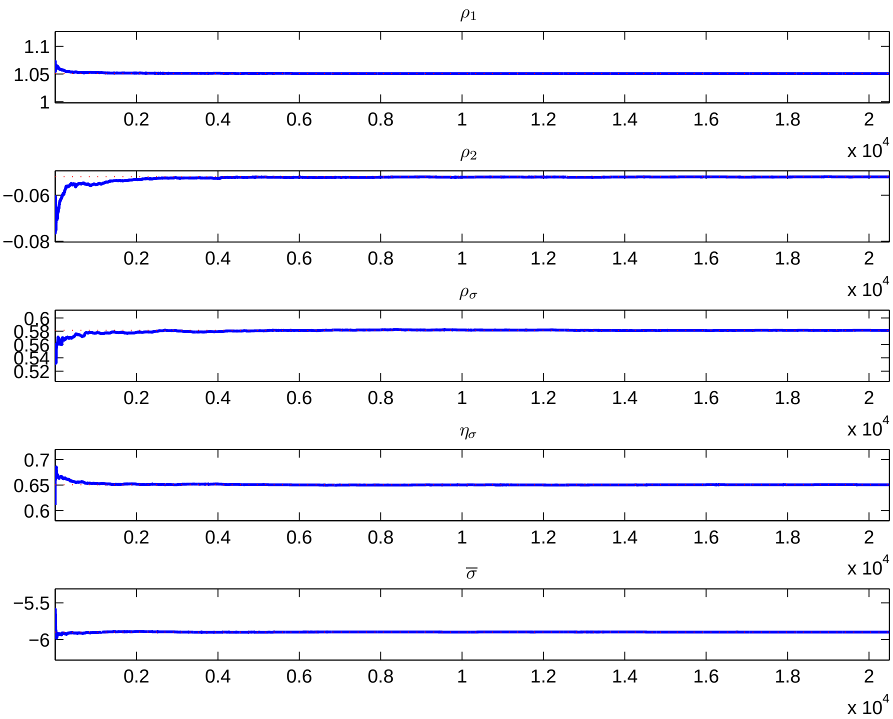


(b) Mean of the parameters over time

Figure 3.9: Evolution of MCMC sampler over time for τ^n

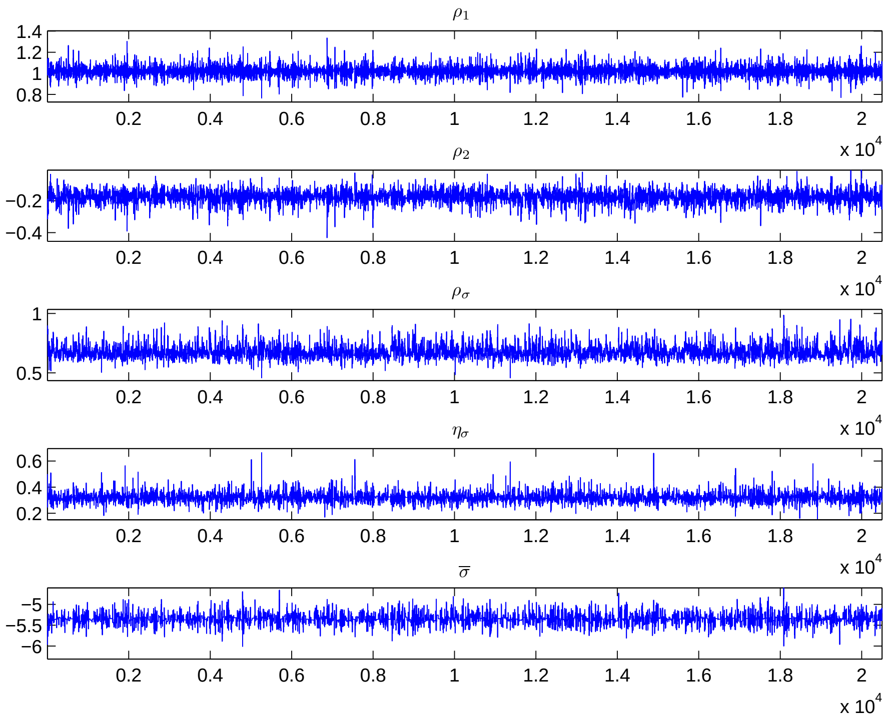


(a) MCMC draws

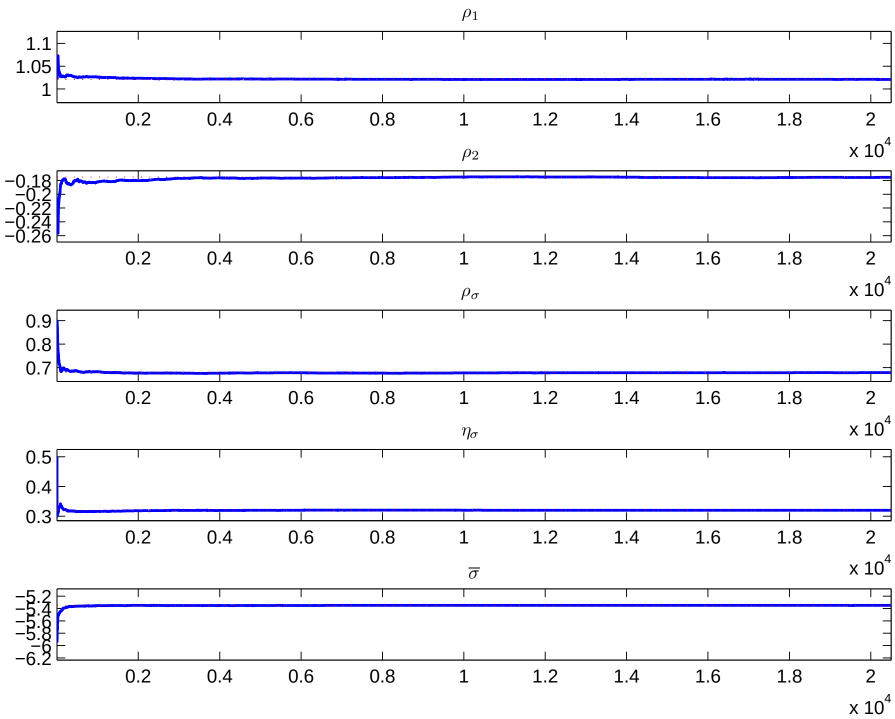


(b) Mean of the parameters over time

Figure 3.10: Evolution of MCMC sampler over time for z

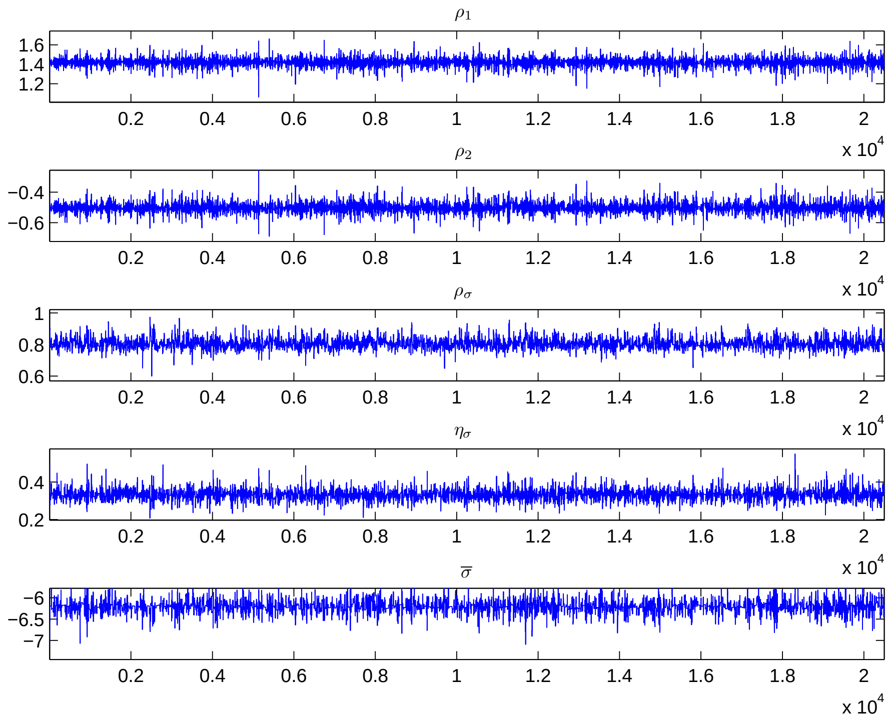


(a) MCMC draws

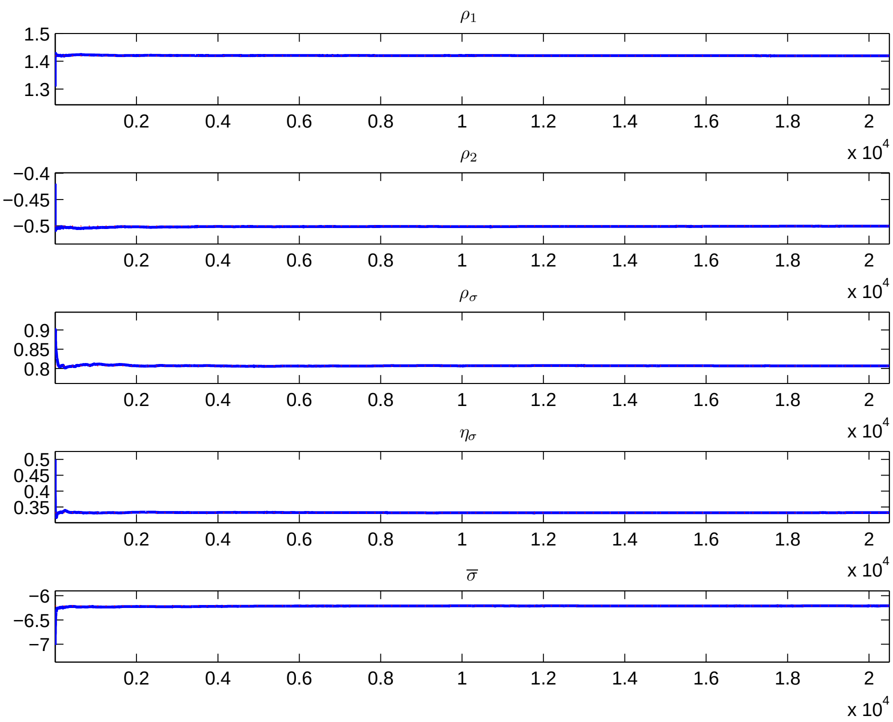


(b) Mean of the parameters over time

Figure 3.11: Evolution of MCMC sampler over time for z^I

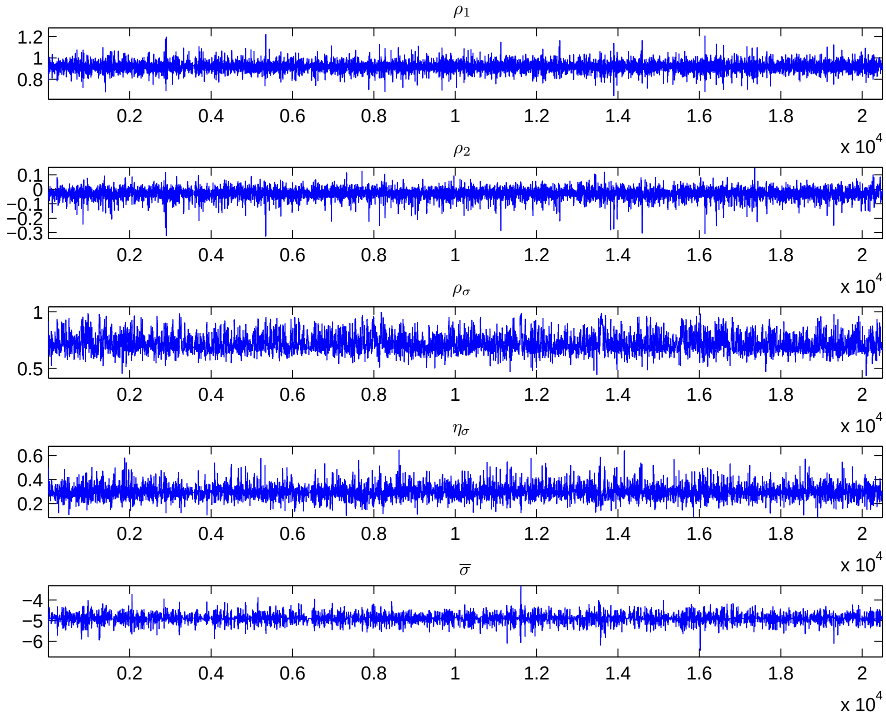


(a) MCMC draws

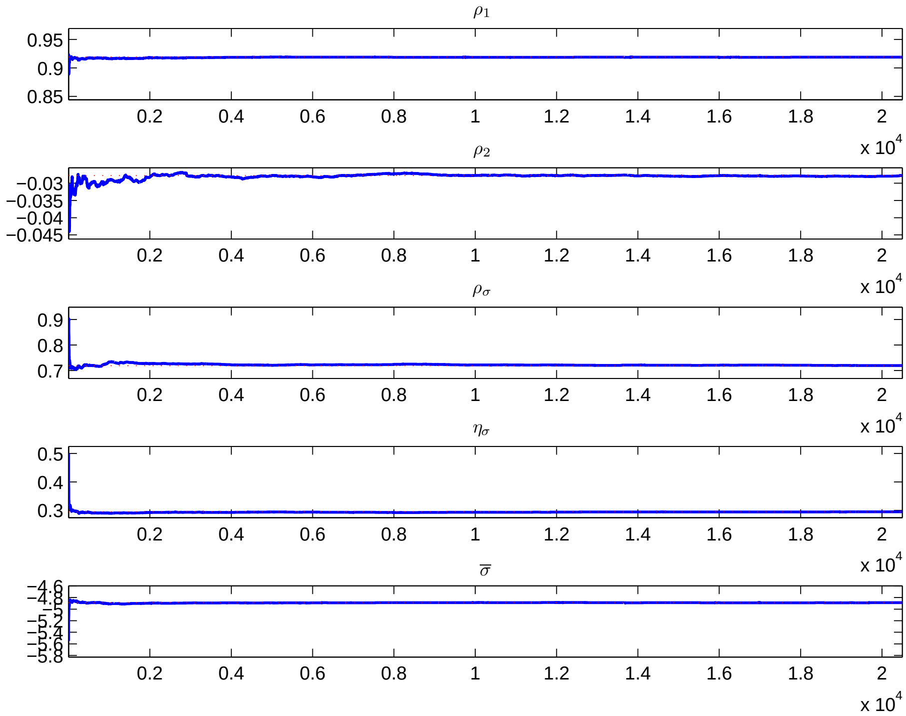


(b) Mean of the parameters over time

Figure 3.12: Evolution of MCMC sampler over time for g

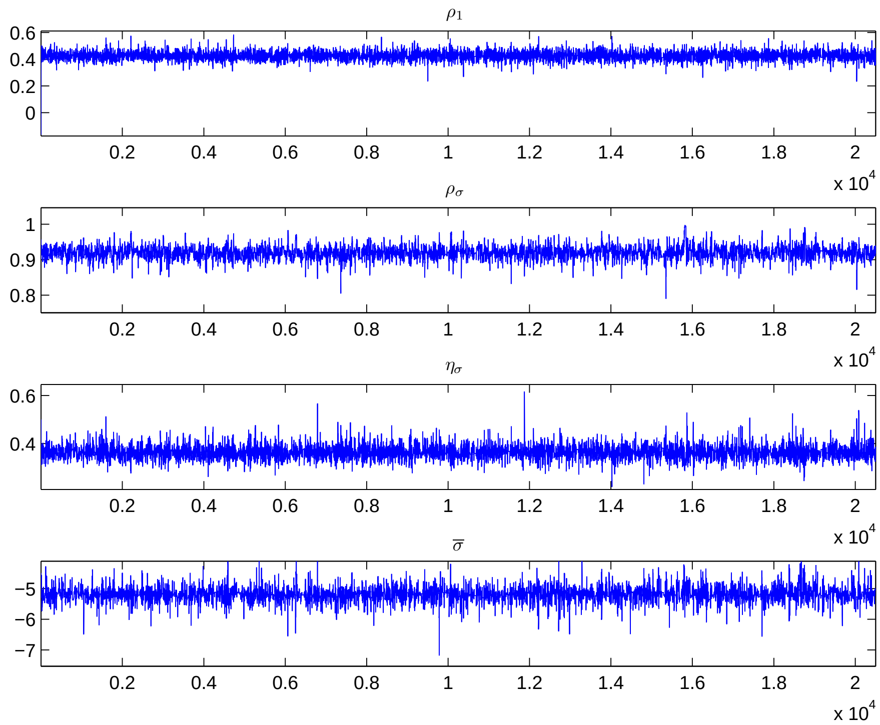


(a) MCMC draws

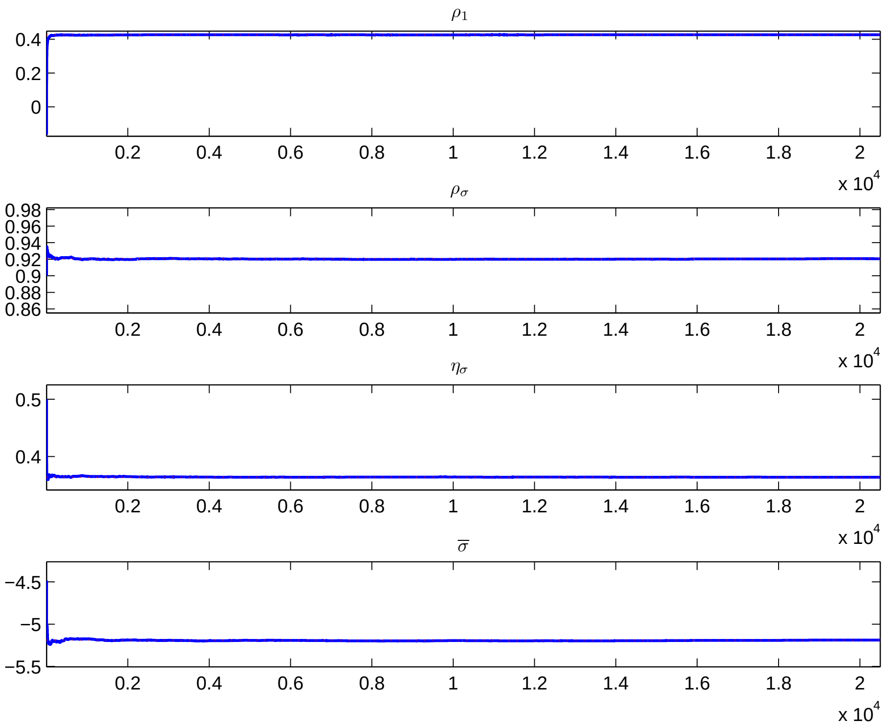


(b) Mean of the parameters over time

Figure 3.13: Evolution of MCMC sampler over time for m



(a) MCMC draws



(b) Mean of the parameters over time

Model Misspecification Diagnostics

Following Kim et al. (1998), we can test the specification of our SV-model. Using N draws from the prediction density $p\left(x_t|x^{t-1};\Theta\right)$, we can compute the probability that x_{t+1}^2 will be less or equal than the actually observed value of $\left(x_{t+1}^{obs}\right)^2$:

$$\Pr\left(x_{t+1}^2 \leqslant \left(x_{t+1}^{obs}\right)^2 \middle| x^t\ ;\Theta\right) \simeq u_{t+1} = \frac{1}{N}\Pr\left(x_{t+1}^2 \leqslant \left(x_{t+1}^{obs}\right)^2 \middle| x^t, \sigma_{t+1|t}\ ;\Theta\right)\ , \quad (3.42)$$

$\forall t = 1, \ldots T-1$. If the SV-model is correctly specified, the sequence of u_t converges in distribution to *i.i.d.* uniform variables as the number of particles N goes to infinity (Rosenblatt, 1952). Under the null hypothesis of a correctly specified model, the u_t can be transformed to *i.i.d.* standard normal variables using the inverse normal CDF. Hence, we can perform a simple test for misspecification by testing the resulting series for their normality. Figure 3.14 shows the corresponding QQ-plots.

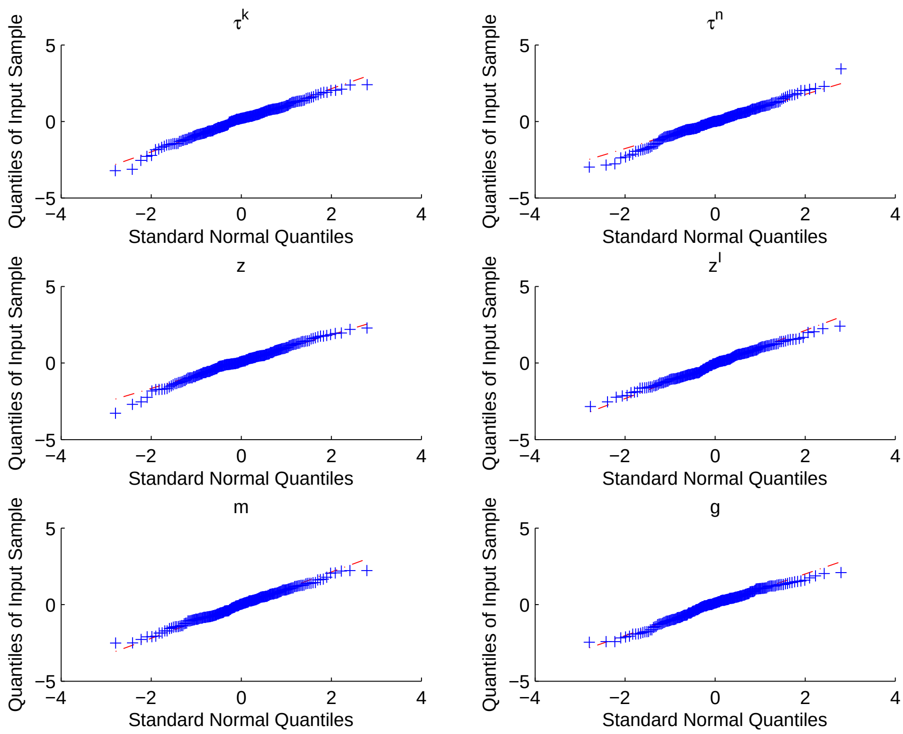
Table 3.12 presents the results from three commonly used normality tests. In general, a correct specification of the model tends to not be rejected. Only for z, the Jarque-Bera and the Kolmogorov-Smirnov tests reject normality. However, this effect is driven by the outliers visible in the bottom left corner of Figure 3.14. In contrast, when shutting off the time-varying volatility and setting the volatility to its unconditional mean, the specification is generally rejected (results are not shown here).

Table 3.12: Tests for model misspecification

	JB	KS	SW
τ^k	0.066	0.039**	0.125
τ^n	0.141	0.960	0.135
z	0.037**	0.035**	0.085
z^I	0.377	0.076	0.586
g	0.500	0.747	0.528
m	0.052	0.377	0.012**

Note: Asterisks indicate significance at the 5% level. JB refers to the Jarque and Bera (1987)-test, KS refers to the Kolmogorov (1933)/Smirnov (1948)-test, and SW refers to the Shapiro and Wilk (1965)-test.

Figure 3.14: QQ-plots



Notes: From left to right and top to bottom: capital taxes, labor taxes, TFP, investment-specific technology, monetary policy shocks, and government spending.

Chapter 4

Central Bank Communication on Financial Stability

4.1 Introduction

The global financial crisis has triggered heated discussions on how best to achieve financial stability in the future. An important role in that regard has been assigned to central banks, many of which have explicit financial stability mandates. In the light of this, a large number of central banks have communicated extensively on financial stability-related matters, e.g. through the publication of Financial Stability Reports (FSRs) and financial stability-related speeches and interviews.

The aim of the current chapter is to shed light on the potential effects of central bank communication about financial stability. It takes a financial market perspective and studies how financial sector stock indices react to the release of such communication, given that the financial sector is one of its main addressees. Doing so, it covers a large number of countries over nearly one and a half decades, and studies the effects of FSRs as well as of speeches and interviews by central bank governors.

An assessment of the effects of financial stability-related communication requires a view on its aims. In line with the aims put forward by Blinder et al. (2008), we focus on the potential of such communication to "create news" and to "reduce noise". A number of central banks have specified the purpose of their FSRs. The ECB's reports, for instance, aim "*to promote awareness in the financial industry and* *among the public at large of issues that are relevant for safeguarding the stability of the euro area financial system. By providing an overview of sources of risk and vulnerability for financial stability, the Review also seeks to play a role in preventing financial crises*" (European Central Bank, 2010, p.7).[1] In light of these statements, it is interesting to study to what extent the views that a central bank expresses in its communications get reflected in the markets. For instance, if the central bank expresses a rather pessimistic view about the prospects for financial stability, and this view gets heard in financial markets, we would expect that stock prices for the financial sector decline. In that sense, these communications "create news". The other motive, to "reduce noise", should then be reflected in market volatility, in the sense that a communication by the central bank should contribute to reducing uncertainty in financial markets, thereby reducing volatility.

But why, and through what channels should central bank communications have an effect on financial markets at all? A number of factors could come into play here. First, the central bank is obviously an important player in financial markets. For instance, if it is ready to change its policy rates, it can directly affect asset prices. Its communication can therefore exert effects through what has been labelled the "signalling channel" in the literature on foreign exchange interventions (e.g. Kaminsky and Lewis, 1996). Second, the analyses that feed into the communications are potentially of high quality, and there are few other institutions communicating about financial stability, such that a central bank publication might indeed contain news. Thus, a co-ordination channel might be at play, whereby communication by the central bank works as a co-ordination device, thereby reducing heterogeneity in expectations and information, and thus inducing asset prices to more closely reflect the underlying fundamentals, a channel that has also been found to be important to explain the effect of foreign exchange interventions (Fratzscher, 2008; Sarno and Taylor, 2001). This channel might imply that communications have longer-lasting effects, as they might change the dynamics in financial markets.

To conduct the empirical analysis, the chapter constructs a unique and novel

[1] In a similar vein, the Bank of England's FSRs aim "*to identify the major downside risks to the UK financial system and thereby help financial firms, authorities and the wider public in managing and preparing for these risks.*" See `http://www.bankofengland.co.uk/publications/fsr/index.htm`.

database on communication comprising more than 1000 releases of FSRs and speeches and interviews by central bank governors from 37 central banks and over the past 14 years. We not only identify the precise timing of these communications, but we also determine their content. We employ a computerized textual-analysis software (called DICTION 5.0), which allows us to grade each of the central bank financial stability statements, based on different semantic features, according to the degree of optimism that is expressed.

A first striking finding from this classification is that the tone of FSRs had continuously become more optimistic after 2000, reaching a peak already in early 2006 and becoming more pessimistic thereafter. This stylized fact, together with formal tests conducted in the chapter, suggests that FSRs comment on the current market environment, but also contain forward-looking assessments of risks and vulnerabilities.

The chapter's findings suggest that communication about financial stability has important repercussions for financial sector stock prices. Moreover, there are clear differences between FSRs, on the one hand, and speeches and interviews, on the other. FSRs clearly create news in the sense that the views expressed in FSRs move stock markets in the expected direction. This effect is quite sizeable as, on average, FSR releases move equity markets by more than 1% during the subsequent month. Another important finding is that FSRs also reduce noise, as market volatility tends to decline in response to FSRs. These effects are particularly strong if the FSR contains an optimistic assessment of the risks to financial stability, when FSRs are found to move equity markets upwards in up to two thirds of the cases. Speeches and interviews, in contrast, have only modest effects on stock market returns, and cannot reduce market volatility.

However, the effects of FSRs and speeches crucially depend on market conditions and other factors. Importantly, during the financial crisis, FSRs were moving financial markets less than before the crisis, while speeches by governors did move financial markets. Finally, the results indicate that financial stability communication of central banks influences financial markets primarily via a coordination channel, i.e. it provides relevant information which exerts a significant and persistent effect on markets.

The chapter shows that while the release schedule of FSRs is pre-scheduled, speeches and interviews are a much more flexible communication tool. For instance, their number is clearly positively correlated with financial market volatility. Given their flexibility, speeches and interviews by definition carry some surprise element. Since it is mostly at the discretion of the central bank governors whether or not to make statements about financial stability, the fact that a governor feels compelled to raise financial stability issues in a speech or an interview can therefore be an important additional news component. In contrast, due to the fixed release schedule for Financial Stability Reports, financial markets expect statements about financial stability issues on the release days. There might be surprising elements in their content, but the mere fact that the FSR is released does not come as a surprise. This difference might be at the heart of the different effects of the two instruments on market volatility.

The empirical findings of the chapter raise a number of policy issues. Communication on financial stability issues by a central bank has been and will likely be watched even more closely in the future, and thus can potentially have an important influence on financial markets. Does this imply that central banks should limit transparency and their communication on certain financial stability issues, as argued by Cukierman (2009), or does this make the case for enhanced transparency and accountability, as argued by others? The findings of the chapter underline that communication by monetary authorities on financial stability issues can indeed influence financial market developments. Yet the findings also show that such communication entails risks as they may unsettle markets. Hence central bank communication on financial stability issues needs to be employed with utmost care, stressing the difficulty of designing a successful communication strategy on these matters.

The chapter proceeds in Section 4.2 by outlining a more general motivation and relating the current chapter to the existing literature. Section 4.3 explains the dataset underlying the empirical analysis. In particular, it reports how the measures for central bank communication have been extracted and quantified. It also shows how the incidence and the content of the communications relate to the external environment, and presents the event study methodology that we employ. Section 4.4 discusses the empirical results and implications, and presents robustness tests. Section 4.5 concludes.

4.2 Motivation and Literature

Given the important role of monetary authorities for financial stability, corresponding central bank communication has always played an important role as a policy instrument, for mainly three reasons. First, financial markets are inherently characterized by asymmetric information and co-ordination problems, characteristics which lie at the heart of the potential risks to financial stability. To address these problems, transparency and communication are crucial. In particular, the central bank can be much more effective in promoting financial stability if it has established a reputation that its analysis and communication are of high quality. Accordingly, communication also serves the role of making the central bank credible. Finally, any body that is entrusted with financial stability tasks will need to be accountable, which calls for a clear mandate, and a transparent conduct of the assigned task. Although Oosterloo and Haan (2004) found that there is often a lack of accountability requirements for central banks' financial stability objectives, this is very likely to change in the future, once financial stability has become a more important and explicit objective of central banks.

As argued by Born et al. (2011b), these aspects of communication for financial stability do therefore closely resemble the role of monetary policy-related communication, as established in the recent literature on central bank communication (see e.g. Blinder et al., 2008; Ehrmann and Fratzscher, 2007a; Gosselin et al., 2009). Also in the monetary policy sphere, communication serves i) to make central banks credible (mirroring the importance of financial stability communication for reputational purposes), ii) to enhance the effectiveness of monetary policy (just like good financial stability communication can contribute to financial stability), and iii) to make central banks accountable.

While being very similar along these three dimensions, there are also differences between monetary policy-related and financial stability-related communication. Central banks have become much more transparent about their conduct of monetary policy over the last decades, along with an increasing importance given to communication. There is a debate on possible limits to central bank transparency (e.g. Mishkin, 2004; Morris and Shin, 2002; Svensson, 2006), but the arguments are much more contentious than in the case of financial stability-related communication. As demonstrated by

Cukierman (2009), a clear case for limiting transparency can be made when the central bank has private information about problems within segments of the financial system. Release of such information may potentially be harmful, e.g. by triggering a run on the financial system. This suggests that policy makers need to be even more careful when designing their communication strategy with regard to their financial stability objectives.

While the literature on central bank communication for monetary policy purposes has been growing rapidly over the recent decade, the communication on financial stability has received considerably less attention. Svensson (2003) argues that through the publication of indicators of financial stability in FSRs, central banks can issue early warnings to economic agents, thereby ideally preventing financial instability from materializing, and thereby ensuring that financial stability concerns do not impose a constraint on monetary policy. Cihak (2006, 2007) provides a systematic overview of FSRs as the main communication channel that central banks use for this purpose. He documents, on the one hand, that the reports have become considerably more sophisticated over time, with substantial improvements in the underlying analytical tools, and on the other hand, that there has been a large increase in the number of central banks that publish FSRs. The frontrunners are the Bank of England, the Swedish Riksbank, and Norges Bank (Norway's central bank), all of which started publication in 1996/1997. It is probably not a coincidence that these three central banks are typically also listed in the group of the most transparent central banks with regard to monetary policy issues (Dincer and Eichengreen, 2009; Eijffinger and Geraats, 2006). In the meantime, around 50 central banks are now releasing FSRs.

A first empirical analysis of FSRs has been conducted by Oosterloo et al. (2007), with the aim to understand who publishes FSRs, for what motives, and with what content. Their results indicate that there are mainly three motives for publication, namely to increase transparency, to contribute to financial stability, and to strengthen co-operation between different authorities with financial stability tasks. They also find that the occurrence of a systemic banking crisis in the past is positively related to the likelihood that an FSR is published.

Even less work has been done with regard to the effects of financial stability-related communication. To our knowledge, the only exception is Allen et al. (2004), who conducted an external evaluation of the Riksbank's work on financial stability issues, and came up with a number of recommendations, such as making the objective of the Riksbank's FSRs explicit, providing the underlying data, or expanding the scope of the FSR to, e.g., the insurance sector. Born et al. (2011b) primarily deal with the conceptual issues of communication about macroprudential issues, making a case for clarity, transparency and predictability, in particular outside crisis times, and stressing the importance to manage expectations by clearly communicating what macroprudential policy can and what it cannot do. While they also provide some empirical evidence as to the role of central bank communication for financial markets, the present chapter goes much deeper in analyzing how central bank communications about financial stability are received in financial markets.

4.3 Measuring Communication and the Effects on Financial Markets

This section introduces the dataset that we develop to study the effects of financial stability-related communication. We start by explaining the choice of data frequency, the sample of countries and time that we use, and the choice of the financial sector stock market indices as our measure for financial markets. Subsequently, we describe the process for identifying the relevant communications, how their content is coded, and the econometric methodology.

4.3.1 Choice of Data Frequency, Data Sample and Relevant Financial Markets

We are interested in the effects of financial stability-related communication on financial markets. A first choice that is required relates to the frequency of the analysis. Given the speed of reactions in financial markets, it is necessary to identify the timing of the events as precisely as possible. Identification of a precise time stamp will allow for an analysis in a very tight time window around the event, thereby ensuring that the market reaction is not distorted by other news. We opted for a daily frequency for two practical reasons. First, given the aim to provide a cross-country study over a relatively long horizon, financial market data are not consistently available at higher frequencies. Second, the identification of the precise days of the release of central bank communications has already not been trivial in many cases, whereas the identification of the exact time of the release within a day is largely impossible. While a higher frequency might have been desirable, it is important to note that the daily frequency is commonly employed in the announcements effect literature - for instance, two classic references with regard to the effect of monetary policy on stock markets, Rigobon and Sack (2004) as well as Bernanke and Kuttner (2005) both use daily data.

The sample of countries and the time period of the study have been determined on the basis of the release of FSRs. We tried to identify the release dates of the FSRs or relevant speeches or interviews by central bank governors for all those central banks listed in Cihak (2006, 2007), i.e. for all central banks which release FSRs. We succeeded to identify such release dates for 35 countries, 24 of which are advanced economies according to the IMF's country classification. Additionally, we included the euro area, as well as the United States as the only country that does not release an FSR, restricting ourselves to studying the effect of speeches and interviews in this case. In total, our sample therefore covers 37 central banks (see Table). Our sample starts in 1996, i.e. the year when the first FSR was released by the Bank of England. The data were extracted in October 2009, such that the sample ends on September 30, 2009.

As to the selection of a financial market that shall be subject of this study, we opted for stock market indices relating to the financial sector, as we expect that empirical effects of financial stability communication should be most easily detectable for this sector. Such data are available from Datastream back to 1996, i.e. to the start of our sample period, for all the countries in our sample. This choice is partially owed to the large cross-country dimension and the need to get historical data for nearly one and a half decades, which limited the availability of less traditional market measures, such as implied volatilities or expected default frequencies (EDFs). While the link of these measures to financial stability would have been relatively direct, we hope that the financial sector stock indices (using MSCI indices) provide a measure that is reasonably closely related to financial stability issues, too. All stock indices are expressed in local currency, given that we are interested in the response of national financial markets to national communication. We will furthermore show that our results are robust to using the overall stock market indices, rather than focusing on the financial sector stocks alone.

4.3.2 Choice and Identification of Communication Events

At the core of this chapter is a measure of communication events that quantifies the content of communication. We focus on the two most important channels of communication about financial stability issues, namely FSRs and speeches and interviews. FSRs are typically relatively comprehensive documents that discuss various aspects of financial stability. They normally begin with an overall assessment of financial stability in the respective country, often including an international perspective. They usually contain an evaluation of current macroeconomic and financial market developments and the assessment of risks to banks and systemically relevant non-banking financial institutions. Cihak (2006) calls these sections the "core" part of an FSR and differentiates them from the "non-core" part that includes research articles on special issues, often written by outside experts. The weights attributed to these two parts vary considerably across central banks. The spectrum ranges from FSRs that only cover the core part (e.g. Norway) to FSRs which only consist of articles covering a special topic (e.g. France). Most central banks lie somewhere in between this range and are usually closer to the first type. Typically, FSRs are published twice a year, i.e. are relatively infrequent communications.

A second important channel for central banks to communicate about financial stability issues is to give speeches and interviews. By their very nature, these are much more flexible than FSRs. Their timing can be chosen flexibly (Ehrmann and Fratzscher (2007b, 2009) have shown this for monetary policy-related speeches), and their content can be much more focused. Of course, this is also due to the fact that they are much shorter than FSRs.

As we are interested in testing the response of financial markets to central bank communication, we need to identify the release dates as a first step (recall that we will conduct the analysis at a daily frequency, hence there is no need to identify the timing within a given day - as long as the release takes place before markets close). As to FSRs, we carefully ensured a proper identification of their release dates, mainly based on information provided on central banks' websites and by central bank press offices, and complemented with information from news reports about the release of FSRs as recorded in Factiva, a database that contains newspaper articles and newswire reports from 14,000 sources. As shown in Table , the dataset contains information on 367 FSRs. The increasing tendency of central banks to publish FSRs is reflected in this database. Starting from less than 10 FSRs per annum in the 1990s, we could identify around 50 FSRs each year in the mid 2000s (note that the drop in numbers in 2009 is entirely due to the fact that the sample ends in September, i.e. covers only three quarters of the year). We tried to be as encompassing as possible with respect to the country coverage. The early publishers are obviously represented more frequently, with 20 and more reports, whereas "late movers" have far fewer observations, down to 1 for the case of the Bank of Greece, which published its first FSR in June 2009.[2]

To identify speeches and interviews is more difficult. Our objective is to extract all relevant public statements that relate to financial stability. For tractability reasons, we restricted our search to speeches by the central bank governor - even in cases where a central bank has a member of its governing body that has an explicit assignment regarding financial stability. We used Factiva and extracted all database entries containing the name of the policy maker together with some keywords that appear with certain regularity in the editorials of the FSRs.[3] From all hits obtained, we extracted those containing statements by the relevant policy maker with a reference to financial stability issues. Since newswire reports typically record the precise time stamp, we were in a position to allocate the speeches and interviews to the appropriate trading days. Communications during weekends were allocated to the subsequent Monday, communications in the evening - such as dinner speeches - to the subsequent trading day. Furthermore, we very carefully chose only the first report about a given statement, which typically originated from a newswire service. This choice has

[2]Although we could not identify the FSR release dates for Indonesia and the Philippines, we kept them in the sample as we were able to identify a number of relevant speeches; note that dropping these two countries from the sample does not affect our results in any substantive way.

[3]To be precise, we used the following search terms: "financial stability or systemic or systemically or crisis or instability or instabilities or unstable or fragile or fragility or fragilities or banking system or disruptive or imbalances or vulnerable or strains".

the advantage that the reporting is very timely, usually comes within minutes of each statement, and that it is mostly descriptive without providing much analysis or interpretation. To avoid double counting, we discarded all subsequent reports or analysis of the same statement.

A number of issues are worth noting about this data extraction exercise. First, the search was conducted only in English language. We might therefore not have discovered all statements, if these were made and reported upon exclusively in other languages. However, due to the fact that Factiva contains also newswire reports and due to the extensive coverage of this topic by newswires, this issue should not be very problematic.

Second, one can easily think of other keywords to use in the database search. We have experimented with larger sets, e.g. including also the terms "volatile", "volatility", "risk", "adverse" or "pressures". However, the additional hits typically related to monetary policy communications (such as central bank governors talking about inflationary "pressures", "risks" to price stability, etc.), such that the resulting dataset on financial stability communications was basically unaltered.

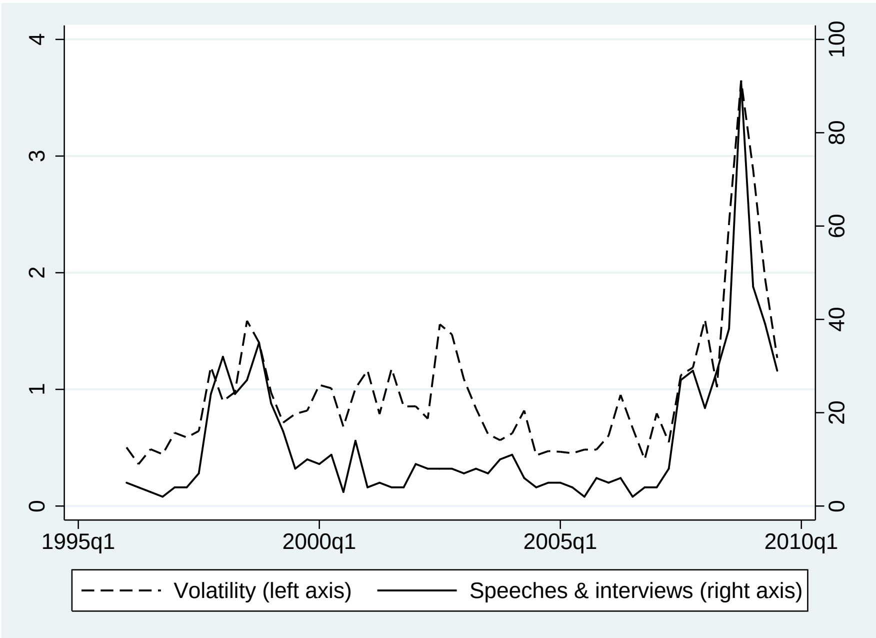
Third, the news sources might be selective in their reporting, thus possibly not covering all relevant statements. However, given the sensitivity of the topic and the importance that it has for financial markets, we are confident that the coverage is close to complete. Furthermore, as we are interested in testing the market response to communication, it makes sense to focus only on those statements that actually reach market participants, and this is best achieved by looking at prominent newswire services.

Fourth, our news sources may wrongly report or misinterpret a statement by policy makers. Again, our objective is to assess communication from the perspective of financial markets and therefore we analyze the information market participants actually receive.

The resulting dataset contains 768 communication events. The breakdown by year in Table reveals large time variations, with a massive increase in the number of speeches in 1998, i.e. during the Asian and the Russian crisis, as well as during the financial crisis of 2007-2010. This suggests that the occurrence of speeches and interviews is responsive to the prevailing circumstances, which is in stark contrast to

FSRs, which are typically released at pre-specified dates. Speeches and interviews do therefore provide the central bank with a very flexible instrument to communicate financial stability concerns, as their timing can be chosen flexibly.

Figure 4.1: Stock market volatility and the occurrence of speeches and interviews



Notes: The figure shows the total number of speeches and interviews in all countries in a given quarter on the right-hand axis (solid line), and the standard deviation of daily returns of the global financial stock index in each quarter on the left-hand axis (dashed line).

Figure 4.1 provides a first stylized fact of the relation between financial markets and the frequency of financial-stability related speeches and interviews, by plotting their total number in all countries in a given quarter on the right-hand axis, and the standard deviation of daily returns of the global financial stock index in each quarter on the left-hand axis. The evolution of the two lines is extremely close, suggesting that communication intensifies in times of financial market turbulence.

Given the low frequency nature of Figure 4.1, we provide a more formal test of the higher frequency-relationship between volatility and the occurrence of speeches in Table 4.2. The table calculates the cumulated stock market returns and the standard deviation of daily stock market returns preceding the communication events, and compares them to equivalent figures for non-event days (with tests for statistically significant differences given in the columns denoted by "Diff"). The left part of the table contains the results for FSRs, the right part for speeches and interviews. The different rows of the table relate to different time windows prior to the event, with the first row measuring returns on the day prior to the event, the second row on the 2 days prior to the event, and so on. Standard deviations are calculated for time windows exceeding 3 days. The non-event comparison figures are calculated for a sample where no communication event has occurred in the preceding 60 business days, and no communication event follows in the subsequent 60 business days. The sample is furthermore restricted to non-overlapping observations.

The picture that resulted from Figure 4.1, i.e. that the occurrence of speeches and interviews is closely related to stock market volatility, is confirmed in the very last set of columns in Table 4.2: on days before an event ("event days"), volatility is substantially higher than on non-event days, with the difference being statistically significant at the 1% level throughout all time windows considered. This is in contrast to the results for the FSRs, the publication schedules of which, as we know, are pre-determined. Even though there are some time windows where the volatility is statistically significantly different, the results are far less consistent. Furthermore, if anything, market volatility tends to be lower on event days than on non-event days, a pattern which is most likely driven by the fact that most central banks started to release their FSRs in the early 2000s, when market volatility was comparatively low.

A similar comparison for the stock market returns also reveals that communication by central banks intensifies during periods of stock market declines. Whereas the average stock return prior to non-event days is typically positive, it is on average negative prior to speeches and interviews, and differences are statistically significant at the 1% level, regardless of the time window. No such pattern is visible for FSRs. The main conclusion from this analysis therefore is that while the release schedule of FSRs is pre-defined, speeches and interviews are a much more flexible communication tool, and react to the current market environment.

In the light of these findings, one might ask whether speeches and their content are predictable, such that financial markets might have priced in the effects already prior to the communication event. In such a case, the subsequent event study methodology would not be appropriate. However, it is important to note that while speeches and interviews occur more frequently in times of high market volatility and declining stock markets, this does not imply any predictability of speeches or their content. Probit models including measures of stock market misalignment, the market trend and its volatility (either directly or their absolute values), do a poor job in predicting the events: the 99*th* percentile of the predicted probabilities of the events is smaller than 0.025.

4.3.3 Measuring the Content of Communications

Once we have identified the communication events, it is necessary to measure their content in order to make the data amenable to econometric analysis. In other words, we want to capture those dimensions and elements of FSRs and speeches/interviews which are relevant for financial market participants and thus will be reflected in asset prices.

A discussion of the various possibilities of achieving this is provided in Blinder et al. (2008). The simplest option consists of assigning a dummy variable that is equal to one on event days, and to zero otherwise. While easily done, this approach limits the analysis severely, namely to a study whether communication affects volatility or absolute returns. If we are interested in the effect of the content of communication, a method for quantification of such content is required. The approach adopted in some part of the literature on monetary policy-related communication, namely to read the communications and code them on various scales, was not feasible for our purposes, given the amount of text that needed to be quantified. We have therefore opted for an automated approach for the current chapter.[4]

We employed the computerized textual-analysis software DICTION 5.0,[5] which searches text for different semantic features by using a corpus of several thousand words, and scores the text along an optimism dimension. This dimension may be

[4] An alternative approach is used by Lucca and Trebbi (2009), where FOMC statements are cut down into small segments of text, the semantic orientation of which is then calculated by checking how often these text segments appear in conjunction with the words dovish or hawkish in a large body of text.

[5] See `http://www.dictionsoftware.com`. Beyond the optimism score, Diction also generates scores on certainty, activity, realism and commonality. As the latter are less immediately relevant for financial stability purposes, and for brevity, we have restricted the analysis to optimism. Results for activity are reported in the working paper version, Born et al. (2010).

important as it provides agents with information about the current state and the prospects of the financial system and underlying risks. The respective scores are computed by adding the standardized word frequencies of various subcategories labelled as optimistic, and by subtracting the corresponding frequencies of pessimistic subcategories. In broad terms, optimism refers to "language endorsing some person, group, concept or event, or highlighting their positive entailments".

This software has been used extensively in communication sciences and in political sciences, e.g. to analyze speeches of politicians (Hart, 2000; Hart and Jarvis, 1997), but has also been applied in the context of central banks (Armesto et al., 2009; Bligh and Hess, 2007). Furthermore, Davis et al. (2006) have used it to measure the reaction of financial markets to earnings announcements, and find a significant incremental market response to optimistic and pessimistic language usage in earnings press releases.

There are a number of advantages of this approach over human coding of the text. First, the software creates a coding that is more mechanical and thus objective, compared to human coding which tends to be more judgmental. While some subjectivity could arise due to the choice of the content of the dictionaries against which a text is assessed, it is important to note that the corpus has been defined based on linguistic theory and without an active participation by the authors of this chapter. Another advantage is the replicability of the coding, which is in stark contrast to human coding, and also allows more text to be added without distorting the scoring process. Third, the automated approach allows a consistent coding of long passages of text, and across a large number of communications. Human coding of long texts with various points is rather difficult, as no part should in principle be given a larger weight in the assessment. Given the breadth of FSRs, this issue is particularly severe in the current application. At the same time, a drawback of the automated approach is that it does not consider the context of the text, and thus cannot generate a "tailor-made" coding for financial stability-related communication.

Based on this computerized textual-analysis software, we computed a score for each individual speech or interview (note that, effectively, we are coding the content of the related news reports, rather than the original source text), and for the overview part of each FSR.[6] Subsequently, we transformed the resulting scores into a discrete variable, which takes the value of -1 for the lowest third of the distribution, a value of 0 for the middle part of the distribution, and the value of $+1$ for the upper third of the distribution. That is, a value of $+1$ denotes a relatively optimistic text, while a value of -1 corresponds to a relatively pessimistic statement. The discretization of scores is required for the subsequent analysis, where we are interested in the market effects of optimistic vs. pessimistic communication, rather than the effect of an incremental change in tone. This transformation was applied for the speeches as well as for the FSRs. Note that we will test for robustness using a very different measurement approach, which also attempts to capture the surprise component contained in the respective communications, as well as (for the parts of the subsequent analysis where a discretization is not required) using the raw optimism scores given by the software.

It is important to note that this implies a *relative* coding, i.e. a given communication is scored in a comparative fashion against the other texts in the sample. However, due to the large sample, both across countries and along the time dimension, our communications cover periods of relative stability and tranquillity, as well as periods of financial market crises or turbulence. Accordingly, the overall sample of text should be relatively balanced, such that text which is coded with plus or minus one should indeed represent a corresponding opinion. We denote the resulting indicators by $I_{it}^{optimism,FSR}$ and $I_{it}^{optimism,speech}$, respectively, where i denotes a given country, and t stands for time. In the appendix, we provide a number of examples of speeches and interviews, and how they were coded.

4.3.4 The Event Study Methodology

What are the effects of FSRs and speeches/interviews on financial markets? The natural econometric approach to test our hypotheses of interest is the event study methodology. We use this methodology because we are interested not only in the

[6] While this overview carries different names across central banks, e.g. editorial, introductory chapter, executive summary, etc., it is rather similar in nature for all FSRs. As mentioned above, the Financial Stability Review of the Bank of France is special in that it only consists of articles covering a special topic. However, given the regular newswire reporting about the Editorial written by the Bank of France Governor, we decided to compute the scores also for the French reports.

contemporaneous effect of financial stability statements, but we also want to know how persistent the effect is over time. We can define the release of an FSR, or the delivery of a speech or an interview as an event. The question we want to address is whether the event affects stock markets in a causal fashion. For that purpose, it is essential that we can compare the stock market evolution following the event to the counterfactual, i.e. a predicted value that we believe would have occurred had the event not happened. A crucial issue in any event study is therefore to find a benchmark model to calculate expected returns, which in turn allows calculation of excess returns.[7] Most event studies look at the effect of events, such as earnings announcements or stock splits, on individual stocks, and use some variant of a factor model, such as the Fama-French (1993) three-factor model, or the Carhart (1997) four-factor model, which extends the previous model by a momentum factor.

Given that we are interested in the evolution of national stock market indices rather than of individual stocks, the book-to-market ratio and the size factor of the Fama-French model are not applicable. Following Edmans et al. (2007) and Pojarliev and Levich (2007), we start by defining normal returns as:

$$\begin{aligned} R_{it} &= \gamma_{0i} + \gamma_{1i} R_{it-1} + \gamma_{2i} R_{mt-1} + \gamma_{3i} R_{mt} + \gamma_{4i} R_{mt+1} \\ &\quad + \gamma_{5i} D_t + \gamma_{6i} T_{it-1} + \gamma_{7i} S_{it-1} + \gamma_{8i} M_{it-1} + \varepsilon_{it} \ , \end{aligned} \tag{4.1}$$

where R_{it} is the daily local currency return on the financial sector stock market index for country i on day t, R_{mt} is the daily US dollar return on Datastream's global financial sector stock market index, and D_t denotes dummy variables for Monday through Thursday. T_{it-1} stands for the trend in stock markets over the 20 days prior to the event, S_{it-1} for the standard deviation of daily stock market returns over the 20 days prior to the event, and M_{it-1} for the "misalignment" of stock indices on the day preceding the event, measured as the percentage deviation of the stock indices from their national average over the entire sample period.

The first 5 factors follow Edmans et al. (2007). The lagged index return controls for possible first-order serial correlation. The global stock market index is meant to capture the effects of international stock market integration, and since some indices

[7] For overviews of the event study literature see, e.g., MacKinlay (1997) or Kothari and Warner (2007).

might be lagging or leading the world index, Edmans et al. (2007) not only include the contemporaneous global returns, but furthermore a lead and a lag. The last three terms are owed to earlier event studies on exchange rates such as Pojarliev and Levich (2007) or Fratzscher (2008). The trend factor attempts to allow for persistence in stock market movements, and is therefore closely related to the momentum factor in the Carhart four-factor model. The inclusion of the standard deviation is an attempt to capture the effect of market volatility. Finally, the misalignment factor is based on the idea that there might be booms or busts in stock markets, and that over a sufficiently long sample, there could be some mean reversion (albeit possibly allowing for a drift). We test for robustness to the exclusion of these last three terms, given that they are derived from the exchange rate literature rather than the stock market event studies, and find our results to be qualitatively unaltered.

Model (4.1) is estimated country by country, only including days that were neither preceding nor preceded by communication events for 60 days (in each direction). Based on the estimated parameters (denoted by hats), it is then possible to calculate excess returns on event days as

$$\begin{aligned}\widehat{\varepsilon}_{it} \;=\; R_{it} - \Big(&\widehat{\gamma}_{0i} + \widehat{\gamma}_{1i}R_{it-1} + \widehat{\gamma}_{2i}R_{mt-1} + \widehat{\gamma}_{3i}R_{mt} + \widehat{\gamma}_{4i}R_{mt+1} \\ &+ \widehat{\gamma}_{5i}D_t + \widehat{\gamma}_{6i}T_{it-1} + \widehat{\gamma}_{7i}S_{it-1} + \widehat{\gamma}_{8i}M_{it-1}\Big) \qquad (4.2)\end{aligned}$$

The hypothesis to be tested is whether communication leads to excess returns in the expected direction, i.e. whether

$$\widehat{\varepsilon}_{it} > 0 \text{ if } I_{it}^{optimism,c} = 1 \quad \text{or} \quad \widehat{\varepsilon}_{it} < 0 \text{ if } I_{it}^{optimism,c} = -1 \ , \qquad (4.3)$$

where the superscript c stands for the two communication types, FSR and speeches or interviews. A more complex approach is required if we want to calculate the *longer-term* effects of communication beyond the event day. While we assume that world markets are exogenous to a communication in an individual country also over extended time windows, this is obviously not the case for the own lag, the recent trend, standard deviation and misalignment: as of the second day, it is necessary to calculate predicted returns for the preceding day, and to plug these into equation

(4.2), thus yielding

$$\begin{aligned}\widehat{\varepsilon}_{it+k} \;=\; & R_{it+k} - \Big(\widehat{\gamma}_{0i} + \widehat{\gamma}_{1i}\widehat{R}_{it+k-1} + \widehat{\gamma}_{2i}R_{mt+k-1} + \widehat{\gamma}_{3i}R_{mt+k} + \widehat{\gamma}_{4i}R_{mt+1+k} \\ & + \widehat{\gamma}_{5i}D_{t+k} + \widehat{\gamma}_{6i}\widehat{T}_{it+k-1} + \widehat{\gamma}_{7i}\widehat{S}_{it+k-1} + \widehat{\gamma}_{8i}\widehat{M}_{it+k-1}\Big)\;. \end{aligned} \tag{4.4}$$

Note that compared to equation (4.2), R_{it-1}, T_{it-1}, S_{it-1} and M_{it-1} have all been replaced by their predicted value in the absence of a communication event. For $k = 0$, the two coincide, whereas for all days $k > 0$, it is important to calculate the appropriate predicted values. Tests for the effects of communication over longer time horizons with a time window of K days then amount to asking whether

$$\sum_{k=0}^{K} \widehat{\varepsilon}_{it+k} > 0 \text{ if } I_{it}^{optimism,c} = 1 \quad \text{or} \quad \sum_{k=0}^{K} \widehat{\varepsilon}_{it+k} < 0 \text{ if } I_{it}^{optimism,c} = -1\;. \tag{4.5}$$

Following common practice in the event study literature, we employ two types of tests for the effects of communications (both described in detail in MacKinlay, 1997). First, we apply a non-parametric sign test to study whether the above conditions hold in more than 50% of all cases. The underlying idea is that by construction - if the factor model is correct - excess returns and cumulated excess returns are on average zero, and that it is equally probable that they are positive or negative. If the events systematically move stock markets in the expected direction, we should find that the excess returns are non-zero, and of the expected sign, in significantly more than 50% of cases. For the second (parametric) test, we compute the average size of the (cumulated) excess returns and test the null hypothesis that they are zero against the alternative.

In a similar vein, to test whether communications reduce noise, i.e. lower stock market volatility, we furthermore test whether

$$\sigma_{\hat{\varepsilon}_{i,t/t+k}} < \sigma_{\hat{\varepsilon}_{i,t-1/t-1-k}} \text{ if } D_{it}^{c} = 1\;, \tag{4.6}$$

with $\sigma_{\hat{\varepsilon}_{i,t/t+k}}$ the standard deviation of daily excess returns in country i from time t to $t + k$, $\sigma_{\hat{\varepsilon}_{i,t-1/t-1-k}}$ their standard deviation over the k days prior to the event, and $D_{it}^{c} = 1$ a dummy variable that is equal to one on the days when a communication of type c is released in country i.[8] Also here, we apply the non-parametric sign test whether the above conditions hold in more than 50% of all cases and the test whether the difference of the two standard deviations is equal to zero.

4.4 The Effects of Financial Stability-Related Communication

This section starts by providing some stylized facts of how the content of FSRs and speeches evolved over time - and to what extent it managed to be forward-looking and identify risks and vulnerabilities rather than reflect market developments (Section 4.4.1). It then proceeds by identifying and testing for the effects of communication on financial markets (Section 4.4.2) and presents a number of sample splits and robustness tests that also sheds further light on the channels trough which communication affects markets (Section 4.4.3).

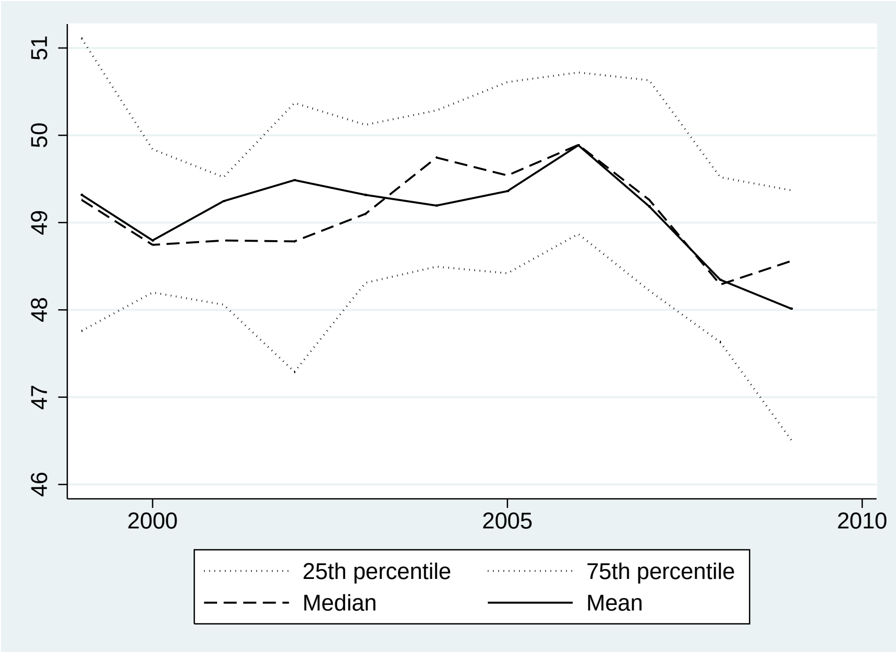
4.4.1 Stylized Facts About Timing and Content of Communication

How did the content of FSRs and speeches evolve over time and across countries? And to what extent was such communication forward-looking rather than reflecting market developments? Figure 4.2 provides an overview of how the optimism expressed in FSRs (upper panel) as well as speeches and interviews (lower panel) has evolved over time. It plots, for each year, the average and median optimism for the respective communication events, as well as the $25th$ and the $75th$ percentiles. Note that the figure for FSRs starts only in 1999, given that in the years before, there were too few FSRs being published to provide a meaningful picture.

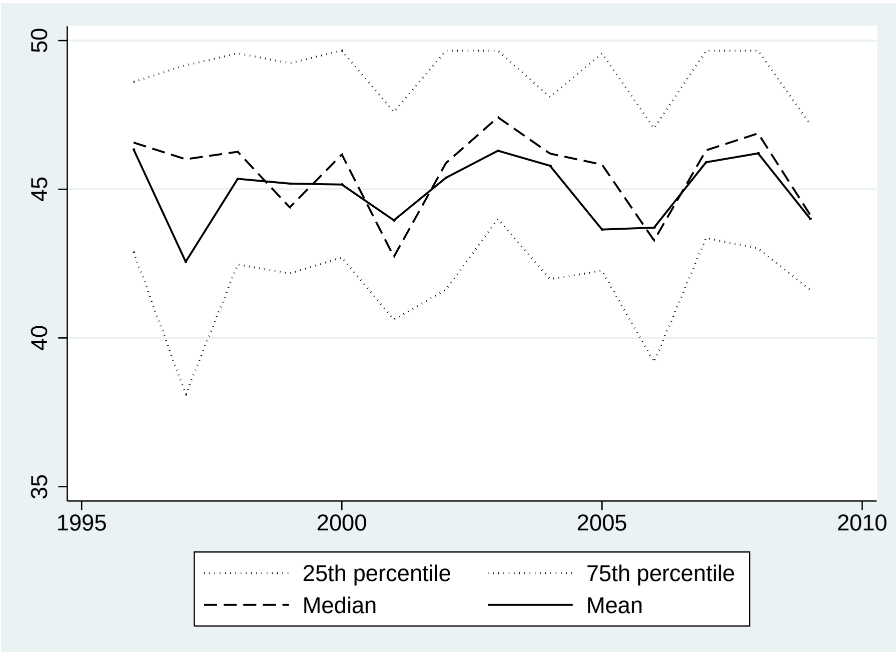
A number of interesting issues emerge from this figure. Most importantly, it is striking that the tone of FSRs had continuously become more optimistic after 2000, reaching a peak in early 2006. This suggests that FSRs contain commentaries on

[8] Excluding the daily excess returns on day t from calculating the post-event standard deviations does not alter our results. This implies that the results are not driven by the initial market reaction on the day of the announcement.

Figure 4.2: Evolution of optimism over time



(a) Financial Stability Reports



(b) Speeches and interviews

Notes: The figure plots the average, median, and 25th and 75th percentiles of the optimism scores for FSRs (upper panel) and speeches and interviews (lower panel) in any given year.

the current market environment, but that they are also forward-looking, with some anticipation of the 2007-2010 crisis. However, there is a relatively large heterogeneity across countries, as shown by the breadth of the scores encompassed by the 25th and the 75th percentiles. This is especially the case for speeches and interviews, which do not seem to follow any obvious pattern over time.[9]

Table 4.3 looks further into the question to what extent the content of communications reflects previous financial market developments, and reports corresponding test results. Separately for FSRs and speeches and interviews, it reports the average return and standard deviation of financial sector stock indices over the usual time windows (from one day to 60 days prior to the event), separately for communications coded as -1, 0 and $+1$ on the optimism scale in columns (1), (2) and (3), respectively. The statistical significance of a test for equality is provided for each pair, i.e. (1) vs. (2), (1) vs. (3), and (2) vs. (3).

The results show that the content of FSRs reflects to some extent prior financial market developments. There is a monotonic relation between the tone of FSRs and the preceding stock market returns: the more optimistic the FSR, the larger have been the preceding returns. However, these differences are typically not statistically significant. At the same time, pessimistic FSRs (i.e. those coded with -1) have, on average, been preceded by considerably larger stock market volatility than neutral or positive FSRs, regardless of the length of the time window, with the differences being highly statistically significant.

Interestingly, no such relations are identifiable for speeches and interviews: there is not a single case where stock market volatility or returns would be related to the content of speeches in a statistically significant manner. If anything, it seems to be the case that there is quite some “leaning against the wind”: the returns preceding optimistic speeches are consistently lower than the returns preceding pessimistic ones, suggesting that a positive picture is given especially in cases of bad stock market performance.

[9]Note that the raw scores cannot be read as direct indications of optimism, as it is not the case that scores below 50 would represent pessimistic text. The interpretation of the scores should be made relative to a large number of texts within the same category.

4.4.2 Effects of FSRs and Speeches/Interviews

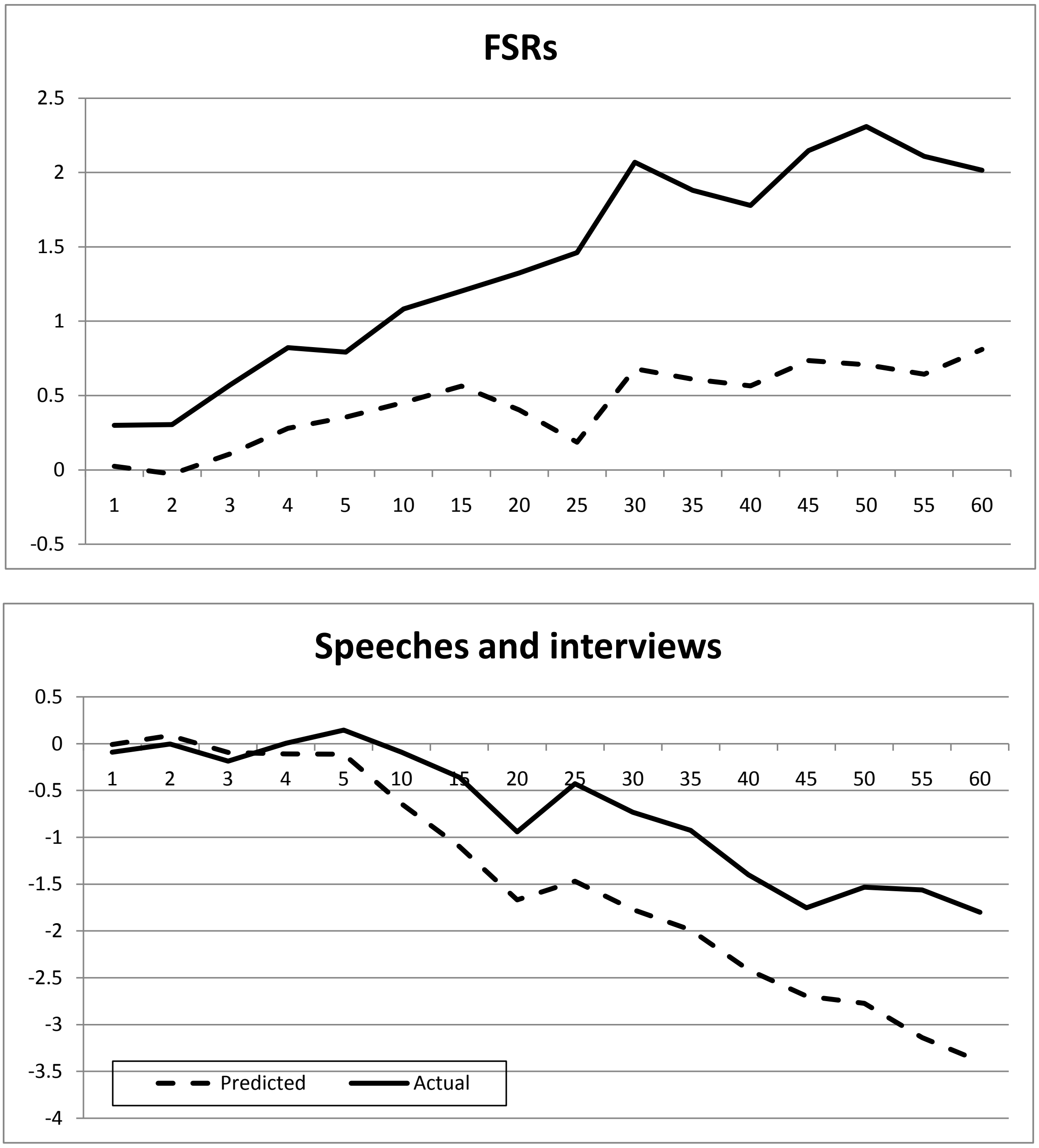
We now turn to the question to what extent central bank communication was affecting financial markets. A first test is provided in Figure 4.3, which compares the actual evolution of stock markets following communication events to the predicted evolution on the basis of the benchmark model (4.1). The upper panel reports the results for the FSRs, the lower panel those for speeches and interviews. The solid line plots the average actual cumulated returns over 60 days following the communication events. The dashed line, in contrast, shows the expected cumulated returns that would result from the benchmark model in the absence of a communication event. To combine pessimistic as well as optimistic communications in one chart, the cumulated returns are multiplied by -1 for pessimistic communications, whereas they are left unchanged for optimistic communications. Accordingly, we would expect the actual returns to lie above the predicted returns after statements if the markets follow the point of view expressed by the central bank (i.e. we observe negative excess returns in response to pessimistic statements, and positive ones in the case of optimistic communications).

The figure provides a compelling picture about the effects of central bank communication. The upper panel for FSRs shows that markets move in the direction of the central bank view, since the actual returns are substantially larger than the predicted returns. Moreover, the effect is quite sizeable economically: for several time windows, FSR releases move equity markets on average by more than 1% in the direction indicated by the FSRs.

Interestingly, expected cumulated returns in this case are relatively close to zero, suggesting the predictions of the benchmark model are close to those of a random walk model. In other words, due to the fact that the release pattern of FSRs is not systematically related to the previous stock market performance, the benchmark model has a hard time in predicting the subsequent returns.

Looking at the lower panel of Figure 4.3, the findings are remarkably different for speeches and interviews. As we have seen above, speeches and interviews typically follow stock market declines, and the model clearly predicts further declines subsequently (the dashed line in the figure). As a matter of fact, actual returns do on average decline after a speech or an interview; however, comparing the expected

Figure 4.3: Cumulated abnormal returns after communication events



Notes: The figure compares the actual evolution of cumulated stock market returns (in %) following communication events to the predicted evolution on the basis of the benchmark model (1). The cumulated abnormal returns are therefore given by the area between the two graphs. The upper panel reports the results for the FSRs, the lower panel those for speeches and interviews. The solid line plots the average actual cumulated returns starting from day 1 after the communication event and up to day 60. The dashed line shows the expected cumulated returns that would result from the benchmark model in the absence of a communication event. The cumulated returns are multiplied by -1 for pessimistic communications, whereas they are left unchanged for optimistic communications.

with the actual evolution, it is also apparent that the stock markets decline by less than expected in the presence of central bank communications. The difference between predicted and actual cumulated returns is substantially smaller than for FSRs, however.

The figure also suggests that central bank communications are potentially affecting financial markets even at very long horizons, given that the gap between predicted and actual cumulated returns is present for the entire horizon of time windows we look at, and begins to narrow only towards the end of the horizon.

The formal test results for the effects of central bank communication are provided in Tables 4 and 5, covering FSRs and speeches and interviews, respectively. The first set of results relates to equation (4.5), i.e. tests whether optimistic statements yield positive excess returns, and pessimistic ones lead to negative excess returns. The first column shows the share of cases in which the condition was met, as well as the results of the non-parametric sign test. Shares above 0.5 would suggest that stock markets move in the direction of the content of communications. The statistical significance is assessed by stars (*** for 1%, ** for 5%, and * for 10% significance) - whereas numbers that are significantly smaller than 0.5 would be characterized by apostrophes (''' for 1%, '' for 5%, and ' for 10% significance).

There is clear evidence that the views represented in FSRs get reflected in financial markets, in significantly more than 50% of all cases. In terms of magnitudes, which are reported in the second column, FSRs generate excess returns on the day of the release of 0.27% on average, and cumulated excess returns up to 1.6% in the longer run, with the largest effects found after 25 to 50 trading days, i.e. after 5 to 10 weeks. Such an effect is indeed sizeable and economically meaningful, in particular when considering that FSRs are generally released twice a year per country.

How are these effects generated? Table 4.4 also provides a breakdown according to the type of the FSR, and reveals that in particular optimistic FSRs affect financial markets. They typically generate positive excess returns, which are furthermore large in magnitude, thus leading to statistically significant estimates. The cumulated excess returns are largest after 55 days, amounting to more than 3%. This suggests that an optimistic assessment provided in FSRs leads to an improvement in stock market sentiment over a fairly long horizon, in a way that is not matched by pessimistic

FSRs leading to a deterioration in sentiment.

Table 4.4 also provides the results for tests whether the release of FSRs lowers stock market volatility, i.e. tests whether condition (4.6) holds, again using both the non-parametric sign test and the parametric test. There is compelling evidence that FSRs do indeed lead to a significant reduction in market volatility.

Moving on to the effect of speeches as reported in Table 4.5, a rather different picture emerges. The effect on returns is less systematic than for FSRs. With some delay, optimistic speeches generate positive excess returns. The effect for pessimistic speeches on returns is, on average, non-existent, however. Of course, this is not to say that no speech would ever exert reactions on financial markets - rather, on average, there seems to be very little effect. At the same time, speeches do not lower stock market volatility - if anything, there is some tendency, especially of optimistic speeches, to somewhat increase it. This suggests that financial stability-related speeches are less able to reduce noise.

To summarize, these findings suggest, first, that communication about financial stability has the potential to affect financial markets. FSRs exert very different effects than speeches and interviews: The views expressed in FSRs get reflected in stock market returns, and in a long-lasting fashion, in particular if the FSR contains an optimistic assessment of the risks to financial stability. FSRs also manage to reduce market volatility somewhat. Speeches and interviews, in contrast, only modestly affect market directions, and do leave market volatility mainly unaffected. An assessment of the effects of these tools therefore needs to clearly distinguish between the two.

4.4.3 Sample Splits and Robustness

We have subjected our benchmark results to a number of sample splits and robustness tests, which we will describe now. There are basically four dimensions to these tests. The first analyzes whether the breadth of the underlying panel dataset masks important heterogeneity, and we test for robustness by introducing various sample splits. The second is concerned with speeches and interviews in particular, and tests whether their effects are different if they are clustered. Third, we test whether our focus on financial stocks is important, or whether the results are robust to using the entire stock market indices. Fourth, we ask whether the split into optimistic and pessimistic content determines our results by providing an alternative way of identifying the content of communication, and by using the raw scores as generated by Diction. All results are provided in Tables 4.6 and 4.7, with FSRs being covered in Table 4.6, and speeches and interviews in Table 4.7. Given the large number of tests, we only show results for a time window of 25 business days.

The first set of results relates to various sample splits. Given the large number of countries and the long time sample, it might be the case that there is substantial heterogeneity across countries or over time that we do not capture in the full sample. The first such split addresses possible cross-country heterogeneity, by re-running the estimation separately for all advanced and all emerging economies (following the IMF's country classification). Results are overall robust. The interesting insight, though, is that there is a reduction in volatility following FSRs by central banks in advanced countries, whereas the main effects on returns originate in emerging countries.

Also the second split along the time dimension reveals interesting patterns. Separate tests for the period prior to the financial crisis 2007-2010 (defining the starting date in September 2007, i.e. with Northern Rock; defining the start of the crisis with Lehman does not affect our results) and the time of the crisis shows that FSRs have exerted no systematic effect on stock markets during the crisis, whereas the effects of speeches and interviews are precisely driven by the period of the crisis, underlining that speeches and interviews may be much more influential during periods of financial stress.

The third sample split intends to identify whether the role of the central bank in financial supervision matters, by testing once for the effects of communication by central banks that do have a formal role in financial supervision, and once for those central banks without such a task. The classification is based on the CBFA index developed in Masciandaro and Quintyn (2009).[10] This differentiation does not seem to play an important role, given that the results are robust, and no major differences overall between the two groups emerge.

[10]This index takes the value 1 if the central bank is not assigned the main responsibility for banking supervision; 2 if the central bank has the main (or sole) responsibility for banking supervision; 3 if the central bank has furthermore responsibility for either insurances or the securities markets; 4 if the central bank has responsibility in all three sectors. We allocate central banks to the group with supervisory functions if their index value is larger than one.

Table 4.7 shows furthermore whether there are differences if speeches and interviews are clustered, i.e. the central bank governor might give a sequence of speeches or interviews in a relatively short time window. Such a sequence might be inherently different from one isolated event. We define a communication event to be part of a cluster if other speeches or interviews occur within 60 days after the event, or have occurred within 60 days before the event. As a matter of fact, these types exert very different effects. Speeches that are part of a cluster are not influencing the market view, and tend to increase market volatility. This is in sharp contrast to the stand-alone speeches, which create news, i.e. move markets along with the views expressed, and tend to do so largely without changing volatility.

The rows of section C in Tables 4.6 and 4.7 present additional robustness tests. First, replacing the financial sector stock indices with the broad national stock market index, we can test whether our results apply more broadly, or are confined to the financial sector. The results are remarkably robust. Furthermore, results are also not sensitive to the precise way we had split the communications into optimistic and pessimistic content. To test for this, we take two routes: First, by defining an alternative approach to discretizing the codes that attempts to control for the expected component contained in the communication, and to construct a surprise measure instead. We do so by means of the following auxiliary regression:

$$C_{it}^{optimism,c} = \alpha_{0i} + \alpha_{1q} + \alpha_2 T_{it-1} + \alpha_3 S_{it-1} + \alpha_4 M_{it-1} + \mu_{it} \ , \tag{4.7}$$

where $C_{it}^{optimism,c}$ denotes the raw Diction coding of a given communication of type c along the optimism dimension, and α_{0i} and α_{1q} are country fixed effects and time fixed effects for each quarter of the sample, respectively. The country fixed effects allow for the possibility that there is a different style in the reporting, thus leading to a different mean coding for each country. Such differences should be well known to observers, and therefore not be a surprise. The time fixed effects control for a common evolution across countries, given that often developments in financial markets are internationally determined. Such common time patterns should also not come as a surprise to financial markets. The last three explanatory factors are as described in benchmark model (4.1), i.e. they control for the trend, for stock market volatility, and for a possible stock market misalignment. We retrieve the residuals $\widehat{\mu}_{it}$ from these regressions, and define a communication to be optimistic if $\widehat{\mu}_{it}$ is above the $66th$ percentile in the distribution, as pessimistic if it is below the $33rd$ percentile, and as neutral otherwise. Even though this classification is very different from the original, unconditional, one, it turns out that the results are remarkably robust. Our second test for the role of our discretization method reverts to the original, raw, scores generated by Diction. Higher scores denote more optimistic communications, such that we would expect stock returns to increase correspondingly. This is indeed what we find, consistently with our earlier results: both FSRs and speeches exert some effects, with those of FSRs being substantially larger than those of speeches. With this measure, we are of course not able to separate out optimistic and pessimistic communications, such that we are neither able to conduct the non-parametric test, nor to fill the tables where we break down the results by the content of the communication.

The final point we address here is the question through which channel communication affects financial markets. Is it that communication affects markets because it contains relevant information, and thus coordinates markets and functions as a focal point - akin to what is known as a coordination channel (e.g. Fratzscher, 2008; Sarno and Taylor, 2001)? Or is it that market participants believe that financial stability communication has a bearing on monetary policy decisions by central banks - or what is referred to as a signalling channel? The evidence discussed so far, in particular the persistence of the effects of communication, strongly points towards the coordination channel being at work (see Sarno and Taylor, 2001). Yet a more direct test of these two channels is to ask whether financial market participants perceive that financial stability communication by central banks could be followed by monetary policy decisions, which should imply that market interest rates are reactive to such communications. As can be seen in the bottom panels of Tables 6 and 7, it is clear that there is no systematic reaction of short (3-month) or long (5 to 10 year) interest rates. Thus, this is further evidence suggesting that there is very little role for a signalling channel, but that it is rather the coordination channel that is at work.

To summarize, the findings suggest that the effects of communication are not universal. Market conditions seem to matter, with different effects during the financial crisis. The origin of the communication also is important, with central banks in advanced economies exerting different effects from those in emerging economies. A sequence of speeches and interviews seems to be affecting stock markets less than an isolated communication by the central bank governor. But importantly, speeches and interviews were moving stock returns during the crisis, while they were not in the pre-crisis period. Finally, the evidence here further supports the conclusion that it is mainly a coordination channel that is at work - i.e. that communication provides relevant information about financial stability itself, rather than giving a signal about monetary policy, thereby affecting financial markets.

4.5 Conclusion

This chapter has provided an empirical assessment of the effects of central bank communication about financial stability, a topic that has remained almost entirely unexplored in the literature to date. The chapter has studied the impact of central bank statements on financial markets, arguably one of the most important target groups of this type of communication. In more detail, it has constructed a unique dataset covering over 1000 communication events (a third of which being FSRs, and two thirds being speeches and interviews by central bank governors) by 37 central banks over a time period from 1996 to 2009, i.e. spanning nearly one and a half decades, and analyzed the reaction of financial sector stocks to these events. The emphasis of the chapter has been to identify whether financial stability-related communication "creates news" and/or "reduces noise".

The chapter's findings suggest that communication about financial stability has important repercussions on financial sector stock prices. However, there are clear differences between FSRs on the one hand and speeches and interviews on the other. FSRs clearly create news in the sense that the views expressed in FSRs get reflected in stock market returns. These effects are furthermore long-lasting. They also reduce noise, as market volatility tends to decline in response to FSRs. These effects are particularly strong if FSRs contain optimistic assessments of the risks to financial stability. Speeches and interviews, in contrast, do on average move financial markets far less. In particular, while having only modest effects on stock market returns, they do not reduce market volatility. However, speeches and interviews were affecting market returns significantly more during the 2007-10 global financial crisis, indicating the potential importance of this communication tool during periods of financial stress.

The mechanism by which the central bank affects financial markets seems to be related to the notion of a co-ordination channel, whereby communication by the central bank works as a co-ordination device, thereby reducing heterogeneity in expectations and information, and thus inducing asset prices to more closely reflect the underlying fundamentals (Sarno and Taylor, 2001). This conclusion is based on the finding that statements have longer-lasting effects, which seems to imply that they have the potential to change the dynamics in financial markets, and based on the result that central bank communication about financial stability does not affect market interest rates in a systematic fashion.

The chapter has also demonstrated how flexibly speeches and interviews can be used as a communication tool, with a higher frequency in times of heightened financial market volatility. In contrast to FSRs with their pre-defined release schedules, the mere occurrence of a speech or an interview can constitute news to financial markets in itself, a fundamental difference that might explain why the two communication channels have so different effects on market volatility. The findings of the chapter therefore underline that communication by monetary authorities on financial stability issues can influence financial market developments, but that it needs to be employed with utmost care, stressing the difficulty of designing a successful communication strategy on financial stability.

Appendix to Chapter 4

Examples of speeches/interviews and their coding

05 March 1996: "Brazil Central Bk President Denies Bank Sector Instability"
"Central bank President Gustavo Loyola Tuesday denied rumors of instability in Brazil's banking sector and said increasing bank investigations and encouragement for bank mergers have quelched any possibility of a crisis [...]" Source: Dow Jones International News
Coded: Optimism =1

27 October 1997: "China c.banker sees more small bank bankruptcies.."
"Some smaller Chinese banks and credit cooperatives could sink into bankruptcy due to bad loans, although a banking crisis was unlikely, central bank governor Dai Xianglong has said." Source: Reuters News
Coded: Optimism =-1

28 January 1998: "U.K. BOE's George Confident Asia Contagion Can Be Avoided"
"Governor of the Bank of England Eddie George said Wednesday he was 'reasonably confident' wider financial contagion from the Asia crisis could be avoided." Source: Dow Jones International News
Coded: Optimism =1

09 November 2000: "Korea markets unstable as worries linger-c.bank."
"South Korea's financial markets continue to show signs of instability as the second phase of financial restructuring progresses, the governor of the central Bank of Korea said on Thursday." Source: Reuters News
Coded: Optimism =-1

19 September 2002: "Mboweni Confident of Financial Stability."
"SA's financial regulators are highly optimistic about the stability of the country's financial system, Tito Mboweni, the SA Reserve Bank governor, said yesterday [. . .]" Source: All Africa
Coded: Optimism =1

10 April 2003: "Fukui says should consider preemptive move on banks."
"Bank of Japan Governor Toshihiko Fukui said on Thursday that Japan should consider ways to provide ailing banks with capital as a preemptive measure before any financial crisis occurred." Source: Reuters News
Coded: Optimism =0

24 September 2003: "Argentina's Central Bank Downplays Big Bank Restructuring"
"Plans to restructure the Argentine financial sector in the wake of last year's financial crisis do not entail a widespread shakeup of the country's banks, top Argentine Central Bank officials said Tuesday." Source: Dow Jones International News
Coded: Optimism =0

17 March 2004: "Greenspan says U.S. banking system healthy."
"Federal Reserve Chairman Alan Greenspan said on Wednesday the U.S. banking system weathered the 2001 recession well, and was in good shape to help finance the economic recovery." Source: Reuters News
Coded: Optimism =1

11 September 2007: "CREDIT WRAPUP 5-Trichet sure major banks sound, Bernanke silent"
"Europe's banks are sound despite the confidence blow from a U.S. subprime crisis, said the head of the European Central Bank on Tuesday, while the [. . .]" Source: Dow Jones International News
Coded: Optimism =1

05 February 2008: "ECB's Noyer: Global Fincl System In Crisis For More Than A Year"
"The global financial system has been in a crisis situation for over a year, and the crisis isn't over, Bank of France Governor Christian Noyer said Tuesday." Source: Dow Jones International News
Coded: Optimism =-1

24 September 2008: "Swedish c.bank head repeats financial system stable"
"Swedish Riksbank Governor Stefan Ingves said on Wednesday Sweden was now feeling the effects of the recent market turmoil more strongly, but repeated reassurances that the financial system was stable." Source: Reuters News
Coded: Optimism =1

03 October 2008: "Bernanke: Fed to do all it can to combat crisis"
"Federal Reserve Chairman Ben Bernanke said on Friday the U.S. central bank will do whatever it can to combat the credit crisis and help the economy." Source: Reuters News
Coded: Optimism =0

06 October 2008: "Turkish banks face narrower credit channels-c.bank"
"Central Bank Governor Durmus Yilmaz said on Monday Turkish banks were facing narrower credit channels due to the global credit crisis, but said they faced no difficulty in renewing external loans." Source: Reuters News
Coded: Optimism =0

Table 4.1: Summary statistics for FSRs and speeches

		FSRs	Speeches & Interviews
By country	Argentina	12	13
	Australia	11	25
	Austria	17	11
	Belgium	7	3
	Brazil	14	9
	Canada	14	22
	Chile	11	15
	China	5	28
	Czech Republic	5	11
	Denmark	11	2
	Euro Area	10	48
	Finland	23	12
	France	13	31
	Germany	5	58
	Greece	1	26
	Hong Kong	12	44
	Hungary	17	17
	Indonesia		6
	Ireland	4	2
	Israel	6	7
	Japan	8	32
	Netherlands	8	17
	New Zealand	10	18
	Norway	20	3
	Philippines		50
	Poland	10	13
	Portugal	5	8
	Singapore	7	1
	South Africa	11	20
	South Korea	9	14
	Spain	14	10
	Sri Lanka	3	2
	Sweden	24	18
	Switzerland	7	16
	Turkey	8	22
	United Kingdom	25	23
	United States		111
By year	1996	1	14
	1997	3	39
	1998	5	118
	1999	7	56
	2000	10	37
	2001	14	17
	2002	18	33
	2003	25	32
	2004	40	26
	2005	53	17
	2006	51	17
	2007	54	68
	2008	51	179
	2009	35	115
Overall		367	768

Notes: The table shows the number of FSRs and speeches that are contained in the database, by country and by year.

Table 4.2: Stock market conditions and the occurrence of communications

	Financial Stability Reports						Speeches & Interviews					
	Returns			**Standard deviation**			**Returns**			**Standard deviation**		
# days	Bench-mark	Event days	Diff	Bench-mark	Event days	Diff	Bench-mark	Event days	Diff	Bench-mark	Event days	Diff
1	0.046	-0.113		--	--	--	0.046	-0.135	**	--	--	--
2	0.110	-0.069		--	--	--	0.110	-0.272	***	--	--	--
3	0.173	-0.087		--	--	--	0.173	-0.522	***	--	--	--
4	0.191	-0.028		4.823	4.211	**	0.191	-0.628	***	4.823	7.820	***
5	0.161	-0.008		4.905	4.361	**	0.161	-0.797	***	4.905	7.848	***
10	0.354	-0.260	*	5.125	4.676	*	0.354	-1.400	***	5.125	7.783	***
15	0.800	-0.148	**	5.308	4.867	**	0.800	-1.476	***	5.308	7.780	***
20	0.949	0.187		5.309	4.981		0.949	-2.235	***	5.309	7.766	***
25	1.313	0.538		5.369	5.023	*	1.313	-2.458	***	5.369	7.808	***
30	1.394	1.272		5.484	5.111	*	1.394	-2.957	***	5.484	7.811	***
35	1.742	1.821		5.503	5.217		1.742	-2.967	***	5.503	7.826	***
40	2.071	2.231		5.474	5.291		2.071	-2.972	***	5.474	7.803	***
45	2.510	2.329		5.682	5.359		2.510	-3.224	***	5.682	7.759	***
50	2.761	2.732		5.547	5.383		2.761	-3.482	***	5.547	7.757	***
55	2.426	2.854		5.682	5.394		2.426	-3.801	***	5.682	7.733	***
60	3.073	3.409		5.600	5.408		3.073	-3.704	***	5.600	7.744	***

Notes: The table shows cumulated stock market returns and the standard deviation of daily stock market returns preceding the communication events (in columns "Event days") and for non-event days (in columns "Benchmark"). Results of mean comparison tests are given in the columns denoted by "Diff". The different rows of the table relate to different time windows prior to the event, starting from a time window of 1 business day to a time window of 60 business days. Standard deviations are only calculated for time windows exceeding 3 business days. The non-event comparison figures are calculated for a sample where no communication event has occurred in the preceding 60 business days, and no communication event follows in the subsequent 60 business days. The sample is furthermore restricted to non-overlapping observations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.3: Stock market conditions and the content of communications

	Financial Stability Reports												Speeches & Interviews											
	Returns						**Standard deviation**						**Returns**						**Standard deviation**					
# days	(1) '-1'	(2) '0'	(3) '1'	(1) vs (2)	(1) vs (3)	(2) vs (3)	(1) '-1'	(2) '0'	(3) '1'	(1) vs (2)	(1) vs (3)	(2) vs (3)	(1) '-1'	(2) '0'	(3) '1'	(1) vs (2)	(1) vs (3)	(2) vs (3)	(1) '-1'	(2) '0'	(3) '1'	(1) vs (2)	(1) vs (3)	(2) vs (3)
1	-0.343	-0.208	0.186		**	**	--	--	--	--	--	--	-0.112	-0.078	-0.212				--	--	--	--	--	--
2	-0.225	-0.222	0.215				--	--	--	--	--	--	-0.248	-0.134	-0.428				--	--	--	--	--	--
3	-0.186	-0.301	0.203				--	--	--	--	--	--	-0.443	-0.287	-0.821				--	--	--	--	--	--
4	-0.157	-0.095	0.151				4.975	3.902	3.806	**	**		-0.612	-0.473	-0.794				8.305	7.714	7.475			
5	-0.022	-0.143	0.132				5.145	4.005	3.985	**	**		-0.703	-0.541	-1.130				8.268	7.783	7.524			
10	-0.948	-0.038	0.156				5.493	4.232	4.352	**	**		-1.448	-0.988	-1.751				7.753	7.956	7.643			
15	-0.765	-0.408	0.655				5.809	4.284	4.559	***	**		-1.245	-1.059	-2.091				7.774	8.035	7.539			
20	-0.744	0.007	1.199		*		5.916	4.413	4.666	***	**		-2.240	-1.645	-2.799				7.773	8.092	7.445			
25	-0.449	0.377	1.584				6.014	4.393	4.714	***	**		-1.759	-2.428	-3.131				7.783	8.123	7.529			
30	0.246	1.264	2.210				6.157	4.479	4.755	***	***		-2.613	-2.451	-3.760				7.755	8.145	7.540			
35	1.349	1.808	2.261				6.356	4.554	4.805	***	***		-2.364	-2.837	-3.648				7.755	8.171	7.559			
40	1.235	2.099	3.258				6.478	4.603	4.861	***	***		-2.136	-2.504	-4.194				7.720	8.161	7.534			
45	1.239	2.142	3.493				6.572	4.665	4.910	***	***		-2.425	-2.791	-4.376				7.700	8.106	7.478			
50	0.928	2.914	4.195				6.602	4.723	4.896	***	***		-2.797	-3.081	-4.502				7.736	8.072	7.471			
55	1.248	2.933	4.237				6.643	4.721	4.892	***	***		-3.197	-3.427	-4.723				7.682	8.064	7.461			
60	2.137	2.897	5.043				6.654	4.751	4.896	***	***		-2.755	-3.438	-4.842				7.701	8.062	7.476			

Notes: The table shows cumulated stock market returns and the standard deviation of daily stock market returns preceding pessimistic (columns (1)), neutral (columns (2)) and optimistic communications (columns (3)). Results of tests for statistically significant differences are given in the columns (1) vs (2), (1) vs (3) and (2) vs (3). The different rows of the table relate to different time windows prior to the event, starting from a time window of 1 business day to a time window of 60 business days. Standard deviations are only calculated for time windows exceeding 3 business days. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.4: Effects of FSRs

	Joint model				Pessimistic FSRs				Neutral FSRs				Optimistic FSRs			
	Returns		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**	
# days	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric
1	0.54	0.27 ***	-- --	-- --	0.44 '	-0.33	-- --	-- --	0.49	-0.09	-- --	-- --	0.53	0.20 *	-- --	-- --
2	0.54	0.33 **	-- --	-- --	0.46	-0.54	-- --	-- --	0.52	0.10	-- --	-- --	0.55	0.14	-- --	-- --
3	0.58 **	0.46 ***	-- --	-- --	0.40 "	-0.75	-- --	-- --	0.53	0.18	-- --	-- --	0.55	0.20	-- --	-- --
4	0.57 **	0.54 ***	0.51	-0.08 *	0.39 "'	-0.73	0.50	-0.10	0.51	0.12	0.52	-0.13 **	0.54	0.37 **	0.51	-0.02
5	0.53	0.44 **	0.53	-0.07 *	0.47	-0.49	0.51	-0.05	0.55	0.21	0.55	-0.11 *	0.54	0.39 *	0.53	-0.05
10	0.53	0.63 **	0.55 **	-0.08 **	0.50	-0.40	0.48	-0.04	0.48	-0.38	0.55	-0.08	0.56 *	0.84 ***	0.61 ***	-0.11 **
15	0.57 **	0.64 **	0.52	-0.06 *	0.44 '	-0.29	0.51	-0.08	0.51	0.02	0.50	-0.02	0.58 **	0.95 ***	0.55 *	-0.08 **
20	0.56 **	0.92 **	0.55 **	-0.05 *	0.50	-0.01	0.56 *	-0.07	0.61 **	0.51	0.55	-0.02	0.61 ***	1.75 ***	0.55	-0.06 **
25	0.57 **	1.27 ***	0.56 ***	-0.07 **	0.48	-0.28	0.60 **	-0.13 **	0.59 **	0.69	0.55	-0.04	0.62 ***	2.18 ***	0.54	-0.05
30	0.58 ***	1.39 ***	0.56 ***	-0.05 *	0.49	-0.36	0.57 **	-0.11 *	0.59 **	0.90 *	0.57 *	-0.02	0.65 ***	2.33 ***	0.55	-0.03
35	0.57 **	1.27 ***	0.56 ***	-0.05 *	0.51	0.11	0.56 *	-0.10 *	0.62 ***	1.50 **	0.55	-0.04	0.64 ***	2.53 ***	0.58 **	-0.01
40	0.53	1.21 **	0.55 **	-0.04	0.54	0.33	0.52	-0.07	0.61 **	1.47 **	0.57 *	-0.03	0.59 **	2.63 ***	0.56 *	-0.01
45	0.56 **	1.41 ***	0.55 **	-0.05 *	0.52	0.17	0.56 *	-0.12 **	0.57 *	1.45 **	0.58 **	-0.04	0.63 ***	2.86 ***	0.52	-0.01
50	0.56 **	1.60 ***	0.56 ***	-0.06 **	0.51	-0.11	0.56 *	-0.12 **	0.59 **	1.46 *	0.58 **	-0.05	0.63 ***	2.97 ***	0.55	-0.01
55	0.56 **	1.47 **	0.56 **	-0.05 *	0.55	0.31	0.58 **	-0.11 **	0.61 **	2.12 **	0.57 *	-0.04	0.66 ***	3.09 ***	0.52	0.00
60	0.55 *	1.21 **	0.55 **	-0.05 *	0.54	0.62	0.58 **	-0.11 **	0.61 **	2.66 ***	0.55	-0.05	0.63 ***	2.87 ***	0.52	0.00

Notes: Notes: The table shows results of the test for communication effects. The first set of results (Returns, non-parametric) tests the share of cases in which $\sum_{k=0}^{K} \widehat{\varepsilon}_{it+k} > 0$ if $I_{it}^{optimism,FSR} = 1$ or $\sum_{k=0}^{K} \widehat{\varepsilon}_{it+k} < 0$ if $I_{it}^{optimism,FSR} = -1$, for different time windows K in the rows of the table. The second column (Returns, parametric) shows the average size of the cumulated excess returns $\frac{1}{N} \sum_{n=1}^{N} \sum_{k=0}^{K} I_{nt}^{optimism,FSR} \widehat{\varepsilon}_{nt+k}$ and tests whether these are different from zero. The columns for "standard deviation" show the share of cases in which the standard deviation of excess returns over K days after the release of an FSR is smaller than the standard deviation during the K days prior to the release, i.e. $\sigma_{\hat{\varepsilon}_{i,t/t+K}} < \sigma_{\hat{\varepsilon}_{i,t-1/t-1-K}}$ if $D_{it}^{c} = 1$ (non-parametric), and their average difference (parametric), and tests these against 0.5 and 0, respectively. The second to fourth panel of the table repeats the exercise for FSRs that have been coded as $I_{it}^{optimism,c} = -1$, $I_{it}^{optimism,c} = 0$ and $I_{it}^{optimism,c} = 1$, respectively. Standard deviations are only calculated for time windows exceeding 3 business days. ***, **, and * indicate statistical significance against the null hypothesis at the 1%, 5%, and 10% levels, respectively. "', ", and ' indicate statistical significance against the alternative hypothesis at the 1%, 5%, and 10% levels, respectively.

Table 4.5: Effects of speeches and interviews

	Joint model				Pessimistic speeches and interviews				Neutral speeches and interviews				Optimistic speeches and interviews			
	Returns		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**	
# days	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric
1	0.45 "	-0.09	-- --	-- --	0.54	0.12	-- --	-- --	0.48	-0.06	-- --	-- --	0.45 "	-0.05	-- --	-- --
2	0.48	-0.10	-- --	-- --	0.57 **	0.38 **	-- --	-- --	0.44 "	-0.27	-- --	-- --	0.52	0.17	-- --	-- --
3	0.49	-0.10	-- --	-- --	0.52	0.28	-- --	-- --	0.46	-0.59	-- --	-- --	0.50	0.07	-- --	-- --
4	0.51	0.11	0.47 '	0.02	0.51	0.07	0.50	-0.18 *	0.45 '	-0.49	0.46	0.08	0.53	0.28	0.46	0.15 '
5	0.53 *	0.26	0.48	0.01	0.49	-0.02	0.53	-0.19 **	0.45 "	-0.47	0.45 '	0.10	0.55 *	0.48 *	0.46	0.11
10	0.55 ***	0.55 **	0.49	0.00	0.46	0.05	0.51	-0.04	0.48	-0.29	0.49	0.04	0.57 **	1.11 ***	0.46	0.01
15	0.54 *	0.74 **	0.48	0.04	0.47	0.06	0.49	0.05	0.49	-0.49	0.50	0.04	0.54	1.47 ***	0.45 '	0.03
20	0.52	0.73 **	0.49	0.06 '	0.50	0.17	0.46	0.06	0.45 '	-0.70	0.53	0.05	0.54	1.54 ***	0.47	0.07
25	0.55 **	1.04 **	0.50	0.06 "	0.47	0.04	0.48	0.04	0.45 '	-0.42	0.56 **	0.02	0.56 **	2.02 ***	0.48	0.12 "
30	0.54 **	1.04 **	0.51	0.06 "	0.50	0.63	0.50	0.04	0.46	-0.39	0.53	0.02	0.57 ***	2.55 ***	0.49	0.12 "
35	0.56 ***	1.06 **	0.50	0.06 '	0.48	0.70	0.51	0.04	0.49	-0.29	0.52	0.01	0.60 ***	2.67 ***	0.49	0.12 "
40	0.54 *	1.01 *	0.50	0.05 '	0.51	0.93	0.52	0.03	0.49	0.07	0.51	0.01	0.57 ***	2.78 ***	0.47	0.11 "
45	0.52	0.95 *	0.50	0.05 '	0.52	1.13	0.53	0.01	0.50	0.05	0.51	0.01	0.56 **	2.83 ***	0.47	0.12 "
50	0.55 ***	1.24 **	0.51	0.04 '	0.49	1.06	0.54 *	0.00	0.50	0.20	0.51	0.02	0.59 ***	3.33 ***	0.48	0.11 "
55	0.55 **	1.58 **	0.52	0.04 '	0.50	0.58	0.53	0.01	0.50	0.06	0.53	0.01	0.59 ***	3.55 ***	0.50	0.10 "
60	0.55 **	1.63 **	0.50	0.03	0.50	0.38	0.51	0.00	0.49	0.25	0.52	0.00	0.60 ***	3.45 ***	0.48	0.09 '

Notes: See notes to Table 4, but all results relate to speeches and interviews rather than FSRs.

Table 4.6: Effects of FSRs – sample splits and robustness

	Joint model				**Pessimistic FSRs**				**Neutral FSRs**				**Optimistic FSRs**			
	Returns		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**	
	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric
A - Benchmark	0.57 **	1.27 ***	0.56 ***	-0.07 **	0.48	-0.28	0.60 **	-0.13 **	0.59 **	0.69	0.55	-0.04	0.62 ***	2.18 ***	0.54	-0.05
B - Sample splits																
1. Country Group																
Advanced economies	0.56 *	0.91 **	0.59 ***	-0.11 ***	0.46	-0.31	0.62 ***	-0.20 ***	0.57 *	1.10 **	0.59 **	-0.09 **	0.58 *	1.62 ***	0.55	-0.03
Emerging economies	0.62 **	2.27 **	0.48	0.05	0.55	-0.14	0.50	0.20	0.64	-0.83	0.40	0.17	0.69 ***	3.21 ***	0.51	-0.07
2. Crisis versus pre-crisis																
Pre-crisis	0.63 ***	2.10 ***	0.55 *	-0.05 **	0.39 "	-1.06	0.61 **	-0.08 *	0.61 **	0.77 *	0.55	-0.01	0.64 ***	2.73 ***	0.51	-0.05 *
Financial crisis 2007-2010	0.45	-0.55	0.60 **	-0.13 *	0.58	0.65	0.58 *	-0.20	0.52	0.47	0.58	-0.11	0.52	-0.32	0.65 **	-0.01
3. Supervisory role																
CB is supervisor	0.56	1.47 **	0.55 *	-0.09 *	0.61	0.96	0.39	-0.15	0.63 *	0.50	0.56	-0.05	0.64 **	2.63 ***	0.63 **	-0.10 *
CB is not supervisor	0.58 **	1.17 **	0.57 **	-0.06 *	0.44	-0.67	0.66 ***	-0.13 *	0.57	0.79	0.55	-0.04	0.59 *	1.80 ***	0.46	0.00
C - Robustness																
All stocks	0.58 ***	1.16 ***	0.55 **	-0.02	0.50	-0.31	0.58 **	-0.05	0.54	0.00	0.48	0.04	0.66 ***	1.96 ***	0.60 ***	-0.06 **
Alternative coding	0.53	0.72 *	0.56 ***	-0.07 **	0.52	0.43	0.52	0.01	0.57 *	0.40	0.63 ***	-0.12 **	0.59 **	1.86 ***	0.54	-0.10 **
Raw Diction scores	--	0.50 ***	0.56 ***	-0.07 **	--	--	--	--	--	--	--	--	--	--	--	--
D - Testing for the signalling channel																
Short-term interest rates	0.54 *	0.05	0.55 **	0.01	0.50	-0.06	0.59 **	0.00	0.58 *	0.12 *	0.51	0.01 '''	0.58 **	0.04	0.55	0.00
Long-term interest rates	0.53	0.02	0.52	0.00	0.45	-0.03	0.51	0.00	0.57 *	0.03	0.50	0.00	0.52	0.00	0.56 *	0.00

Notes: See notes to Table 4. All results relate to the effect of FSRs at a time window of 25 days. Row 1 reports the benchmark results, each subsequent row reports results of a specific sample split or robustness test. Sample splits for advanced/emerging economies, pre-crisis/financial crisis, CB as supervisor or not. Robustness tests relate to using overall stock indices rather than financial sector stocks indices, as well as to using an alternative coding of the content of the communications, or using the raw Diction optimism scores directly, rather than their discretized versions. The last panel shows the effects on short- and long-term interest rates.

Table 4.7: Effects of speeches and interviews – sample splits and robustness

	Joint model				**Pessimistic speeches and interviews**				**Neutral speeches and interviews**				**Optimistic speeches and interviews**			
	Returns		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**		**Returns**		**Standard deviation**	
	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric	non-parametric	parametric
A - Benchmark	0.55 **	1.04 **	0.50	0.06 "	0.47	0.04	0.48	0.04	0.45 '	-0.42	0.56 **	0.02	0.56 **	2.02 ***	0.48	0.12 "
B - Sample splits																
1. Country Group																
Advanced economies	0.54 *	1.02 **	0.48	0.07 "	0.50	0.56	0.47	0.04	0.45 '	-0.34	0.52	0.02	0.58 **	2.53 ***	0.46	0.15 "
Emerging economies	0.57 *	1.10	0.56 **	0.04	0.36 "	-1.74	0.50	0.03	0.47	-0.63	0.65 ***	0.03	0.52	0.60	0.52	0.06
2. Crisis versus pre-crisis																
Pre-crisis	0.52	0.48	0.51	0.02	0.48	-0.10	0.47	0.06	0.48	-0.88	0.55 *	0.01	0.52	0.84	0.52	0.01
Financial crisis 2007-2010	0.59 ***	1.87 **	0.49	0.11 "	0.45	0.29	0.48	0.02	0.43 "	0.10	0.56 *	0.03	0.62 ***	3.58 ***	0.43 '	0.27 "
3. Supervisory role																
CB is supervisor	0.54	0.75	0.50	0.06	0.50	0.61	0.49	-0.03	0.48	0.11	0.50	0.08	0.56 *	1.81 **	0.50	0.12
CB is not supervisor	0.56 **	1.37 **	0.51	0.06	0.44 '	-0.49	0.47	0.10 '	0.43 '	-0.97	0.62 ***	-0.04	0.57 *	2.30 ***	0.45	0.12
4. Clustering																
Speeches as part of cluster	0.51	-0.04	0.35 "'	0.43 "'	0.47	-0.20	0.24 "'	0.50 "'	0.56	1.84	0.37 '	0.36 "'	0.50	-0.22	0.41	0.45 "
Speeches outside cluster	0.56 **	1.25 ***	0.53 **	-0.01	0.47	0.08	0.52	-0.03	0.43 "	-0.86	0.59 ***	-0.05	0.58 **	2.50 ***	0.49	0.05
C - Robustness																
All stocks	0.51	0.87 ***	0.48	0.06 "	0.51	-0.22	0.48	0.05	0.48	-0.28	0.53	0.01	0.52	1.46 ***	0.44 "	0.13 "'
Alternative coding	0.53	0.75 *	0.50	0.06 "	0.48	0.37	0.47	0.09 '	0.48	-0.51	0.53	0.01	0.53	1.84 ***	0.51	0.08
Raw Diction scores	--	0.13 **	0.50	0.06 "	--	--	--	--	--	--	--	--	--	--	--	--
D - Testing for the signalling channel																
Short-term interest rates	0.49	-0.07	0.53 *	-0.03 *	0.44 "	-0.01	0.53	-0.02	0.41 "'	-0.17	0.54 *	-0.02	0.42 "'	-0.15	0.52	-0.04 *
Long-term interest rates	0.49	-0.06	0.50	0.00	0.48	-0.02	0.50	-0.01 **	0.42 "'	-0.04	0.53	0.00	0.47	-0.12	0.48	0.01

Notes: See notes to Table 8, but all results relate to speeches and interviews rather than FSRs. The table also contains test results for speeches and interviews that are part of a cluster or not.

Bibliography

Abel, Andrew B. (1983). “Optimal Investment under Uncertainty”. *American Economic Review 73* (1), 228–33.

Abel, Andrew B., Avinash K. Dixit, Janice C. Eberly, and Robert S. Pindyck (1996). “Options, the Value of Capital, and Investment”. *Quarterly Journal of Economics 111* (3), 753–777.

Alexopoulos, Michelle and Jon Cohen (2009). “Uncertain Times, Uncertain Measures”. Working Papers. University of Toronto, Department of Economics.

Andreasen, Martin (2010). “How to Maximize the Likelihood Function for a DSGE Model”. *Computational Economics 35* (2), 127–154.

Anselin, Luc (1988a). “Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity”. *Geographical Analysis 20*, 1–17.

– (1988b). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.

– (2007). “Spatial Econometrics in RSUE: Retrospect and Prospect”. *Regional Science and Urban Economics 37*, 450–456.

– (2010). “Thirty Years of Spatial Econometrics”. *Papers in Regional Science 89* (1), 3–25.

Armesto, Michelle T., Rubén Hernández-Murillo, Michael T. Owyang, and Jeremy Piger (2009). “Measuring the Information Content of the Beige Book: A Mixed Data Sampling Approach”. *Journal of Money, Credit and Banking 41* (1), 35–55.

Arraiz, Irani, David M. Drukker, Harry H. Kelejian, and Ingmar R. Prucha (2010). "A Spatial Cliff-Ord-type Model with Heteroskedastic Innovations: Small and Large Sample Results". *Journal of Regional Science 50*, 592–614.

Arulampalam, M. Sanjeev, Simon Maskell, Neil Gordon, and Tim Clapp (2002). "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking". *IEEE Transactions on Signal Processing 50*, 174–188.

Auerbach, Alan J. (1989). "Tax Reform and Adjustment Costs: The Impact on Investment and Market Value". *International Economic Review 30* (4), 939–962.

Bachmann, Rüdiger and Christian Bayer (2011). "Uncertainty Business Cycles - Really?" NBER Working Papers 16862.

Bachmann, Rüdiger, Steffen Elstner, and Eric R. Sims (2010). "Uncertainty and Economic Activity: Evidence from Business Survey Data". NBER Working Papers 16143.

Baltagi, Badi H. (2008). *Econometric Analysis of Panel Data*. 4th. Chichester: Wiley.

Baltagi, Badi H. and Dong Li (2001). "Double Length Artificial Regressions for Testing Spatial Dependence". *Econometric Reviews 20*, 31–40.

Baltagi, Badi H. and Qi Li (1991). "A Joint Test for Serial Correlation and Random Individual Effects". *Statistics & Probability Letters 11* (3), 277–280.

– (1995). "Testing AR(1) Against MA(1) Disturbances in an Error Component Model". *Journal of Econometrics 68* (1), 133–151.

Baltagi, Badi H. and Ping X. Wu (1999). "Unequally Spaced Panel Data Regressions With AR(1) Disturbances". *Econometric Theory 15* (06), 814–823.

Beaudry, Paul and Bernd Lucke (2010). "Letting Different Views about Business Cycles Compete". *NBER Macroeconomics Annual 2009, Volume 24*, 413–455.

Beaudry, Paul and Franck Portier (2006). "Stock Prices, News, and Economic Fluctuations". *American Economic Review 96* (4), 1293–1307.

Bekaert, Geert, Marie Hoerova, and Marco Lo Duca (2010). "Risk, Uncertainty and Monetary Policy". NBER Working Papers 16397.

Bernanke, Ben S. (1983). "Irreversibility, Uncertainty, and Cyclical Investment". *Quarterly Journal of Economics 98* (1), 85–106.

Bernanke, Ben S. and Kenneth N. Kuttner (2005). "What Explains the Stock Market's Reaction to Federal Reserve Policy?" *Journal of Finance 60* (3), 1221–1257.

Bhargava, Alok, Luisa Franzini, and Wiji Narendranathan (1982). "Serial Correlation and the Fixed Effects Model". *Review of Economic Studies 49* (4), 533–49.

Binsbergen, Jules van, Jesús Fernández-Villaverde, Ralph S.J. Koijen, and Juan F. Rubio-Ramírez (2010). "The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences". NBER Working Papers 15890.

Bligh, Michelle C. and Gregory D. Hess (2007). "The Power of Leading Subtly: Alan Greenspan, Rhetorical Leadership, and Monetary Policy". *The Leadership Quarterly 18* (2), 87–104.

Blinder, Alan S., Michael Ehrmann, Marcel Fratzscher, Jakob De Haan, and David-Jan Jansen (2008). "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence". *Journal of Economic Literature 46* (4), 910–45.

Bloom, Nicholas (2009). "The Impact of Uncertainty Shocks". *Econometrica 77* (3), 623–685.

Bloom, Nicholas, Max Floetotto, and Nir Jaimovich (2010). "Really Uncertain Business Cycles". Mimeo, Stanford University.

Boehner, John (2010). "Republicans Will Oppose Democrats' Job-Killing Tax Hike on Small Businesses". *Press release.*

Born, Benjamin and Jörg Breitung (2011a). "Simple Regression Based Tests for Spatial Dependence". *Econometrics Journal 14* (2), 330–342.

– (2011b). "Testing for Serial Correlation in Fixed-Effects Panel Data Models". Mimeo, University of Bonn.

Born, Benjamin and Johannes Pfeifer (2011). "Policy Risk and the Business Cycle". Bonn Econ Discussion Papers No. 06/2011.

Born, Benjamin, Michael Ehrmann, and Marcel Fratzscher (2010). "Macroprudential Policy and Central Bank Communication". CEPR Discussion Papers No. 8094.

– (2011a). "Central bank Communication on Financial Stability". ECB Working Paper Series No. 1332.

Born, Benjamin, Michael Ehrmann, and Marcel Fratzscher (2011b). “How Should Central Banks Deal with a Financial Stability Objective? The Evolving Role of Communication as a Policy Instrument”. *Handbook of Central Banking, Financial Regulation and Supervision After the Financial Crisis*. Ed. by Sylvester C.W. Eijffinger and Donato Masciandaro. Cheltenham: Edward Elgar, 245–268.

Breusch, Trevor S. and Adrian R. Pagan (1979). “A Simple Test for Heteroscedasticity and Random Coefficient Variation”. *Econometrica 47* (5), 1287–1294.

Burridge, Peter (1980). “On the Cliff-Ord Test for Spatial Autocorrelation”. *Journal of the Royal Statistical Society B 42*, 107–108.

Calvo, Guillermo A. (1983). “Staggered Prices in a Utility-Maximizing Framework”. *Journal of Monetary Economics 12* (3), 383–398.

Cantor, Eric (2010). “Beware of the Obama Tax Increases”. *USA Today*. Sept. 2, 2010.

Carhart, Mark M (1997). “On Persistence in Mutual Fund Performance”. *Journal of Finance 52* (1), 57–82.

Carroll, Christopher D. and Miles S. Kimball (2008). “Precautionary Saving and Precautionary Wealth”. *The New Palgrave Dictionary of Economics*. Ed. by Steven N. Durlauf and Lawrence E. Blume. Basingstoke: Palgrave Macmillan.

Chang, Roberto and Andrés Fernández (2010). “On the Sources of Aggregate Fluctuations in Emerging Economies”. NBER Working Papers 15938.

Chib, Siddharta and Srikanth Ramamurthy (2010). “Tailored Randomized Block MCMC Methods with Application to DSGE Models”. *Journal of Econometrics 155* (1), 19–38.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005). “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy”. *Journal of Political Economy 113* (1), 1–45.

Christiano, Lawrence J., Mathias Trabandt, and Karl Walentin (2007). “Introducing Financial Frictions and Unemployment into a Small Open Economy Model”. Sveriges Riksbank Working Paper Series No. 214.

Chugh, Sanjay K. (2011). “Firm Risk and Leverage-Based Business Cycles”. Mimeo, University of Maryland.

Cihak, Martin (2006). “How Do Central Banks Write on Financial Stability?” IMF Working Papers 06/163.

– (2007). “Central Banks and Financial Stability: A Survey”. Mimeo.

Clarida, Richard, Jordi Galí, and Mark Gertler (2000). “Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory”. *Quarterly Journal of Economics 115* (1), 147–180.

Cliff, Andrew D. and J. Keith Ord (1972). “Testing for Spatial Autocorrelation Among Regression Residuals”. *Geographical Analysis 4*, 267–284.

– (1981). *Spatial Processes: Models and Applications*. London: Pion.

Cochrane, John H. (2005). *Asset Pricing*. Revised Edition. Princeton, NJ: Princeton University Press.

Cox, David R. and Patty J. Solomon (1988). “On Testing for Serial Correlation in Large Numbers of Small Samples”. *Biometrika 75* (1), pp. 145–148.

Cukierman, Alex (2009). “The Limits of Transparency”. *Economic Notes 38* (1-2), 1–37.

Davidson, Russel and James G. MacKinnon (1984). “Model Specification Tests Based on Artificial Linear Regressions”. *International Economic Review 25*, 485–502.

– (1988). “Double-Length Artificial Regressions”. *Oxford Bulletin of Economics and Statistics 50*, 203–217.

Davidson, Russell and Emmanuel Flachaire (2008). “The Wild Bootstrap, Tamed at Last”. *Journal of Econometrics 146*, 162–169.

Davidson, Russell and James G. MacKinnon (2004). *Econometric Theory and Methods*. New York: Oxford University Press.

Davis, Angela K., Jeremy M. Piger, and Lisa M. Sedor (2006). “Beyond the Numbers: An Analysis of Optimistic and Pessimistic Language in Earnings Press Releases”. Working Papers 2006-005. Federal Reserve Bank of St. Louis.

Dincer, Nergiz and Barry Eichengreen (2009). “Central Bank Transparency: Causes, Consequences and Updates”. NBER Working Papers 14791.

Doucet, Arnaud and Adam M. Johansen (2009). “A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later”. *Handbook of Nonlinear Filtering.*

Drukker, David M. (2003). "Testing for Serial Correlation in Linear Panel-Data Models". *Stata Journal 3* (2), 168–177.

Duffie, Darrell and Kenneth J. Singleton (1993). "Simulated Moments Estimation of Markov Models of Asset Prices". *Econometrica 61* (4), 929–952.

Eberly, Janice, Sergio Rebelo, and Nicolas Vincent (2008). "Investment and Value: A Neoclassical Benchmark". NBER Working Papers 13866.

Edmans, Alex, Diego García, and Øyvind Norli (2007). "Sports Sentiment and Stock Returns". *Journal of Finance 62* (4), 1967–1998.

Ehrmann, Michael and Marcel Fratzscher (2007a). "Communication by Central Bank Committee Members: Different Strategies, Same Effectiveness?" *Journal of Money, Credit and Banking 39* (2-3), 509–541.

– (2007b). "The Timing of Central Bank Communication". *European Journal of Political Economy 23* (1), 124–145.

– (2009). "Purdah-On the Rationale for Central Bank Silence around Policy Meetings". *Journal of Money, Credit and Banking 41* (2-3), 517–528.

Eicker, Friedhelm (1963). "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions". *Ann. Math. Statist. 34*, 447–456.

– (1967). "Limit Theorems for Regressions with Unequal and Dependent Errors". *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability.* Ed. by L. LeCam and J. Neyman. Berkeley: University of California Press, 59–82.

Eijffinger, Sylvester C.W. and Petra M. Geraats (2006). "How Transparent Are Central Banks?" *European Journal of Political Economy 22* (1), 1–21.

Engle, Robert F. (1982). "A General Approach to Lagrange Multiplier Model Diagnostics". *Journal of Econometrics 20*, 83–104.

European Central Bank (2010). Financial Stability Review December 2010.

Fama, Eugene F. and Kenneth R. French (1993). "Common Risk Factors in the Returns on Stocks and Bonds". *Journal of Financial Economics 33* (1), 3–56.

Fernández-Villaverde, Jesús and Juan F. Rubio-Ramírez (2007). "Estimating Macroeconomic Models: A Likelihood Approach". *Review of Economic Studies 74* (4), 1059–1087.

– (2010). "Macroeconomics and Volatility: Data, Models, and Estimation". Mimeo, University of Pennsylvania.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, and Juan F. Rubio-Ramírez (2010). "Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in U.S. Data". NBER Working Papers 15928.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Keith Kuester, and Juan F. Rubio-Ramírez (2011). "Fiscal Volatility Shocks and Economic Activity". Unpublished Manuscript.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Juan F. Rubio-Ramírez, and Martín Uribe (forthcoming). "Risk Matters: The Real Effects of Volatility Shocks". *American Economic Review*.

Fields, Gary (2009). "Political Uncertainty Puts Freeze on Small Businesses". *The Wall Street Journal*.

Fisher, Jonas D.M. (2006). "The Dynamic Effects of Neutral and Investment Specific Technology Shocks". *Journal of Political Economy 114* (3), 413–451.

Fratzscher, Marcel (2008). "Oral Interventions Versus Actual Interventions in Fx Markets - An Event-Study Approach". *Economic Journal 118* (530), 1079–1106.

Garcia, Carlos, Jorge Restrepo, and Evan Tanner (2007). "Designing Fiscal Rules for Commodity Exporters". ILADES-Georgetown University Working Papers inv199.

Garcia-Cicco, Javier, Roberto Pancrazi, and Martín Uribe (2010). "Real Business Cycles in Emerging Countries?" *American Economic Review 100* (5), 2510–31.

Geweke, John (1992). "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments". *Bayesian Statistics*. Ed. by José. M Bernardo, James O. Berger, A.Philip Dawid, and Adrian F. M. Smith. Vol. 4. Oxford: Clarendon Press, 641–649.

Geweke, John, Joel Horowitz, and Hashem Pesaran (2008). "Econometrics". *The New Palgrave Dictionary of Economics*. Ed. by Steven N. Durlauf and Lawrence E. Blume. Basingstoke: Palgrave Macmillan.

Godsill, Simon J., Arnaud Doucet, and Mike West (2004). "Monte Carlo Smoothing for Nonlinear Time Series". *Journal of the American Statistical Association 99* (465), 156–168.

Gosselin, Pierre, Aileen Lotz, and Charles Wyplosz (2009). "Interest Rate Signals and Central Bank Transparency". *NBER International Seminar on Macroeconomics 2007*, 9–51.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman (1988). "Investment, Capacity Utilization, and the Real Business Cycle". *American Economic Review 78* (3), 402–17.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997). "Long-Run Implications of Investment-Specific Technological Change". *American Economic Review 87* (3), 342–362.

– (2000). "The Role of Investment-Specific Technological Change in the Business Cycle". *European Economic Review 44* (1), 91–115.

Hansen, Nikolaus, Sybille D. Müller, and Petros Koumoutsakos (2003). "Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES)". *Evolutionary Computation 11* (1), 1–18.

Hart, Roderick P. (2000). *Campaign Talk: Why Elections Are Good For Us*. Princeton, NJ: Princeton University Press.

Hart, Roderick P. and Sharon E. Jarvis (1997). "Political Debate: Forms, Styles, and Media". *American Behavioral Scientist 40*, 1095–1122.

Hartman, Richard (1972). "The Effects of Price and Cost Uncertainty on Investment". *Journal of Economic Theory 5* (2), 258 –266.

Hayashi, Fumio (1982). "Tobin's Marginal q and Average q: A Neoclassical Interpretation". *Econometrica 50* (1), 213–24.

Hsiao, Cheng (2003). *Analysis of Panel Data*. 2nd ed. Cambridge, UK: Cambridge University Press.

Imrohoroglu, Ayse (2010). "Experts Weigh in: Can the Economy Be Saved? - Uncertainty is Killing Confidence". *Los Angeles Times*.

Ingersoll, Jonathan E. and Stephen A. Ross (1992). "Waiting to Invest: Investment and Uncertainty". *Journal of Business 65* (1), 1–29.

Jaimovich, Nir and Sergio Rebelo (2009). "Can News about the Future Drive the Business Cycle?" *American Economic Review 99* (4), 1097–1118.

Jarque, Carlos M. and Anil K. Bera (1987). "A Test for Normality of Observations and Regression Residuals". *International Statistical Review / Revue Internationale de Statistique 55* (2), pp. 163–172.

Jones, John Bailey (2002). "Has Fiscal Policy Helped Stabilize the Postwar U.S. Economy?" *Journal of Monetary Economics 49* (4), 709–746.

Judd, Kenneth L. (1998). *Numerical Methods in Economics*. Cambridge, MA: MIT Press.

Justiniano, Alejandro and Giorgio E. Primiceri (2008). "The Time-Varying Volatility of Macroeconomic Fluctuations". *American Economic Review 98* (3), 604–41.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti (2010). "Investment Shocks and Business Cycles". *Journal of Monetary Economics 57* (2), 132 –145.

Kaminsky, Graciela L. and Karen K. Lewis (1996). "Does Foreign Exchange Intervention Signal Future Monetary Policy?" *Journal of Monetary Economics 37* (2-3), 285–312.

Kelejian, Harry H. and Ingmar R. Prucha (2001). "On the Asymptotic Distribution of the Moran I Test Statistic with Applications". *Journal of Econometrics 104*, 219–257.

Kelejian, Harry H. and Dennis P. Robinson (1998). "A Suggested Test for Spatial Autocorrelation and/or Heteroskedasticity and Corresponding Monte Carlo Results". *Regional Science and Urban Economics 28*, 389–417.

– (2004). "The Influence of Spatially Correlated Heteroskedasticity on Tests for Spatial Correlation". *Advances in Spatial Econometrics*. Ed. by Luc Anselin, Raymond J. G. M. Florax, and Sergiu J. Rey. Berlin: Springer, 26–97.

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A. Sims (2008). "Calculating and Using Second Order Accurate Solutions of Discrete Time Dynamic

Equilibrium Models". *Journal of Economic Dynamics and Control 32* (11), 3397 –3414.

Kim, Sangjoon, Neil Shephard, and Siddhartha Chib (1998). "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models". *Review of Economic Studies 65* (3), 361–93.

Kimball, Miles S. (1990). "Precautionary Saving in the Small and in the Large". *Econometrica 58* (1), 53–73.

King, M. (1981). "A Small Sample Property of the Cliff-Ord Test for Spatial Correlation". *Journal of the Royal Statistical Society B 43*, 263–264.

King, Robert G., Charles I. Plosser, and Sergio T. Rebelo (1988). "Production, Growth and Business Cycles: I. The Basic Neoclassical Model". *Journal of Monetary Economics 21* (2-3), 195–232.

Klein, Ezra (2010). "Where Was the Policy Uncertainty in 1993?" *Washington Post*.

Koenker, Roger (1981). "A Note on Studentizing a Test for Heteroscedasticity". *Journal of Econometrics 17* (1), 107–112.

Kolmogorov, Andrey N. (1933). "Sulla Determinazione Empirica di una Legge di Distribuzione". *Giornale dell'Istituto Italiano degli Attuari 4* (1933), 83–91.

Kothari, S.P. and J.B. Warner (2007). "Econometrics of Event Studies". *Handbook of Corporate Finance: Empirical Corporate Finance Volume 1*. Ed. by B.E. Eckbo. Amsterdam: Elsevier, 3–36.

Krämer, Walter (2003). "The Robustness of the F-test to Spatial Autocorrelation Among Regression Disturbances". *Statistica 3*, 435–440.

– (2005). "Finite Sample Power of Cliff-Ord-Type Tests for Spatial Disturbance Correlation in Linear Regression". *Journal of Statistical Planning and Inference 128*, 489 –496.

Krämer, Walter and C. Donninger (1987). "Spatial Autocorrelation Among Errors and the Relative Efficiency of OLS in the Linear Regression Model". *Journal of the American Statistical Association 82*, 577–579.

Lee, Lung-fei (2007). "GMM and 2SLS estimation of Mixed Regressive, Spatial Autoregressive Models". *Journal of Econometrics 137*, 489–514.

Leeper, Eric M., Michael Plante, and Nora Traum (2010). "Dynamics of Fiscal Financing in the United States". *Journal of Econometrics 156* (2), 304–321.

Leland, Hayne E. (1968). "Saving and Uncertainty: The Precautionary Demand for Saving". *Quarterly Journal of Economics 82* (3), pp. 465–473.

Liu, Regina Y. (1988). "Bootstrap Procedures Under Some Non-I.I.D. Models". *Annals of Statistics 16*, 1696–1708.

Lowrie, Annie (2010). "The Uncertainty Mandate - Think Congressional Republicans Can Reduce Economic Uncertainty? Don't Be so Sure." *Slate*.

Lubik, Thomas A. and Frank Schorfheide (2007). "Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation". *Journal of Monetary Economics 54* (4), 1069–1087.

Lucca, David O. and Francesco Trebbi (2009). "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements". NBER Working Papers 15367.

MacKinlay, A. Craig (1997). "Event Studies in Economics and Finance". *Journal of Economic Literature 35* (1), 13–39.

Martellosio, Federico (2010). "Power Properties Of Invariant Tests For Spatial Autocorrelation In Linear Regression". *Econometric Theory 26*, 152–186.

Masciandaro, Donato and Marc Quintyn (2009). "After the Big Bang and Before the Next One? Reforming the Financial Supervision Architecture and the Role of the Central Bank. A Review of Worldwide Trends, Causes and Effects (1998-2008)". Paolo Baffi Centre Research Paper 2009-37.

McGrattan, Ellen R. (1994). "The Macroeconomic Effects of Distortionary Taxation". *Journal of Monetary Economics 33* (3), 573–601.

McKinnon, John D. (2010). "High-Income Households Spend Less Due To Tax Uncertainty". *The Wall Street Journal*.

Mendoza, Enrique G., Assaf Razin, and Linda L. Tesar (1994). "Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption". *Journal of Monetary Economics 34* (3), 297–323.

Mertens, Karel and Morten O. Ravn (2011). "Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks". *Review of Economic Dynamics 14* (1), 27–54.

Mishkin, Frederic S (2004). "Can Central Bank Transparency Go Too Far?" *The Future of Inflation Targeting*. Ed. by Christopher Kent and Simon Guttmann. Sydney: Reserve Bank of Australia.

Monacelli, Tommaso and Roberto Perotti (2008). "Fiscal Policy, Wealth Effects, and Markups". NBER Working Papers 14584.

Moran, P.A.P. (1948). "The Interpretation of Statistical Maps". *Biometrika 35*, 255–260.

Morris, Stephen and Hyun Song Shin (2002). "Social Value of Public Information". *American Economic Review 92* (5), 1521–1534.

Oosterloo, Sander and Jakob de Haan (2004). "Central Banks and Financial Stability: A Survey". *Journal of Financial Stability 1* (2), 257–273.

Oosterloo, Sander, Jakob de Haan, and Richard Jong-A-Pin (2007). "Financial Stability Reviews: A First Empirical Analysis". *Journal of Financial Stability 2* (4), 337–355.

Pojarliev, Momtchil and Richard M. Levich (2007). "Do Professional Currency Managers Beat the Benchmark?" NBER Working Papers 13714.

Popescu, Adina and Frank Smets (2010). "Uncertainty, Risk-taking, and the Business Cycle in Germany". *CESifo Economic Studies 56* (4), 596–626.

Primiceri, Giorgio E. (2005). "Time Varying Structural Vector Autoregressions and Monetary Policy". *Review of Economic Studies 72* (3), 821–852.

Reeve, Elspeth (2010). "Is Uncertainty About Obama's Policies to Blame for the Lame Economy?" *The Atlantic Wire*.

Rigobon, Roberto and Brian Sack (2004). "The Impact of Monetary Policy on Asset Prices". *Journal of Monetary Economics 51* (8), 1553–1575.

Romer, Christina D. and David H. Romer (2010). "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks". *American Economic Review 100* (3), 763–801.

Rosenblatt, Murray (1952). "Remarks on a Multivariate Transformation". *Annals of Mathematical Statistics 23* (3), 470–472.

Rudebusch, Glenn D. (1995). "Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure". *Journal of Monetary Economics 35* (2), 245 –274.

Ruge-Murcia, Francisco J. (2010). "Estimating Nonlinear DSGE Models by the Simulated Method of Moments". CIREQ Working Paper Series 19-2010.

Sarno, Lucio and Mark P. Taylor (2001). "Official Intervention in the Foreign Exchange Market: Is It Effective and, If So, How Does It Work?" *Journal of Economic Literature 39* (3), 839–868.

Schmitt-Grohé, Stephanie and Martín Uribe (2004). "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function". *Journal of Economic Dynamics and Control 28* (4), 755–775.

– (2006). "Optimal Fiscal and Monetary Policy in a Medium-Scale Macroeconomic Model". *NBER Macroeconomics Annual 2005, Volume 20*, 383–462.

– (2010). "What's News in Business Cycles". Mimeo, Columbia University.

– (2011). "Business Cycles With A Common Trend in Neutral and Investment-Specific Productivity". *Review of Economic Dynamics 14* (1c), 122–135.

Searle, S.R. (1971). *Linear Models*. New York: Wiley.

Shapiro, Samuel S. and Martin B. Wilk (1965). "An Analysis of Variance Test for Normality (Complete Samples)". *Biometrika 52* (3/4), pp. 591–611.

Shephard, Neil (2008). "Stochastic Volatility Models". *The New Palgrave Dictionary of Economics*. Ed. by Steven N. Durlauf and Lawrence E. Blume. Basingstoke: Palgrave Macmillan.

Smets, Frank and Rafael Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". *American Economic Review 97* (3), 586–606.

Smirnov, Nickolay (1948). "Table for Estimating the Goodness of Fit of Empirical Distributions". *Annals of Mathematical Statistics 19* (2), pp. 279–281.

Stock, James H. and Mark W. Watson (1999). "Forecasting Inflation". *Journal of Monetary Economics 44* (2), 293–335.

Svensson, Lars E. O. (2003). “Monetary Policy and Real Stabilization”. NBER Working Papers 9486.

– (2006). “Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con”. *American Economic Review 96* (1), 448–452.

White, Halbert (1980). “A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity”. *Econometrica 48*, 817–838.

Wingfield, Brian (2010). “Continuing The Bush Tax Cuts May Create More Uncertainty”. *Forbes*.

Wooldridge, Jeffrey M. (1990). “A Unified Approach to Robust, Regression-Based Specification Tests”. *Econometric Theory 6* (01), 17–43.

– (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wu, Guiying (2009). “Uncertainty, Investment and Capital Accumulation: A Structural Econometric Approach”. PhD thesis. Oxford University,

Yang, Shu-Chun Susan (2005). “Quantifying Tax Effects Under Policy Foresight”. *Journal of Monetary Economics 52* (8), 1557–1568.

Yun, Tack (1996). “Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles”. *Journal of Monetary Economics 37* (2), 345–370.