



Essays in Applied Econometrics and Finance

Reinhard Ellwanger

Thesis submitted for assessment with a view to obtaining the degree
of Doctor of Economics of the European University Institute

Florence, 9 July 2015

European University Institute
Department of Economics

Essays in Applied Econometrics and Finance

Reinhard Ellwanger

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Peter Hansen, Supervisor, EUI

Prof. Juan Dolado, EUI

Prof. Christian Brownlees, Universitat Pompeu Fabra

Dr. Christiane Baumeister, University of Notre Dame

© Reinhard Ellwanger, 2015

No part of this thesis may be copied, reproduced or transmitted without prior permission of the author

Abstract

This thesis comprises three essays. The first two chapters address topics in commodity markets and their interaction with derivative and other asset markets. The third essay deals with the effects to and from fiscal policy that arise due to the structure of the relationship between central and regional governments. Finance and applied econometrics constitute the common thread for these articles. The first two take a financial economics and financial econometrics perspective, while the third essay addresses a topic of public finance with an empirical approach.

The first chapter offers an explanation for volatile oil prices. Using information from options and futures I document economically large jump tail premia in the crude oil market which can be related to investors' "fear". These premia vary substantially over time and significantly forecast crude oil futures and spot returns. The results suggest that oil futures prices overshoot (undershoot) in the presence of upside (downside) tail fears in order to allow for smaller (larger) risk premia thereafter.

The second essay relates the comovement of stock and commodity prices to increased participation of financial investors in commodity future markets. I present a partial equilibrium model in which demand for futures by financial investors transmits stock market shocks into commodity prices via a time varying risk premium. Empirically, I find that commodity index investors react systematically to stock market shocks by adjusting their commodity risk exposure.

In the third chapter, joint with Abián García Rodríguez, we investigate the relationship between fiscal decentralization - the share of government spending and taxation carried out at the subnational level - and fiscal policy effects. Using a cross-section of countries, we document a positive relationship between decentralization and the effectiveness of fiscal policy as measured by the size of fiscal multipliers. We also present a case study for the decentralization process in Spain and find that it had a positive impact on output growth.

Acknowledgements

I first would like to thank my supervisor, Peter Reinhard Hansen, for his support and guidance during these years, and particularly for never failing to encourage me to pursue my path. I am also grateful to Fabio Canova, Juan Dolado and Evi Pappa who have been very supportive and always available for academic and personal inquiries. My gratitude goes to Christiane Baumeister and Christian Brownlees as well, for accepting to be part of my committee and for their valuable comments which have improved (and will) the essays the thesis is comprised of. I am also very thankful to all the faculty members of the Economics Department at the EUI, and to the staff members of Villa San Paolo, especially Jessica, Loredanna and Sonia, for supporting my journey during the last years. This thesis would have not been possible without all the comments and discussions during our lunch meetings and presentations in the “Econometrics and beyond working group”.

The last five years would have not been the same without the valuable support, discussions and company of my friends and colleagues, Abián, who is also my co-author, Ilya, Kirill, Marek, Romanos, Srecko, Stefan, Vasja, Wojciech and all the marvelous people at the EUI who have not only helped me with academic discussions but have also made my everyday life enjoyable in the past years. I am confident there will be opportunities to continue sharing experiences with all of them in the future. A special thank goes to Clodomiro, my friend since the first moment I arrived to Florence and flatmate during most of my stay here; and to Lian, for his advice, mock presentations and mock interviews for the job market.

I am deeply indebted to my parents for their unconditional love and support. Thanks to my sister Regina and my brother Gottfried with whom I could share many adventures. I always felt reassured knowing they were there for me despite the distance. Above all I would like to thank Gaby for her support, encouragement, quiet patience and the marvelous experience of sharing her life and love with me every day.

Contents

Abstract	i
Acknowledgements	ii
1 Driven by Fear? The Tail Risk Premium in the Crude Oil Futures Market	1
1.1 Introduction	1
1.2 Theoretical Setup	4
1.2.1 Setup and Definitions	4
1.2.2 Risk Premia and Spot Prices	7
1.3 Empirical Implementation	9
1.3.1 Estimation of the Variation Measures	9
1.3.2 Data Description	12
1.3.3 Empirical Tail Risk Measures	13
1.4 Predictability of Futures and Spot Returns	16
1.4.1 Predictive Regression Framework and Control Variables	16
1.4.2 Forecasting Results	18
1.4.3 Risk or Fear: Where does the predictability come from?	20
1.4.4 VAR Estimates on the Impact of the Oil Fear	21
1.5 Interaction with the Macroeconomy	23
1.5.1 Oil and Aggregate Uncertainty	23
1.5.2 Oil Uncertainty and Real Activity	24
1.6 Conclusion	25
1.A Appendix: Computation of the Jump Measures	26
1.B Appendix: Data and Empirical Measures	28

1.C	Appendix: Predictability of Futures and Spot Returns	30
1.D	Appendix: Interaction with the Macroeconomy	31
2	The Role of Commodity Index Investment in Commodity and Asset Price Comovement	37
2.1	Introduction	37
2.2	Related Literature	40
2.3	A Model of Commodity Futures Market with Financial Investment	43
2.4	The Empirical Effect of Stock Market Movements on CITs' Positions	48
2.5	Conclusion	53
2.A	Appendix: Proof of Proposition 1 - 3	54
2.B	Appendix: Figures	55
2.C	Appendix: Tables	57
3	Regional Decentralization and Fiscal Policy Effects - International and Intranational Evidence	64
3.1	Introduction	64
3.2	International Evidence	66
3.2.1	Measuring Decentralization	66
3.2.2	Econometric Framework and Data Description	70
3.2.3	Identification of Structural Innovations	71
3.2.4	Results	73
3.2.5	Robustness	76
3.3	Intranational Evidence from Spain	78
3.3.1	The Decentralization Process in Spain (1975-2007)	78
3.3.2	Episodes of Decentralization	80
3.3.3	Justification of the Narrative Approach	82
3.3.4	Empirical Framework	84
3.3.5	Results	84
3.3.6	Discussion	87
3.4	Conclusion	90

3.A	Appendix: Decentralization Measures and Cross-country Comparison	91
3.B	Appendix: Intranational Evidence	98

Chapter 1

Driven by Fear? The Tail Risk Premium in the Crude Oil Futures Market

1.1 Introduction

Volatile oil spot and futures prices have drawn a lot of attention from academics, policy makers and investors over the last years. The origins of this volatility are hotly debated because the changes in oil demand and supply appear too smooth to explain the large swings in oil prices (see e.g. Tang and Xiong 2012; Alquist and Kilian 2010; Baumeister and Peersman 2013). One explanation proposed by Kilian (2009) is that precautionary demand shocks are an important driver of short run fluctuations in oil prices. He suggests that a key source of precautionary demand movements is uncertainty about shortfalls of expected supply relative to expected demand. A natural question is therefore how uncertainty shocks can be accurately measured and whether they help to predict the future evolution of oil prices. Yet another is if and how they relate to the risk premia embedded in oil derivatives. The explanations building on competitive storage models such as the one proposed by Alquist and Kilian (2010) are typically based on the assumption of risk neutrality and ignore potential effects arising from temporal variations in these premia. This stands in contrast to the convincing evidence of significant and time-varying risk premia in oil markets from the commodity finance literature (see e.g. Gorton and Rouwenhorst 2004; Erb and Harvey 2006; Singleton 2013; Hamilton and Wu 2014; Szymanowska et al. 2014). We bridge this gap by presenting a novel uncertainty measure that is explicitly related to the risk premium embedded in oil market derivatives.

The contribution of this paper is threefold. First, we present novel estimates of oil

market uncertainty based on the left jump and right jump variation premia ($LJVP$ and $RJVP$, respectively) embedded in crude oil futures and options. Bollerslev and Todorov (2011b) and Bollerslev, Todorov, and Xu (2014) show that these premia - defined as the part of the variance risk premium that is due to large sized upward and downward jumps - can be estimated in an essentially model free manner and contain important information about market participants' sentiments and expected stock market returns. We demonstrate that for the oil market, $LJVP$ and $RJVP$ are economically large and significantly vary over episodes of documented supply and demand uncertainty. The variation measures have strong predictive power for both crude oil futures and spot prices that is not contained in traditional oil price predictors. Our results suggest that oil futures and spot prices overshoot (undershoot) in the presence of upside (downside) tail fears in order to allow for smaller (larger) risk premia thereafter.

Second, we use a stylized no-arbitrage model of storage in order to show that relative to the futures price movement, this overshooting (undershooting) is amplified for the spot price due to the time varying value of holding physical oil in inventory. A relative increase in $RJVP$ is also associated with a rise in this value that pushes the spot price in the same direction as the futures risk premium. Consistently, $RJVP$ and $LJVP$ exhibit larger in-sample and out-of-sample predictability for spot price returns than for futures returns. This finding complements the model of Alquist and Kilian (2010), in which a similar no-arbitrage condition in the storage market is used to show that an increase in oil production volatility leads to an overshooting of spot prices. Their model is derived under the assumption of risk neutrality and the overshooting is due to convex adjustment costs of inventories. Since in our model the fear shock is associated with an additional movement of the risk premium this overshooting is magnified. Under risk aversion, the spot price not only increases with respect to the spot price in some future period, but also with respect to the rise implied by storage models that assume risk neutrality. Taken together, these results represent new evidence for the importance of time varying risk premia in explaining the perceived excess volatility in oil spot and futures prices.

Finally, to put our jump risk measures in perspective, we also investigate the link between $LJVP$, $RJVP$ and real economic activity as well as macroeconomic uncertainty. The oil fear measures appear not to be spanned by aggregate uncertainty. Also, there is little evidence of a stable linear relationship between stock market returns, $LJVP$ and $RJVP$. The results are consistent with the idea that the oil tail risk measures aggregate different types of uncertainty that are relevant for oil prices, e.g. oil supply and oil demand uncertainty, that individually have a very distinct relationship with aggregate uncertainty.

There are several reasons for focusing on the tails of the oil price distribution to quantify uncertainty. First, recent theoretical works show that models with tail risks can account for

the high equity risk premium and excess market return volatility (Barro 2006; Wachter 2013). There is also increasing evidence that the index option implied compensation for aggregate market volatility and tail risks are closely connected to economic uncertainty and temporal variation in risk aversion (Bollerslev, Tauchen, and Zhou 2009; Bollerslev, Todorov, and Xu 2014). In contrast, most of the commodity finance literature has focused on futures, ignoring potential information from the related option prices. Among the exceptions are Trolle and Schwartz (2010), who document a significant and time-varying variance risk premium in the crude oil option market and Kang and Pan (2013) who show that this premium forecasts short term futures returns. We find that most, if not all of the variance risk premium and its forecasting power is due to time varying compensation for tail variations. One of the advantage of considering tail risk premia is that they are naturally separated into upside and downside uncertainty, thus providing additional information beyond that contained in the variance risk premium. We show that in particular the time varying asymmetry of the jump premia across the two tails improves the prediction spot and futures returns. Based on the variance risk decomposition proposed by Bollerslev and Todorov (2011b), we present evidence that this asymmetry reflects changes in effective risk aversion of oil market participants. On a more general level, these findings suggest that time varying disaster fears embedded in option prices of individual assets, not only on market indexes, convey important information on the their return dynamics.

Formally, $RJVP$ and $LJVP$ are defined as the difference of the conditional expected variation of jump tails under the statistical, objective probability measure and the risk neutral measure. In other words, they represent the difference of the expected actual jump variations and the option implied market price for the insurance against these jump variations. Our implemented model for the option implied jump-measures follows Bollerslev, Todorov, and Xu (2014). It is semi-parametric and flexible, allowing the tail distributions to differ across the left and the right tail and for independent time variation in the shape and level of the tails. The empirical estimation is based on panel of out-of-money (OTM) call and put options. Intuitively, short-maturity OTM options are most sensitive to large jumps, which allows us to separate the jump risk from the diffusive risk. The statistical jump variation is based on intraday futures data and non-parametric methods developed by Bollerslev and Todorov (2011a). Empirically, we find that the statistical jump variations in oil futures are significantly smaller than their risk neutral counterparts. Moreover, the actual jumps are approximately symmetric, while the option implied prices for these jumps display time-varying asymmetries. Thus the observed time variation in the relative size of $RJVP$ and $LJVP$ are largely void of influences from the actual jumps, and can be interpreted as a direct measure of investor fears (Bollerslev and Todorov 2011b).

On average, the $LJVP$ tends to be larger than $RJVP$, implying that oil investors are

on average requesting a higher premium when downside-risks predominate. This is largely consistent with the theory of normal backwardation (Keynes 1976). Accordingly, when producers of the physical commodity want to hedge their price risk using derivatives, then arbitrageurs who take the other side of the contract have to be compensated in form of a risk premium order to take on the risk. Our empirical estimates imply that on average about half futures risk premium is due to large jump risks.

Last, we contribute to a strand of literature going back to the idea of Bernanke (1983) that oil market uncertainty - rather than oil price changes alone - is a key variable to understand the relationship between the oil market and macroeconomic outcomes. Our measures provide precise definitions and estimates for oil price uncertainty, and we discuss some interactions with aggregate variables towards the end of the paper.

The rest of this paper is structured as follows. The next section provides a formal definition of the tail risk variation measures, and explains the spot price overshooting. In section 3 we discuss the empirical implementation and data as well as the properties of our estimates. Section 4 presents the forecasting results and section 5 the interaction of the jump tail premia with aggregate outcomes. Section 6 concludes.

1.2 Theoretical Setup

1.2.1 Setup and Definitions

In this section, we present the general setup and provide formal definitions for our tail risk measures. This setup is inspired by Bollerslev and Todorov (2011b) and Bollerslev, Todorov, and Xu (2014). Instead of considering the (aggregate) stock market, we will focus on the dynamics of an individual asset, namely oil futures.

To fix ideas, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a filtered probability space with the filtration $(\mathcal{F}_t)_{t \geq 0}$ and let F_t denote the price of a crude oil futures contract. The dynamics of the futures price are described by the following jump diffusion process

$$\frac{dF_t}{F_{t-}} = \alpha_t dt + \sigma_t dW_t + \int_{\mathbb{R}} (e^x - 1) \tilde{\mu}(dt, dx), \quad (1.1)$$

where the drift α_t and the stochastic volatility σ_t are assumed to be locally bounded càdlàg processes, and W_t is a standard Brownian Motion. Here $\tilde{\mu}(dt, dx) = \mu(dt, dx) - v_t^{\mathbb{P}}(dx)dt$ denotes a compensated jump measure, with $\mu(dt, dx)$ the counting measure and $v_t^{\mathbb{P}}(dx)dt$ the compensator of jumps, where \mathbb{P} denotes the statistical, objective measure.¹

Under standard non-arbitrage assumptions, there exists a risk-neutral measure denoted

¹The compensator $v_t^{\mathbb{P}}(dx)dt$ ensures that the jump measure $\tilde{\mu}(dt, dx)$ is a martingale.

\mathbb{Q} , under which the futures price follows a martingale of the form

$$\frac{dF_t}{F_{t-}} = a_t dt + \sigma_t dW_t^{\mathbb{Q}} + \int_{\mathbb{R}} (e^x - 1) \tilde{\mu}^{\mathbb{Q}}(dt, dx), \quad (1.2)$$

where a_t denotes the drift, $dW_t^{\mathbb{Q}}$ is a Brownian motion with respect to the risk neutral measure and $\tilde{\mu}^{\mathbb{Q}} = \mu(dt, dx) - v_t^{\mathbb{Q}}(dx)dt$ denotes the jump measure under \mathbb{Q} following the previous decomposition. In general, the change of measure alters both the drift and the jump intensity describing the dynamics of the futures price while the volatility associated with the Brownian motions remains the same under both measures. This reflects the special pricing of jumps in comparison with continuous movements.

Our interest will be in both the futures risks premium (FRP) - a premium reflecting risk associated with holding a (long) futures contract - and the variance risk premium (VRP) - a premium reflecting risks associated with holding a (long) variance swap - that are associated with the jump part of the futures price. Following Bollerslev and Todorov (2011b), the FRP at time t and for some $T > t$ is defined as

$$FRP_{t,T} \equiv \frac{1}{T-t} \left(E_t^{\mathbb{P}} \left(\frac{F_T - F_t}{F_t} \right) - E_t^{\mathbb{Q}} \left(\frac{F_T - F_t}{F_t} \right) \right). \quad (1.3)$$

Since the futures price F_t is a martingale under the \mathbb{Q} -measure, $FRP_{t,T}$ is effectively determined by the difference of the objective expectation of the futures price at some future date T and the current futures price.

Given our jump diffusion model in equation (1.1) we can, without loss of generality, define the FRP due to large jumps above some threshold $k_t > 0$,

$$\begin{aligned} FRP_{t,T}(k_t) &\equiv \frac{1}{T-t} E_t^{\mathbb{P}} \left(\int_t^T \int_{|x| > k_t} (e^x - 1) v_s^{\mathbb{P}}(dx) ds \right) - \\ &\quad \frac{1}{T-t} E_t^{\mathbb{Q}} \left(\int_t^T \int_{|x| > k_t} (e^x - 1) v_s^{\mathbb{Q}}(dx) ds \right). \end{aligned} \quad (1.4)$$

Going one step further, we can decompose $FRP_{t,T}(k_t)$ into the contributions from large positive and large negative jumps

$$FRP_{t,T}(k_t) = FRP_{t,T}^+(k_t) + FRP_{t,T}^-(k_t), \quad (1.5)$$

where $FRP_{t,T}^+(k_t)$ captures the futures risk premia due to $x > k_t$ and $FRP_{t,T}^-(k_t)$ captures the premia due to $x < -k_t$.

The variability of the futures price is measured by the quadratic variation QV of its

log-price process of the interval $[t, T]$

$$QV_{[t,T]} = \int_t^T \sigma_s^2 ds + \int_t^T \int_{\mathbb{R}} x^2 \mu(ds, dx). \quad (1.6)$$

Similar to the futures risk premium, $VRP_{t,T}$ is formally defined as the difference of the expected quadratic variation over the $T - t$ period under the respective probability measure.

$$VRP_t = \frac{1}{T-t} (E_t^P(QV_{[t,T]}) - E_t^Q(QV_{[t,T]})). \quad (1.7)$$

Under this definition of the variance risk premium, VRP_t equals the expected payoff from a long variance swap contract (Carr and Wu 2009). The variance risk premium is also naturally decomposed into a part associated with the continuous-time stochastic volatility process σ_s and a part that is due to jumps. We denote $RJV_{t,T}^{\mathbb{P}}(k_t)$ and $LJV_{t,T}^{\mathbb{P}}(k_t)$ the predictable component of the quadratic variation arising through large positive and large negative jumps under the \mathbb{P} measure

$$RJV_{t,T}^{\mathbb{P}}(k_t) = \int_t^T \int_{x > k_t} x^2 v_s^{\mathbb{P}}(dx) ds, \quad LJV_{t,T}^{\mathbb{P}}(k_t) = \int_t^T \int_{x < -k_t} x^2 v_s^{\mathbb{P}}(dx) ds \quad (1.8)$$

and their counterparts under the risk neutral measure \mathbb{Q}

$$RJV_{t,T}^{\mathbb{Q}}(k_t) = \int_t^T \int_{x > k_t} x^2 v_s^{\mathbb{Q}}(dx) ds, \quad LJV_{t,T}^{\mathbb{Q}}(k_t) = \int_t^T \int_{x < -k_t} x^2 v_s^{\mathbb{Q}}(dx) ds. \quad (1.9)$$

The part of the variance risk premium due to large positive jumps is then

$$RJVP(k_t) \equiv \frac{1}{T-t} (E_t^{\mathbb{P}}(RJV_{t,T}^{\mathbb{P}}(k_t)) - E_t^{\mathbb{Q}}(RJV_{t,T}^{\mathbb{Q}}(k_t))), \quad (1.10)$$

while the part due to large negative jumps is

$$LJVP(k_t) \equiv \frac{1}{T-t} (E_t^{\mathbb{P}}(LJV_{t,T}^{\mathbb{P}}(k_t)) - E_t^{\mathbb{Q}}(LJV_{t,T}^{\mathbb{Q}}(k_t))). \quad (1.11)$$

As suggested by Bollerslev and Todorov (2011b), for equity index option the difference between $LJVP(k_t)$ and $RJVP(k_t)$ is naturally associated with investors' fear. In this paper we investigate this hypothesis for the oil market and define

$$FI_t(k_t) \equiv LJVP(k_t) - RJVP(k_t). \quad (1.12)$$

The index FI_t measures the asymmetry between the premium requested for the downside variance risk and the premium charged for upside variance risk that is due to large jumps.

Under the above definition, a relatively large left jump variation premia or "downside fear" is associated with a low value of $FI_t(k)$.

The respective premia can then be estimated in an essentially model free manner using high-frequency returns and options data. The next section discusses this procedure in more detail. Before, we turn to the relationship between risk premia and spot prices.

1.2.2 Risk Premia and Spot Prices

In this section, we present a theoretical framework based on no-arbitrage conditions that relates the risk premia to oil spot prices. We show that in general, an increase in the futures risk premium associated with a rise in market participants fears - as proxied by the tail risk premia - will drive a temporary wedge between the current and expected prices for both futures and spot prices. Since this wedge subsequently reverts to zero, our model predicts that futures and spot prices overshoot with respect to future prices in the wake of disaster fears.

For our analysis we draw on non-arbitrage conditions derived from two different approaches to commodity derivative pricing that allow to relate the current futures price and the current spot price to the expected spot price some period ahead.² The first approach is based on the basic definition of the futures risk premium as described in equation (1.3). By non-arbitrage, the value of the futures price at the time of maturity must be equal to the spot price of the commodity. Hence $F_{T,T} = S_T$, where S_T stands for the spot price of oil at time T , and where in slight abuse of notation we let T denote the contract's terminal date for the remainder of this subsection. Moreover, since $F_{t,T}$ is a martingale under the \mathbb{Q} -measure, it follows that

$$1 + (T - t)FRP_{t,T} = \frac{E_t(S_T)}{F_{t,T}} = \frac{E_t(F_{t,T})}{F_{t,T}}. \quad (1.13)$$

Equation (1.13) reflects the non-arbitrage condition that the price of a futures has to be equal to the expected spot price discounted by the premium associated with holding the futures contract. This decomposition of the futures price into the expected spot price and a risk premium has been used frequently in the analysis of the futures risk premium. The empirical and theoretical evidence overwhelmingly suggests that in the oil market, and in commodity markets in general, the (net) premium is on average positive and fluctuating over time (e.g. Keynes 1976; Bessembinder 1992; Hamilton and Wu 2014; Baumeister and Kilian 2014).

²The setup is similar to the one in Gospodinov and Ng (2013). In our model, however, we explicitly focus on the effect of the tail premia and their interaction with the futures risk premium and the convenience yield, whereas Gospodinov and Ng (2013) focus exclusively on the latter.

The second non-arbitrage condition is based on the theory of storage, going back to the works of Kaldor (1939) and Working (1949). A distinguishing feature of commodities as an asset class is the significance of the convenience yield, defined as the benefit of immediate availability of a physical commodity rather than a time T contingent claim on the commodity (see e.g. Fama and French 1987). This benefit links the futures to the spot price through the following relationship

$$F_{t,T} = S_t(1 - (T - t)CY_{t,T}), \quad (1.14)$$

where $CY_{t,T}$ is the (net of storage costs and interest rate outlays) equilibrium convenience yield, in annualized terms. This equation has to hold under non-arbitrage, since the price of a futures contract has to be equal to the cost of buying the commodity now minus the net benefits of carrying the commodity to maturity. Such benefits can arise due to a variety reasons such as temporary stock-outs and associated price spikes or convex adjustment costs. Importantly, the convenience yield can also be interpreted as a call option on a futures contract (Milonas and Thomadakis 1997). Given that our fear measured is derived from actual option prices on the futures, we expect a close link between oil market fears and the convenience yield.

From the decomposition of the futures risk premium into contributions from positive and negative large jumps, equations (1.4) and (1.5), it follows that an exogenous, relative shift in the tail risk premium associated with the right tail above that of the left tail will decrease $FRP_{t,T}$. Equation (1.13) then implies that the futures price will rise above the expected spot price, and is expected to decline towards to spot price thereafter. Moreover, the size of the shift in the futures price will be determined by the change in the futures risk premium. In order to investigate the relationship between current and expected spot prices, we combine equation (1.13) and (1.14), obtaining the following relationship between the current and the expected spot price

$$\frac{E_t(S_T)}{S_t} = (1 + (T - t)FRP_{t,T})(1 - (T - t)CY_{t,T}). \quad (1.15)$$

Thus under no-arbitrage, the (relative to the expected spot price) price of current spot oil is determined by the net value of holding the physical oil in storage and the futures risk premium. Moreover, a tail fear induced the change in the spot price is unambiguously larger than the change in the futures price if the net convenience yield varies negatively with the relative size of the tail risk measures. In this case, the spot prices not only overshoot with respect to expected spot prices, but also with respect to the current futures price. The expected spot and futures returns, $r_{(T-t),S}$ and $r_{(T-t),F}$, respectively, that are associated

with a relative increase in the right tail variation measure should therefore exhibit

$$E_t(r_{(T-t),S}|RJVP_t > LJVP_t) < E_t(r_{(T-t),F}|RJVP_t > LJVP_t) < 0, \quad (1.16)$$

with an inverse relationship for relative increases in the left tail measure. Considering a temporary, mean preserving change in the tail risk measures, i.e. a change such that $E_t(S_T|RJVP_t > LJVP_t) = E_t(S_T|RJVP_t = LJVP_t)$, equation (1.16) implies that both the spot and the futures price have to adjust immediately by an upward movement.³ Moreover, the underlying mechanism does not rely on any shifts in inventories. This is important, since some of the large price fluctuations, e.g. around the 1991 Gulf War, are difficult to reconcile with the observed changes in inventory (Kilian and Murphy 2014).

Of course, equation (1.16) rests on the assumption that the net convenience yield varies negatively with FI_t . Subsequently we show that FI_t is inversely related to the market prices of out-of-the-money options. The interpretation of the convenience yield as an option suggests therefore that we should expect this inverse relationship to hold theoretically. Empirically this assumption is justified by our data, as we find a strong and statistically significant correlation of almost -40% between our measure of FI_t and the log of the net convenience yield measure as implied by equation (1.14).

Taken together, the non-arbitrage conditions linking the contemporaneous futures price and spot price with future prices, suggest that oil futures and spot prices overshoot relative to the expected spot price in the wake of upside (downside) fears. This implies that our tail variation measures should forecast futures and spot market returns. Moreover, this overshooting, reflected by the expected return, is larger for the spot price than for the futures price. We address this hypothesis by direct forecasts and a structural VAR analysis after presenting the estimation methodology and results for the tail variation measures.

1.3 Empirical Implementation

1.3.1 Estimation of the Variation Measures

As noted by Carr and Wu (2009), the implied total variation $QV_{t,T}^Q$, also known as the variance swap rate, can be well approximated by a portfolio of out-of-the-money put and

³This comparative static thought experiment resembles the setting of Alquist and Kilian (2010), who describe the effects from a change in the conditional variance of oil supply shocks. In their model, an increase in the conditional variance of these shocks lead to a similar overshooting of the spot price. However, these results are derived under the assumption of risk neutrality so that the futures price is an unbiased predictor of the spot price. In contrast, our framework allows for an interaction of a time-varying risk premia with the convenience yield and suggest that conditional on $RJVP_t$, $LJVP_t$, futures prices are not necessarily an unbiased predictor.

call options. For our calculation we follow the methodology for the CBOE Volatility Index (VIX).⁴

Our specification of the jump tails follows Bollerslev, Todorov, and Xu (2014). In particular, the jump distribution and intensity under the \mathbb{Q} -measure are based on the semi-parametric model

$$v_t^{\mathbb{Q}}(dx) = \left(\phi_t^+ \times e^{-\alpha_t^+ x} 1_{(x>0)} + \phi_t^- \times e^{-\alpha_t^- |x|} 1_{(x<0)} \right) dx. \quad (1.17)$$

Relative to other existing models, this specification imposes only minimal restrictions on the jump tail dynamics since (a) the left and right jump tails are allowed to differ and (b) the level shift parameters ϕ^\pm and the shape parameters α^\pm are allowed to vary independently over time.

The estimation of $\alpha^+(\alpha^-)$ and $\phi^+(\phi^-)$ is based on the observation that for $(T - t) \downarrow 0$ and $k \uparrow \infty (k \downarrow -\infty)$

$$\frac{e^r O_{t,T}(K)}{(T - t)F_{t,T}} \approx \frac{\phi_t^\pm e^{k(1 \pm \alpha_t^\pm)}}{\alpha_t^\pm (\alpha_t^\pm \pm 1)}, \quad (1.18)$$

where $O_{t,T}(K)$ denotes the price of a call (put) option with strike K and $k = \log(K/F_{t,T})$. This reflects the intuition that for close to maturity, deep OTM options the risks associated with the diffusive part become negligible and their price therefore reflects jump risks.

From equation (1.18) it follows that the ratio of two OTM options does not depend on ϕ_t^\pm , leading to the natural estimator suggested by Bollerslev and Todorov (2013):

$$\hat{\alpha}_t^\pm = \operatorname{argmin}_{\alpha^\pm} \frac{1}{N_t^\pm} \sum_{i=1}^{N_t^\pm} \left| \log \left(\frac{O_{t,\tau}(k_{t,i})}{O_{t,\tau}(k_{t,i-1})} \right) (k_{t,i} - k_{t,i-1})^{-1} - (1 \pm (-\alpha^\pm)) \right|, \quad (1.19)$$

where $O_{t,\tau}$ is the time t price of an OTM option on the futures with log-moneyness k , N_t^\pm denotes the total number of options used in the estimation and $0 < |k_{t,1}| < \dots < |k_{t,N_t^\pm}|$. In practice we will pool options such that t refers to a given month which implicitly assumes that α^\pm is approximately constant during this period.

For a given α^\pm , we then use equation (1.18) to estimate

$$\hat{\phi}_t^\pm = \operatorname{argmin}_{\phi^\pm} \frac{1}{N_t^\pm} \sum_{i=1}^{N_t^\pm} \left| \log \left(\frac{e^{r\tau} O_{t,\tau}(k_{t,i})}{(T - t)F_{t-\tau}} \right) + (\pm \hat{\alpha}_t^\pm - 1)k_{t,i} + \log(\hat{\alpha}_t^\pm \mp 1) + \log(\hat{\alpha}_t^\pm) - \log(\phi^\pm) \right|. \quad (1.20)$$

From the definition of the tail risk premia in equation (1.9) and our assumptions for the

⁴See the white paper on the CBOE website for details regarding the VIX methodology.

large jumps dynamics in (1.17), it follows that for time to maturity $T - t$ and threshold k_t

$$RV_{t,T}^{\mathbb{Q}} = (T - t)\phi_t^+ e^{-\alpha_t^+ k_t} (\alpha_t^+ k_t (\alpha_t^+ k_t + 2) + 2) / (\alpha_t^+)^3 \quad \text{and} \quad (1.21)$$

$$LJV_{t,T}^{\mathbb{Q}} = (T - t)\phi_t^- e^{-\alpha_t^- k_t} (\alpha_t^- k_t (\alpha_t^- k_t + 2) + 2) / (\alpha_t^-)^3. \quad (1.22)$$

The \mathbb{Q} tail measures are then computed by replacing the population quantities in (1.21) by their estimates.

The estimation of the corresponding quantities under the objective measure are based on high-frequency intraday data. We use the notation of Bollerslev and Todorov 2011b and divide the trading day t into the $[t, t + \pi_t]$ overnight period and the $[t + \pi_t, t + 1]$ active trading period. Hence π_t denotes the length of the close to open interval.⁵ Dividing the effective trading time in equally spaced intervals, we obtain n returns $\Delta_{t,i}f \equiv f_{t+\pi+i} - f_{t+\pi+i-1}$, where f denotes the logarithm of the futures price. We denote RV_t the realized variation on day t , which is consistently estimated by summing the squared intraday returns

$$RV_t \equiv \sum_{i=1}^n (\Delta_{t,i}f)^2 \xrightarrow{p} \int_{t+\pi_t}^{t+1} \sigma_s^2 ds + \int_{t+\pi_t}^{t+1} \int_{\mathbb{R}} x^2 \mu(ds, dx). \quad (1.23)$$

Realized jumps under the statistical measure are estimated using the threshold technique first proposed by Mancini (2001). Under the threshold estimation, we first compute an estimate for the continuous part of the volatility, σ_t , and then filter out jumps by identifying a threshold separating jumps that appear incompatible with the underlying normal distribution. The truncation threshold for large jumps is time-varying and captures the effects of well-described volatility clustering as well as intraday volatility. Out of the returns that are identified as jumps, we select the large and medium-sized ones for the tail estimation. The reader is referred to Appendix A regarding for further details regarding the estimation procedure.

Due to the lack of observations for sufficiently large jumps under the statistical measure it is infeasible to estimate the same flexible jump tail specification as under the \mathbb{Q} -measure. Instead we follow Bollerslev and Todorov (2011b) in assuming that⁶

$$v_t^{\mathbb{P}} = ((\alpha_0^- 1_{x<0} + \alpha_0^+ 1_{x>0}) + (\alpha_1^- 1_{x<0} + \alpha_1^+ 1_{x>0}) \sigma_t^2) v^{\mathbb{P}}(x) dx, \quad (1.24)$$

where $v^{\mathbb{P}}(x)dx$ is the time-invariant distribution of empirical jumps and the level parameters

⁵Although the trading hours are non-stochastic, it is convenient to treat π_t as stochastic. The theoretical derivations presented below are valid under mild conditions regarding the stochastic process for π_t . See Bollerslev and Todorov 2011b for details.

⁶Empirical evidence that the jump distribution is in the oil market is approximately proportional to the continuous volatility is presented in Doran and Ronn (2008).

α_0^\pm and α_1^\pm relate linearly to the time varying continuous volatility σ_t^2 .⁷

The estimation of the time invariant jump distribution draws on the insight by Bollerslev and Todorov (2011a) that the tails of an arbitrary distribution are approximately distributed according to a Generalized Pareto Distribution. Given our empirical jumps, the two parameters of the Generalized Pareto Distribution along with α_0 and α_1 are estimated separately for each tail via GMM as outlined in Appendix A.

1.3.2 Data Description

Our empirical analysis is based on light sweet crude oil (West Texas Intermediate) futures and options.⁸ Crude oil derivatives are traded in extremely liquid markets and available historical data goes back to the 1980s. Trading of oil futures started in April 1983 for contracts with maturities up to three months, and for options on futures in November 1986. For the estimation of the \mathbb{Q} jump tails we use an option data set obtained from the Chicago Mercantile Exchange (CME group, formerly NYMEX) that contains all historical end-of-the-day settlement prices.⁹ Conveniently these crude oil options are quoted for a variety of strike prices and expiration dates - one for each calendar month of the year - thus ensuring a sufficient number of short maturity deep OTM options for the empirical implementation of our estimator. The derivation of the tail parameters formally relies on a decreasing time to maturity $(T - t) \downarrow 0$. We therefore only take the contract with the shortest time to maturity, whenever the maturity is larger than 9 days. The last trading days of a given option contract can be characterized by pricing abnormalities due to the lack of trading volume, which makes it necessary to discard this data and resort to the first back contract for those days.¹⁰

In order to mitigate potential influences from the diffusive risk, we retain only OTM call (put) options with at log-moneyness more than plus (minus) twice the maturity-normalized Black-Scholes at-the-money implied volatility. We clean the data by discarding all options with a settlement price of less than 3 cents and those violating the monotonicity condition in the strike dimension. The dataset comprises daily data from December 1987 to December 2013. For the estimation of the jump tails we pool all clean deep OTM options for a given

⁷The specification for the jump tails under the \mathbb{P} -measure is significantly more restrictive than the corresponding \mathbb{Q} -measure specification. Empirically the \mathbb{P} jump tails are dwarfed by the risk neutral analogues and we will therefore only work with the more general \mathbb{Q} -measures in the later part of this paper. The empirical evidence for a significant difference in the expected jump tail variations under the different measures are presented in the following section.

⁸The CME tickers for the futures and options is CL and LO, respectively.

⁹NYMEX crude oil options are American style. For short maturity, deep OTM options the difference between European and American options is negligible, so we use the original options for the jump tail estimation. For the computation of the expected quadratic variation under the \mathbb{Q} -measure we convert the option prices into corresponding European style values following Barone-Adesi and Whaley (1987).

¹⁰This “cleaning procedure” is standard, see e.g. Trolle and Schwartz (2010) or Bollerslev and Todorov (2011b).

calendar month. This leaves us with 313 monthly sets of pooled option data with maturities 9-40 days.¹¹

Our futures data comprises 5-minute intraday price quotes from Tickdatamarket. We retain the future contracts corresponding to the option contracts used for the jump tail estimation under the risk neutral measures in order to ensure an exact matching. After standard data cleaning procedures our sample comprises 6,327 trading days.¹² For the computation of the realized measures we use the part of the day where trading was actively carried out throughout the sample. Thus the first price observation is taken to be 10:00 (CST), and the last price observation 14:30. This leaves us with 54 price observations for each trading day.

1.3.3 Empirical Tail Risk Measures

The estimates for the tail variation measures under \mathbb{Q} implied by equation (1.21) require a choice of the threshold k_t that separates large from small jumps. Similar to Bollerslev, Todorov, and Xu (2014) we allow k_t to vary as a linear function of the implied volatility. This form of time variation mimics the estimation procedure for the statistical jumps and accounts for the idea that what is classified as an “extreme” event can differ with economic conditions and corresponding market volatility. The following results are presented for the threshold k_t equal to three times the at-the-money Black-Scholes (BS) implied volatility.¹³ This specification of k_t corresponds to a median threshold of 25 %.

The estimates for FI_t are presented in figure (1.5). Most of the noticeable movements in the series correspond to well known periods of oil price or aggregate uncertainty, such as the 1st Gulf War in 1990 and 1991, the financial crisis in 2008 and NBER recessions. Similar intuitive results are presented in figure (1.5) for the individual tail measures, alongside the risk neutral and statistical expected quadratic variations.

Similar to results from the equity index market presented by Bollerslev and Todorov (2011b), we find that in the crude oil futures markets the \mathbb{P} -tail distribution implied by the futures data is dwarfed by the corresponding \mathbb{Q} measures. The estimates presented in table (1.1) indicate that on average, the statistical variation measure for the left tail is about 200 times smaller than the \mathbb{Q} measure counterpart. The corresponding ratio for the right tail is around 50 and also sufficiently small in order to conclude that changes in the tail premia are almost entirely driven by movements in the tail variations under \mathbb{Q} measure. Thus the tail

¹¹The monthly pooling ensures a sufficient number of options for estimation throughout the sample. It also has the advantage that potential monthly seasonalities are averaged out.

¹²The original sample constitutes 6519 trading days. Some days around Christmas, Thanksgiving and July 4th feature irregular trading hours and were discarded.

¹³We also experimented with other thresholds, obtaining qualitatively similar results. Table (1.6) in Appendix C displays the result for k_t four times the at-the-money Black-Scholes implied volatility.

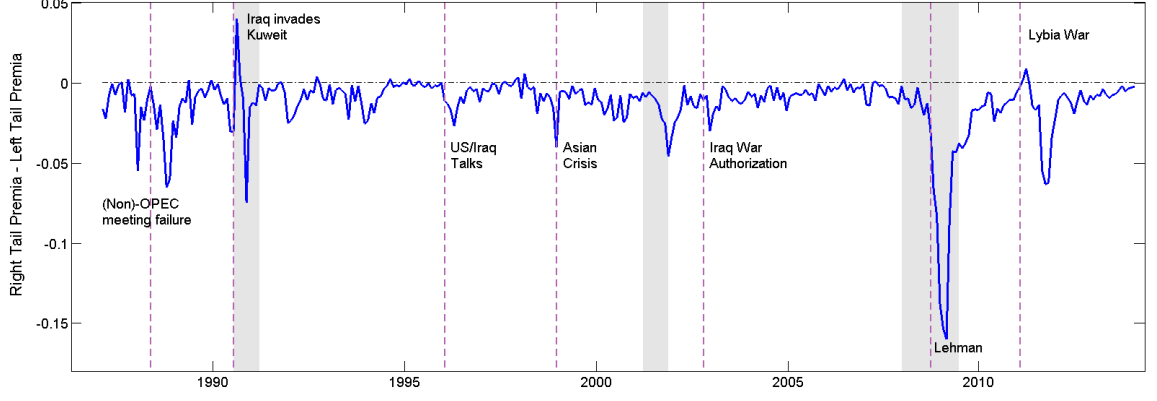


Figure 1.1: This figure displays the difference between the tail measures, $RJPV_{t,T}^Q - LJPV_{t,T}^Q$, in annualized form. The tail measures are computed for the threshold $k_t = 3 \times \text{BS implied volatility}$. Shaded areas represent NBER recessions.

variation premia appear well approximated by the risk neutral variation measures only:¹⁴

$$RJPV(k_t) \approx -RJPV_{t,T}^Q \quad \text{and} \quad (1.25)$$

$$LJPV(k_t) \approx -LJPV_{t,T}^Q. \quad (1.26)$$

Interestingly, this indicates that changes in the objective jump distributions play a minor role in explaining the time variation in the size of the tail premia.¹⁵

A second important finding, reported in Appendix A, is that the statistical left and right tail variation measures are approximately symmetric. The symmetry implies that in each point of time, the conditional probabilities of a large upward jump is roughly equal to the conditional probability of a large downward jump. Together with strong time variation of the difference of $RJPV_{t,T}^Q$ and $LJPV_{t,T}^Q$ documented in figure (1.5), this provides additional evidence that the tail risk premia are only loosely connected to the statistical tail variation measures.

On average, the variation risk premia for the left tail is much larger than the premia for the right tail (table 1.1). This is consistent with commodity futures markets being in normal “backwardation”, an idea first put forth by Keynes (1976). He postulated that producers of the physical commodity that want to hedge their output will have to pay a risk premium for speculators that take on the matching long positions in futures markets. Accordingly, speculators will demand a larger premium for their exposure to downside tail risk.

¹⁴A similar approximation empirically holds for the equity market. As Bollerslev, Todorov, and Xu (2014) point out, this conveniently avoids peso-type estimation problems.

¹⁵There is a similar finding for the equity index market. See e.g. Bollerslev and Todorov (2011a), Bollerslev and Todorov (2011b) and Bollerslev, Todorov, and Xu (2014).

	Mean	SD	AR(1)
$LJV_{t,T}^{\mathbb{Q}}$	0.0233	0.0253	0.8444
$RJV_{t,T}^{\mathbb{Q}}$	0.0081	0.0101	0.7730
$LJV_{t,T}^{\mathbb{P}}$	0.0001	0.0001	0.6698
$RJV_{t,T}^{\mathbb{P}}$	0.0002	0.0002	0.7103
OILVIX2	0.1353	0.1293	0.8374
VRP	-0.0244	0.0741	0.1897

Table 1.1: Summary statistics for the monthly estimates of the tail variation measures and the traditional variation measures. SD stands for the standard deviation, AR(1) for first order autocorrelation. The sample period is 1987 to 2013, comprising 324 observations. All measures are presented in annualized form. The tail variation measures are evaluated at $k_t = 3 \times$ Black-Scholes implied volatility.

The implied total variation $QV_{t,T}^{\mathbb{Q}}$ is computed using all OTM options on a given day. For comparability with the tail measures, the monthly series we present in figure (1.5) is calculated by taking the average over the respective calendar month. In the computation of the realized total variation we also account for the overnight returns. The contribution of the squared overnight returns to the entire daily observation, π_t , is about 50% on average.¹⁶ We compute the daily series by an appropriate scaling of the intraday realized variance RV_t and obtain the monthly series by averaging over days.¹⁷ Our estimates of VRP_t are then based on the difference between the expected quadratic variation under the \mathbb{Q} -measure and the realized variation for the respective month.¹⁸ Over our sample period, VRP_t is about -2%. This figure right between the numbers of Trolle and Schwartz (2010), who report an average premium of almost -3% over the 1996 - 2006 sample period and Kang and Pan (2013), who report a premium of -1.65% for a sample period slightly shorter than ours.

For a threshold of $k_t = 3 \times$ BS-implied volatility used here, the sum of the two tail premia is on average about 3% and slightly larger than the absolute value of the average VRP_t . Of the total VRP_t , about $E(RJV^{\mathbb{Q}})/E(VRP_t) = 33\%$ come from the right tail and about $E(LJV^{\mathbb{Q}})/E(VRP_t) = 94\%$ from the left tail. This suggests that the entire variance risk premium is due to compensation for tail variations, while variations due to continuous

¹⁶This number is larger than then the average contribution of the squared overnight to the daily volatility, see e.g. Ahoniemi and Lanne (2013). Part of this is due to the relatively small active trading window we are considering for the realized measures in the previous section. For additional details the reader is referred to Appendix A.

¹⁷Since there is no consensus in the literature, we experimented also with different forms of the scaling the contributions of the overnight returns. The level of the average variation presented in this section is of course not affected by the scaling, while the results for the monthly series used for forecasting in the next sections are qualitatively similar.

¹⁸This obviously differs from the definition of VRP_t which is based on the expectation of the realized variation under the \mathbb{P} -measure rather than the ex-post realized variation. For the unconditional estimates tabulated here this distinction does not matter if we assume that the expectation error is zero.

	FI_t	FI_{t-1}	FI_{t-2}	FI_{t-3}
$r_{S,t}$	0.1426** (0.0552)	-0.0319 (0.0558)	-0.119** (0.0555)	-0.148*** (0.0594)
$r_{F,t}$	0.2117*** (0.0545)	0.0491 (0.0557)	-0.0462 (0.0558)	-0.1002* (0.0562)

Table 1.2: Contemporaneous and 1-3 months spot return ($r_{S,t}$) and futures return ($r_{F,t}$) correlation with $FI_t \approx RJV^{\mathbb{Q}} - LJV^{\mathbb{Q}}$. The estimates are based on monthly observations from 1988 to 2013.

price movements and small jumps earn no or even a small positive premium.

In order to investigate how the tail premia interact with the spot prices, we also document the correlations of the relative movements in the premia - captured by the fear index FI_t - with current and future spot returns. Table (1.2) shows that the contemporaneous spot and future returns and FI_t are positively correlated, while this correlation turns negative for future spot and futures returns. Thus, these prices seem to exhibit an “overshooting” in the wake of large upside risks, allowing the price to decrease over the subsequent periods. We will address this issue more carefully in the next section where we discuss the forecasting properties of our indicators.

It is also interesting to see how the estimated jump intensities relate to the futures risk premium due to large jumps. Using our separate estimates for the left and the right tails under the respective measure, we obtain this premium through equation (1.5). Similar to our results from the tail variation premia, the futures premia for the left tail tends to be larger than that for the right tail. The average of the (continuously compounded) futures premium due to large jumps period is 3.7% over our sample period. In comparison, the average (continuously compounded) total futures risk premium computed from the 1st back contract is 7.5%, suggesting that about half of the short maturity futures risk premium is due to tail risk.

1.4 Predictability of Futures and Spot Returns

1.4.1 Predictive Regression Framework and Control Variables

This section presents new predictability evidence of the tail premia for crude oil spot and futures returns. The baseline forecasts are performed in-sample, while cross-validation techniques are employed for the out-of-sample robustness check. If the premia capture oil market agents’ attitudes toward tail risks, we would expect a large upside (downside) tail variation risk premium to be associated with relatively small (large) returns. We test this hypothesis

in a regression framework for predictability of futures and spot returns of the form

$$r_{j,t+i} = \beta_{0,i} + \beta_{1,i} \cdot LJV_{t,T}^{\mathbb{Q}} + \beta_{2,i} \cdot RJV_{t,T}^{\mathbb{Q}} + Controls_t' \cdot \beta_{3,i} + \epsilon_{t,i}, \quad j = \{S, F\}, \quad (1.27)$$

where $r_{S,t+i}$ is the i th-month ahead spot return and $r_{F,t+i}$ is the i th-month ahead futures return, while $LJV_{t,T}^{\mathbb{Q}}$ and $RJV_{t,T}^{\mathbb{Q}}$ represent the left and the right tail premia, respectively, approximated by \mathbb{Q} -measure variations. The vector $Controls_t$ include both macroeconomic-financial variables and crude oil market specific variables that are potential predictors of commodity spot and futures returns (see e.g. Bessembinder and Chan 1992; Hong and Yogo 2012). For $i > 1$, we employ overlapping regressions in order to enhance the efficiency of our estimates, using robust Newey-West standard errors so as to account for the autocorrelation in the residuals induced by the overlap. The lag length for the computation of the standard errors is chosen twice the length of the overlap.

We first describe the set of oil market specific control variables. As suggested by the recent literature, we include the estimated VRP_t . We compute VRP_t as the difference between the lagged expected variation under the risk neutral and the actual variation in period t as suggested in Bollerslev, Tauchen, and Zhou (2009) and as described in the previous section. For additional robustness analysis, we also include the contemporaneous expectation $QV_t^{\mathbb{Q}}$ and the contemporaneous realized variation separately. The oil market specific variables further include changes in oil inventories, obtained as the monthly storage level from the web site of the U.S. Energy Information Administration (EIA) and open interest growth. The computation of the open interest variable is based on the open interest of futures and options combined obtained from the CFTC website and computed as the 12-month growth rates taking geometric averages as suggested by Hong and Yogo (2012). Finally, we include the slope of the term structure as measured by the net ratio of the current spot price over the 1st back futures contract.¹⁹

We control for macroeconomic conditions by including the short term interest rate, computed as the yield of a 3-month T-Bill, and the Aruoba-Diebold-Scotti Business Conditions Index (ADS) published by the Federal Bank of Philadelphia. The index is based on may economic indicators in the U.S. and a higher value is associated with better economic conditions. Since crude oil prices might be also driven by global rather than US-specific factors, we include the Real Activity Index developed in Kilian (2009). The other variables include the yield spread, computed as the difference between Moody's Aaa and Baa corporate bond yields and the CBOE VIX, a measure of the implied volatility of S&P 500 index options.

¹⁹We also experimented with other definitions of the slope of the term structure, e.g. the net ratio of the spot price with the 3rd back futures contract and the net ratio of the first two futures contracts, yielding almost identical results.

1.4.2 Forecasting Results

We first discuss the forecasting results for crude oil futures returns, presented in columns (5) - (8) of table (1.3). For all regressions, the coefficients have the expected sign. A relatively high right tail premium is associated with negative futures returns, and the left tail premium with positive returns. For the model without controls, all coefficients are statistically significant at the 5% significance level and most at the 1% significance level. This confirms our intuition that the tail risk premia are associated with a substantial change in the oil futures premia. The results are robust to the inclusion of standard predictors of crude oil prices. The only exception is the noticeable rise in the standard error for the left tail premium in the three month horizon regression, which renders the coefficient statistically insignificant at conventional significance levels. This indicates certain degree of correlation with some of the predictor variables, which is also noted through the increasing coefficient for the coefficient associated with the right tail premium when controlling for the other predictors. Jointly, the tail premia are always significant at the 1% significance level as measured through F-test. The adjusted R^2 is 3.7% for the three month horizon and 6.8% for the six month horizon regression, which amounts to almost one fourth of the predictability associated with all regressors.

Variables	(1) $r_{S,t+3}$	(2) $r_{S,t+3}$	(3) $r_{S,t+6}$	(4) $r_{S,t+6}$	(5) $r_{F,t+3}$	(6) $r_{F,t+3}$	(7) $r_{F,t+6}$	(8) $r_{F,t+6}$
L3. $RJV_{t,T}^Q$	-7.70*** (2.044)	-10.13*** (2.254)			-5.27*** (1.872)	-9.54*** (2.364)		
L3. $LJV_{t,T}^Q$	2.56*** (0.510)	1.85* (1.059)			1.59*** (0.546)	1.49 (1.025)		
L6. $RJV_{t,T}^Q$			-11.03*** (2.845)	-16.83*** (2.901)			-9.76** (3.875)	-18.32*** (3.539)
L6. $LJV_{t,T}^Q$			5.27*** (0.698)	3.80*** (1.413)			3.47*** (0.860)	3.99*** (1.460)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R^2	0.0877	0.2371	0.1419	0.3131	0.0374	0.1634	0.0685	0.2827
Obs.	321	283	318	282	321	283	318	282
Wald test (p-value)	<0.001	<0.001	<0.001	<0.001	0.005	<0.001	<0.001	<0.001

Newey-West standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1.3: Forecasting results for $i = 3$ and $i = 6$ months. The dependent variable $r_{S,t+i}$ stands for spot returns, $r_{F,t+i}$ for futures returns. Wald test stands for an F-test of the joint significance of $RJV_{t,T}^Q$ and $LJV_{t,T}^Q$. The estimation period is 1987 - 2013. The differences in the number of observations is due to data availability for the control variables.

We now describe the forecasting results for crude oil spot returns, presented in columns (1) - (4) of table (1.3). For all regressions, the coefficients associated with the tail premia

have similar sizes to those of the futures regressions. Again, the significance of the coefficient for the left tail premium in the three month regression when the other predictor variables are is slightly decreased due to the rise in the standard error relative to the model without predictors. For all other regressions, the individual coefficients are significant at the 1% significance level. Again, a joint F-test for the premia allows us to reject the null of no significance at all conventional significance levels for all regressions. The adjusted R^2 s associated with the spot return regressions are much larger than those for futures returns. At the three month horizon, the tail measures account for about 9% of the variability in the spot return, while this number increases to 14% at the six month horizon, constituting almost half of the in-sample variability explained by all regressors.²⁰ Thus the tail risk measures have strong predictability for futures and spot price returns in the expected direction, and, judging from the fit of the regression, this predictability is larger for the spot than for the futures price. This is consistent with the idea that fears about large oil price fluctuations affect both the futures and the spot price, and that the convenience yield moves in the same direction as the futures risk premium.

For forecasting purposes, a high R^2 is not always indicative of a good model since it is always possible to increase the in-sample fit by adding additional regressors without improving out-of-sample forecasting power. In order to safeguard against a potential in-sample-overfit, we perform an out-of-sample cross-validation. Our cross-validation procedure uses the entire series as both in-sample and out-of-sample data. Specifically, we run repeating regressions using a single return observation from the original sample as the validation data, and the remaining observations as the training data.²¹ This is repeated until each observation in the sample is used once as the validation data, so that finally every return observation was treated “out-of-sample” once. The cross-validation statistics is based on the mean squared prediction error, hence lower values indicate a better fit. In order to evaluate the predictive out-of-sample performance of our predictors we estimate the forecasting model presented in equation (1.27) in the previous section first including the tail premia and then excluding it. The results are displayed in table (1.7) in Appendix C. Including the tail premia as predictive regressors reduces the cross validation statistics substantially for all models considered, indicating that their forecasting power is not only due to in-sample properties. Importantly, the out-of-sample MSPE for the model including $LJV_{t,T}^Q$ and $RJV_{t,T}^Q$ without additional control variables as regressors is always lower than the forecast assuming a con-

²⁰We also assessed whether the tail risk measures possess incremental forecasting power with respect to the control variables. This is indeed the case: The incremental adjusted R^2 for the regressions including all variables in comparison with regressions using the control variables only is about 10% and 11% for futures and spot returns, respectively, for the six months horizon and about 5% and 6% for the three months horizon.

²¹We assure that these returns are completely non-overlapping with the remaining returns by leaving an out of sample window of ± 4 months around the respective return for $i = 3$ months and ± 7 months for $i = 6$ months.

stant return and the no-change forecast. In contrast, the model including only the control variables displays a higher out-of-sample MSPE against the no-change forecast for the spot price regressions, suggesting that the R^2 related to these variables is driven at least in part by in-sample overfitting.

1.4.3 Risk or Fear: Where does the predictability come from?

After having provided evidence for the forecasting power of our novel predictors, we now turn to the question whether the premia a compensation for potential risks or rather describe the oil market's attitudes towards risks. From the definition of FI_t it follows that for approximately symmetric jump tails under the \mathbb{P} measure,

$$\begin{aligned} FI_t &= \frac{1}{T-t} [(E_t^{\mathbb{P}}(LJV_{t,T}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(LJV_{t,T}^{\mathbb{Q}})) - (E_t^{\mathbb{Q}}(LJV_{t,T}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(LJV_{t,T}^{\mathbb{Q}}))] \\ &\approx \frac{1}{T-t} [E_t^{\mathbb{Q}}(LJV_{t,T}^{\mathbb{Q}}) - E_t^{\mathbb{Q}}(RJV_{t,T}^{\mathbb{Q}})] . \end{aligned} \quad (1.28)$$

In this case, as shown by Bollerslev, Todorov, and Xu (2014), FI_t will be largely void of risk compensation associated with the temporal changes in the jump intensities and is therefore naturally interpreted as a proxy of oil market fears. The notion of a “fear index” is also warranted because the premia are not directly related to asymmetries in future actual jump probabilities. In table (1.8) presented in Appendix C we show that the \mathbb{Q} tail variation measures do not display any forecasting power for the realized jump variation beyond that predicted by symmetric measures such as the realized variance. In contrast, the forecasting results shown in table (1.4) indicate that the return predictability through our novel predictors is mainly driven by its asymmetry. Here we use the 6 month ahead forecasting regression and decompose the tail measure in their difference, FI_t and a level component $RJV_{t,T}^{\mathbb{Q}} + LJV_{t,T}^{\mathbb{Q}}$. Using these variables as single regressors, displayed in columns (2) and (3), shows that the regression with the fear index as a predictor variable yields a statistically significant coefficient that indicates that upside fears are associated with decreasing spot price returns. The R^2 for this regression is almost 10%, while the R^2 for the regression with the tail risk level measure is less than 2% and yields an statistically insignificant coefficient.²² These results are consistent with the idea that time variation in the effective risk aversion of oil market participants is an important driver of the short run fluctuation of the price of oil. Taken together, the results presented in this section are confirm evidence for the predictability of both futures and spot returns by our novel risk indicators.

²²Including both the fear index and the level component jointly yields, as in column (4), obviously yields the same R^2 as the baseline.

Variables	(1) $r_{S,t+6}$	(2) $r_{S,t+6}$	(3) $r_{S,t+6}$	(4) $r_{S,t+6}$
L6. $RJV_{t,T}^Q$	-11.03*** (2.626)			
L6. $LJV_{t,T}^Q$	5.27*** (0.615)			
L6. FI_t		-3.84*** (0.569)		-8.15*** (1.559)
L6. $(RJV_{t,T}^Q + LJV_{t,T}^Q)$			1.01 (0.881)	-2.88*** (1.099)
Adj. R^2	0.1419	0.0950	0.0180	0.1419
Obs	318	318	318	318

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.4: Forecasting results for 6 month oil spot returns, denoted by $r_{S,t+6}$. The estimation period is 1987 - 2013.

1.4.4 VAR Estimates on the Impact of the Oil Fear

In this section, we compare our forecasting results with estimates from VAR models, where we include FI_t as a directional measure of oil market uncertainty. The VAR framework has shown useful for structural analysis of oil price shocks and forecasting the real price of oil (Kilian 2009; Baumeister and Kilian 2012). Building on the reduced form version of the model proposed by Kilian and Murphy (2014), the variables in the estimation order are oil production (in percentage changes), the global real activity proposed by Kilian (2009), changes in crude oil inventories, FI_t , the six month futures spread ($CY_{t,6}$) and the nominal price of oil (in percentage changes).²³ In a second exercise, we replace the price of oil by the percentage change of the first back futures contract.²⁴ The identification of shocks to FI_t is based on a Cholesky-decomposition of the reduced form errors. Including production, real activity and inventory as the first variables in the VAR ensures the impact of these variables is already controlled when looking at the impact of shocks to FI_t on the oil and futures price.²⁵ Given the tradeoff between overparameterization and allowing for sufficient lags to account for business cycle effects, we estimate the VAR with 12 lags.

Figure (1.2) plots the impulse response function of oil spot prices and the six months futures spread to a shock in the oil fears index. Oil spot prices react with an instantaneous

²³Expressing the price of in log percentage changes ensures that the cumulative impulse responses reflect the percentage change of oil over the entire period.

²⁴The calculation of the percentage change of the first back futures contract is always based on the same contract for a given period. However across periods, contracts will differ.

²⁵As pointed out by Kilian and Murphy (2014), inventory will in general also adjust quickly in response to uncertainty shocks. The results presented here are robust to a change in the ordering of the variables such that FI_t is included before inventories.

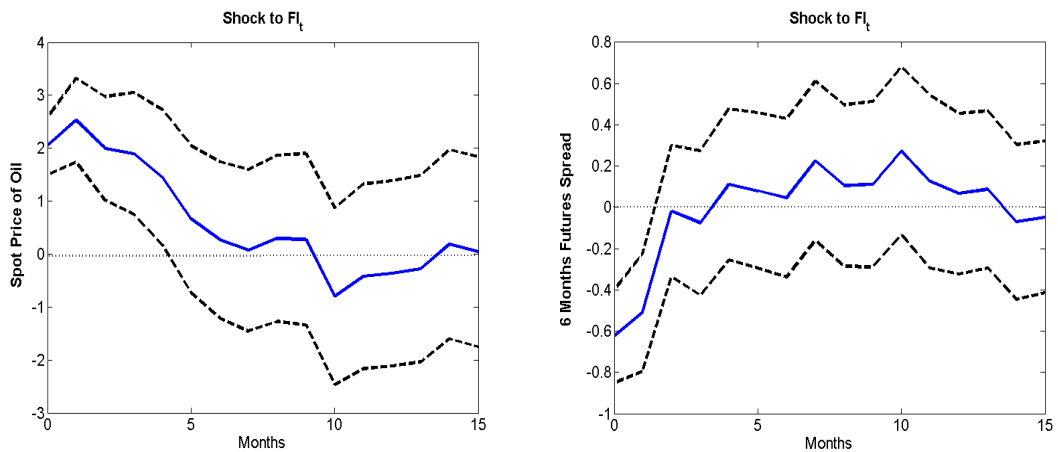


Figure 1.2: Responses of the spot price of oil and the six months futures spread to a one standard deviation shock. The estimates are based on monthly data, 1987:1 - 2012:6. Dashed lines represent 90% confidence intervals.

increase of about 2%, and a gradual reversion to previous levels over the following months. The 90 % confidence intervals are plotted around this, highlighting that this impact is statistically significant over the first months. The six months futures spread reacts with an immediate, yet much smaller decline and a faster reversion to previous levels. Instead, the response of inventories to fear shocks, presented in Appendix C, do not provide any evidence for a systematic reaction of inventories to FI_t .²⁶ These effects are consistent with the idea that shocks to upside (downside) fears are associated with an immediate increase (decrease) of the price of oil that is due to the combination of an increase (decrease) in the net convenience yield and an decrease (increase) in the risk premium. As such, changes in relative uncertainty do not require an immediate response of inventory in order to display discernible effects on prices.

For comparison, the responses of spot prices to a shock to real activity and implied oil price volatility are presented in figure (1.3). The reaction of the price of oil to real activity appears more persistent, yet smaller on impact, whereas the estimated effect of the implied oil price volatility shocks is largely insignificant. The results are consistent with the idea that changes in the relative uncertainty about upside and downside fears are an important driver of short run fluctuations in oil prices and futures returns.

²⁶This might also be an indication of the redundancy of inventories when the futures spread is included. I thank Christiane Baumeister for the comment.

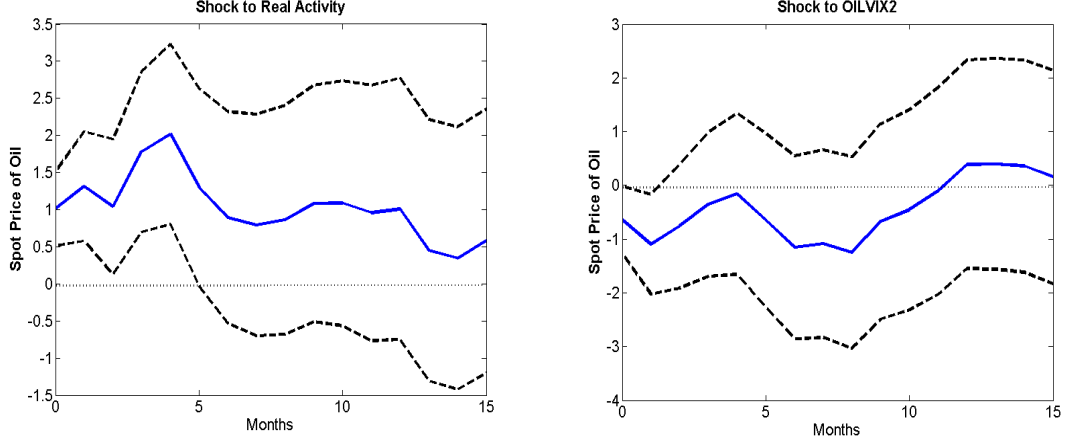


Figure 1.3: Responses of the spot price of oil to shocks in real activity and the option implied variance, OILVIX2. The estimates for the responses to the implied variance are based on a VAR model that contains the same ordering as described in the text and OILVIX2 instead of FI_t . The estimates are based on monthly data, 1987:1 - 2012:6. Dashed lines represent 90% confidence intervals.

1.5 Interaction with the Macroeconomy

1.5.1 Oil and Aggregate Uncertainty

Thus far, we have treated the oil risk factors in isolation from the aggregate asset uncertainty. In this section we address their interaction. A recent strand of literature, building on Bollerslev, Tauchen, and Zhou (2009), has shown that the variance risk premium embedded in stock market index options and futures is a suitable measure of aggregate uncertainty. This work has been extended by Bollerslev and Todorov (2011b) and Bollerslev, Todorov, and Xu (2014), suggesting that a substantial fraction of the variance risk premium and its forecasting power for stock market and portfolio returns is due to the aggregate “fears” as measured through the stock market fear index. This raises the question whether the aggregate risk measures proposed by these authors predicts oil spot and future prices beyond the oil risk measures proposed in this paper. Table (1.9) in Appendix D shows that the oil and stock market tail variation measures are indeed highly correlated.²⁷ The left tail oil variation measure $LJV_{t,T}^Q$ and the stock market fear index $FI_{t,SPX}$ exhibit the highest correlation of 65%.

In table (1.10) in Appendix D we present the forecasting results for six months futures

²⁷The stock market variance risk premium $VRP_{t,SPX}$ and fear index $FI_{t,SPX}$ are based on S&P 500 index options and futures following the methodology of Bollerslev, Todorov, and Xu (2014) for the years 1996 to 2013. I would like to thank Lai Xu for providing the data on the S&P 500 fear index and Marek Raczko for providing the data on the S&P 500 variance risk premium.

and spot price returns, respectively. The aggregate fear index and variance risk premium appear to contain some explanatory power for the 6 month spot price return regression, with an R^2 of about 5%.²⁸ However, this effect is completely dominated by the oil specific tail risk variation measures, implying that the oil specific measures already entail relevant information from the aggregate measures. On a more general level these results also suggest that time-varying disaster fears embedded in option prices on individual assets, not only on market indexes, convey important information on individual premia beyond that implied by the market.

In addition to these results, we also address the question whether $LJV_{t,T}^Q$ and $RJV_{t,T}^Q$ predict stock market returns. We find little evidence of a stable relationship, since these forecasting results depend crucially on the sample period, in particular on the inclusion of data during and after the financial crises. Table (1.11) in Appendix D shows that prior to the outbreak of the financial crisis, a relatively high left tail variation measure was associated with lower stock market returns, and a relatively high right tail variation with higher stock market returns. However, extending the sample beyond the outset of the financial crisis this relationship breaks down. These results - in terms of both the significance and signs of the coefficients - are consistent with the idea that $LJV_{t,T}^Q$ and $RJV_{t,T}^Q$ aggregate different types of uncertainty that are relevant for oil prices, e.g. oil supply and oil demand uncertainty, that individually have a very distinct relationship with the aggregate stock market. Prior to the financial crisis, the most notable event in terms of oil fears was the 1991 Gulf War episode, which is clearly identified with supply risk (Alquist and Kilian 2010). On the other hand, the financial crisis and its aftermath were associated mainly with uncertainty about future demand for oil.²⁹

1.5.2 Oil Uncertainty and Real Activity

Bernanke (1983) and more recently Elder and Serletis (2010) and Jo (2014) pointed out that uncertainty about oil prices, rather than the fluctuations per se, can have important effects on real economic activity. In this section we investigate this hypothesis by using our novel uncertainty measures to predict industrial production, which is available at a monthly frequency.

Our regressions, presented in table (1.12) in Appendix D, indicate that the right tail

²⁸The important predictor seems to be $VRP_{t,SPX}$, a result that is consistent with the work of Bollerslev, Todorov, and Xu (2014) who show that $VRP_{t,SPX} - FI_{t,SPX}$ is mainly associated with economic uncertainty, while $FI_{t,SPX}$ captures attitudes toward risk.

²⁹A different explanation for this result might be that oil derivative markets have only recently become more integrated in the broader financial system, and that this integration went alongside fundamental changes between the stock market and oil price relationship (Christoffersen and Pan 2014). We leave this hypothesis for further research

variation measure $RJV_{t,T}^Q$ has a statistically significant impact on growth in industrial production, while the left tail variation measure does not contain any additional information. The incremental R^2 from the inclusion of the right tail measure is about 5%, is robust to the inclusion of oil market and aggregate control variables, and extends also to the 6 month horizon.³⁰ Intuitively, a large $RJV_{t,T}^Q$ is always bad news for the oil importing economy because it can be due to (i) fears related to supply cuts (2) uncertainty about (global) economic growth, both of which should have a negative effect on economic activity. In contrast, a large premium for the left tail might be good news for the economy if the corresponding fears are related to oil market specific events.

1.6 Conclusion

Oil prices are difficult to forecast and exhibit wild swings or “excess volatility” that are difficult to rationalize by changes in fundamentals alone. We find that the jump risk premia embedded on crude oil future options contain important information on oil market fears and contribute to the explanation of oil price volatility. These premia are economically large, vary substantially over time and significantly forecast crude oil futures and spot returns. This result is robust after controlling for macro-finance and oil market specific variables, and importantly, for time-varying aggregate disaster fear as measured by *S&P500* option implied tail risk. Instead, our oil uncertainty measures appear to conveniently aggregate oil price uncertainty derived from different sources, e.g. oil supply and oil demand uncertainty, that individually have a very distinct relationship with aggregate uncertainty.

We show that oil futures prices overshoot (undershoot) in the presence of upside (downside) tail fears in order to allow for smaller (larger) risk premia thereafter. Consistent with the theory of storage, this overshooting (undershooting) is amplified for the spot price because of time varying benefits from holding inventory that work in the same direction. These results are complementary to storage models using risk neutrality, and stress the importance of time varying risk premia in explanations of large swings in oil prices. On a more general level it is shown that time-varying disaster fears embedded in option prices on individual assets, not only on market indexes, convey important information on the risk premia and price dynamics of these assets.

³⁰These results are confirmed by VAR estimates in a framework proposed by Bloom (2009). The results are available upon request

1.A Appendix: Computation of the Jump Measures

Estimating realized and expected jumps under the statistical measure

This section provides further details on the estimation of the jump properties under the statistical (objective) measure, which are computed from data on 5-minute intraday returns. We use the notation of Bollerslev and Todorov (2011b) and divide the trading day t into the $[t, t + \pi_t]$ overnight period and the $[t + \pi_t, t + 1]$ active trading period, comprising $n + 1 = 54$ price observations. Denoting $\Delta_{t,i}f \equiv f_{t+\pi+i} - f_{t+\pi+i-1}$, where f denotes the logarithm of the futures price, we have for a suitable threshold $\alpha_{t,i}$

$$\sum_{i=1}^n (\Delta_{t,i}f)^2 \xrightarrow{p} \int_{t+\pi_t}^{t+1} \sigma_s^2 ds + \int_{t+\pi_t}^{t+1} \int_{\mathbb{R}} x^2 \mu(ds, dx) \quad \text{and} \quad (1.29)$$

$$\sum_{i=1}^n (\Delta_{t,i}f)^2 1_{|\Delta_{t,i}f| \leq \alpha_{t,i}} \xrightarrow{p} \int_{t+\pi_t}^{t+1} \sigma_s^2 ds \equiv CV_t. \quad (1.30)$$

We allow the truncation levels $\alpha_{t,i}$ to vary with both the daily and the intraday volatility following Bollerslev and Todorov (2011a). The time-of-day factor, TOD_j is then computed via

$$TOD_j = \frac{\sum_{m=0}^N (\Delta_{m(\pi+n),i}f)^2 1_{|\Delta_{m(\pi+n),i}f| < \bar{\alpha}}}{\sum_{m=0}^N 1_{|\Delta_{m(\pi+n),i}f| < \bar{\alpha}}} / \frac{\sum_{m=0}^N \sum_{j=1}^n (\Delta_{m(\pi+n),i}f)^2}{\sum_{m=0}^N \sum_{j=1}^n 1_{|\Delta_{m(\pi+n),i}f| < \bar{\alpha}}}, \quad (1.31)$$

where $\bar{\alpha} = 3\sqrt{\Pi \cdot 0.5} \sqrt{\frac{1}{N} \sum_{m=0}^N \sum_{j=1}^{n-1} |\Delta_{m(\pi+n),j}f| |\Delta_{m(\pi+n),j+1}f|}$. Given the time-of-day factor, the time-varying threshold $\alpha_{j,t}$ is then computed as

$$\alpha_{j,t} = 3\left(\frac{1}{n}\right)^{0.49} \sqrt{CV_{t-n,t}} TOD_j. \quad (1.32)$$

The dynamics of our empirical jumps in equation (1.24), require an estimate of $v^{\mathbb{P}}$. This estimate is based on medium and large sized jump tails using the EVT proposed by Bollerslev and Todorov (2011a). In particular, defining $\psi^+(x) = e^x - 1$ and $\psi^-(x) = e^{-x}$, $v_{\psi}^+(y) = \frac{v(\ln(y+1))}{y+1}$ and $v_{\psi}^-(y) = \frac{v(-\ln y)}{y}$, $y > 0$, the jump tail measures are

$$\bar{v}_{\psi}^{\pm}(x) = \int_x^{\infty} v_{\psi}^{\pm}(u), \quad (1.33)$$

with $x > 0$ for $\bar{v}_{\psi}^+(x)$, and $x > 1$ for $\bar{v}_{\psi}^-(x)$. Under the assumption that \bar{v}_{ψ}^{\pm} belong to the domain of attraction of an extreme value distribution (see Bollerslev and Todorov 2011b),

it follows that

$$1 - \frac{\bar{v}_\psi^\pm(u+x)}{\bar{v}_\psi^\pm} \sim G(u; \sigma^\pm, \xi^\pm), \quad u > 0, x > 0, \quad (1.34)$$

where $G(u; \sigma^\pm, \xi^\pm)$ is the CDF of a generalized Pareto distribution with

$$G(u; \sigma^\pm, \xi^\pm) = \begin{cases} 1 - (1 + \xi^\pm u / \sigma^\pm)^{-1/\xi^\pm}, & \xi^\pm \neq 0, \sigma^\pm > 0 \\ 1 - e^{-u/\sigma^\pm}, & \xi^\pm = 0, \sigma^\pm > 0. \end{cases} \quad (1.35)$$

Now, for a large threshold tr^\pm , the integrals corresponding to the jump tail measures under \mathbb{P} are a function of the parameter vector

$$\Theta \equiv [\sigma^\pm, \xi^\pm, \alpha_0^\pm \bar{v}_\psi^\pm(tr^\pm), \alpha_1^\pm \bar{v}_\psi^\pm(tr^\pm)], \quad (1.36)$$

which are estimated using the exact GMM framework suggested by Bollerslev and Todorov (2011a). The moment conditions used for estimation are

$$\frac{1}{N} \sum_{t=1}^N \sum_{i=1}^n \phi_j^\pm(\psi^\pm(\Delta_{t,if}) - tr^\pm) 1_{\psi^\pm(\Delta_{t,if}) > tr^\pm} = 0, \quad j = 1, 2 \quad (1.37)$$

$$\frac{1}{N} \sum_{t=1}^N \sum_{i=1}^n 1_{\psi^\pm(\Delta_{t,if}) > tr^\pm} - \alpha_0^\pm \bar{v}_\psi^\pm - \alpha_1^\pm \bar{v}_\psi^\pm(tr^\pm) CV_t = 0 \quad (1.38)$$

$$\frac{1}{N} \sum_{t=2}^N \left(\sum_{i=1}^n 1_{\psi^\pm(\Delta_{t,if}) > tr^\pm} - \alpha_0^\pm \bar{v}_\psi^\pm(tr^\pm) - \alpha_1^\pm \bar{v}_\psi^\pm(tr^\pm) CV_t \right) CV_{t-1} = 0, \quad (1.39)$$

where

$$\phi_1^\pm(u) = -\frac{1}{\sigma^\pm} + \frac{\xi^\pm}{(\sigma^\pm)^2} \left(1 + \frac{1}{\xi^\pm}\right) \left(1 + \frac{\xi^\pm u}{\sigma^\pm}\right)^{-1} \quad \text{and} \quad (1.40)$$

$$\phi_2^\pm(u) = \frac{1}{(\xi^\pm)^2} \ln \left(1 + \frac{\xi^\pm u}{\sigma^\pm}\right) - \frac{u}{\sigma^\pm} \left(1 + \frac{1}{\xi^\pm}\right) \left(1 + \frac{\xi^\pm u}{\sigma^\pm}\right)^{-1} \quad (1.41)$$

are the scores associated with the log-likelihood function of the generalized Pareto distribution. Facing the trade off between a sufficient number of observations of medium and large jumps on the one hand, and the approximation of the jump tails by the generalized Pareto distribution on the other, our choice of tr^\pm corresponds to a jump in the log price of $\pm 1.2\%$. In total, we detect 3266 (3756) positive (negative) jumps, out of which 134 (198) are above the threshold. The large jumps, displayed in figure (1.4), seem to cluster and the occurrences of positive and negative jumps appear relatively symmetric.

The parameter estimates for our specification of the statistical large jumps' dynamics,

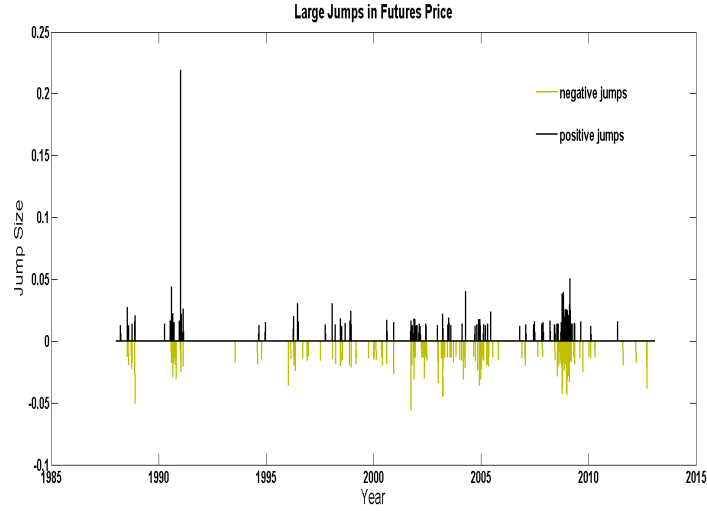


Figure 1.4: Detected large jumps in the log price. The threshold is $\pm 1.2\%$, the estimates are based on 5-minute intraday futures data from 1988 to 2013.

presented in table (1.5) provide further evidence for this symmetry.

Left Tail			Right Tail		
Parameter	Estimate	St. Error	Parameter	Estimate	St. Error
ξ^-	0.3107	0.0707	ξ^+	0.3603	0.1054
$100 \cdot \sigma^-$	0.4317	0.0429	$100 \cdot \sigma^+$	0.3446	0.0457
α_0^-	-0.1194	0.0059	α_0^+	-0.0906	0.0046
α_1^-	11.7926	0.4172	α_1^+	8.7404	0.3248

Table 1.5: Estimates for \mathbb{P} tail parameters. The estimates are based on 5-minute intraday futures data from 1988 to 2013.

1.B Appendix: Data and Empirical Measures

Tail Risk and Traditional Volatility Measures

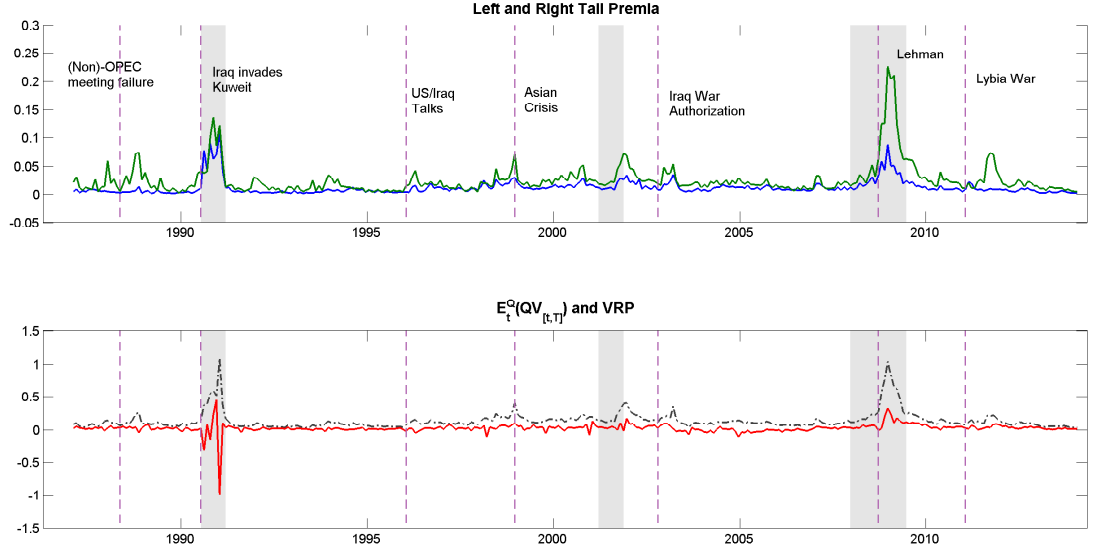


Figure 1.5: The figure on top displays the left tail (green) and right tail (blue) variation measures under \mathbb{Q} . The figure on the bottom displays the traditional measures $E_t^{\mathbb{Q}}(QV_{t,T})$ (dashed line) and $VRP_t \equiv E_t^{\mathbb{Q}}(QV_{t,T}) - RV_{t,T}$ (red line). All measures are presented in annualized form. The tail measures are computed for the threshold $k_t = 3 \times$ BS implied volatility. Shaded areas represent NBER recessions.

Robustness to the choice of k_t

Table (1.6) displays the estimated tail risk premia for a larger thresholds, $k_t = 4 \times$ the at-the-money Black-Scholes implied volatility and $k_t = 6.8 \times$ the at-the-money Black-Scholes implied volatility.

	Obs.	Mean	S.d.
$k_t = 4 \times$ ATM Black-Scholes implied volatility			
$LJV_{t,T}^{\mathbb{Q}}$	324	0.0113	0.0130
$RJV_{t,T}^{\mathbb{Q}}$	324	0.0032	0.0041
$k_t = 6.8 \times$ ATM Black-Scholes implied volatility			
$LJV_{t,T}^{\mathbb{Q}}$	324	$1.060 \cdot 10^{-3}$	$1.467 \cdot 10^{-3}$
$RJV_{t,T}^{\mathbb{Q}}$	324	$0.188 \cdot 10^{-3}$	$0.334 \cdot 10^{-3}$

Table 1.6: Summary Statistics for $LJV_{t,T}^{\mathbb{Q}}$ and $RJV_{t,T}^{\mathbb{Q}}$ based on pooled monthly data; evaluated at $k_t = 4 \times$ and $k_t = 6.8 \times$ Black-Scholes implied volatility.

1.C Appendix: Predictability of Futures and Spot Returns

Cross validation

	Futures				Spot			
	$r_{F,t+3}$	$r_{F,t+3}$	$r_{F,t+6}$	$r_{F,t+6}$	$r_{S,t+3}$	$r_{S,t+3}$	$r_{S,t+6}$	$r_{S,t+6}$
Controls	No	Yes	No	Yes	No	Yes	No	Yes
No change forecast	30.5	-	55.0	-	34.3	-	59.7	-
With $RJV_{t,T}^{\mathbb{Q}}, LJV_{t,T}^{\mathbb{Q}}$	29.1	26.7	48.9	41.0	33.0	33.6	50.8	57.1
Without $RJV_{t,T}^{\mathbb{Q}}, LJV_{t,T}^{\mathbb{Q}}$	29.8	28.1	51.8	46.7	33.9	35.8	52.2	65.2

Table 1.7: Cross validation statistics for forecasting models with and without $RJV_{t,T}^{\mathbb{Q}}$ and $LJV_{t,T}^{\mathbb{Q}}$. The dependent variable $r_{F,t+3}$ stands for 3 month futures returns, $r_{F,t+6}$ for six month futures returns, $r_{S,t+3}$ for three month spot returns and $r_{S,t+6}$ for six month spot returns. Each return is evaluated out of sample once, with a 7 (13 in the case of the 6 months prediction) out-of-sample window around the corresponding return. The control variables are those described in section 4. Each model is evaluated twice: Once including the the predictors $RJV_{t,T}^{\mathbb{Q}}$ and $LJV_{t,T}^{\mathbb{Q}}$, and once excluding them. A lower statistics indicates a lower out-of-sample MSPE. Values are multiplied by 100. Sample period is 1989 - 2013.

Forecasting empirical jump variations

	(1)	(2)	(3)	(4)
Variables	Emp. LJV	Emp. LJV	Emp. RJV	Emp. RJV
$L.RV2$	0.00*** (0.001)	0.00*** (0.001)	0.00*** (0.001)	0.00** (0.001)
$L.RJV_{t,T}^Q$		-0.01 (0.016)		0.02* (0.011)
$L.LJV_{t,T}^Q$		0.01 (0.004)		-0.00 (0.004)
Obs.	312	312	312	312
Adj. R^2	0.44	0.46	0.27	0.28

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1.8: Forecasting results for empirical jumps. Emp. LJV stands for the sum of squared log returns from large negative jumps in the futures price, Emp. RJV stands for the sum of squared log returns from positive negative jumps. $RV2$ is the average monthly realized variance.

1.D Appendix: Interaction with the Macroeconomy

Robustness to Aggregate Uncertainty and Fears

	$RJV_{t,T}^Q$	$LJV_{t,T}^Q$	$FI_{t,SPX}$	$VRP_{t,SPX}$
$RJV_{t,T}^Q$	1.00			
$LJV_{t,T}^Q$	0.85	1.00		
$FI_{t,SPX}$	0.60	0.65	1.00	
$VRP_{t,SPX}$	0.48	0.62	0.37	1.00

Table 1.9: Correlation between monthly oil risk measures and monthly stock market uncertainty measures. $FI_{t,SPX}$ is the fear index derived from S&P500 index options, $VRP_{t,SPX}$ the variance risk premium derived from S&P500 index options and futures as described in Bollerslev, Todorov, and Xu (2014). The sample period is 1996:1 to 2013:8.

Variables	(1) $r_{S,t+6}$	(2) $r_{S,t+6}$	(3) $r_{S,t+6}$	(4) $r_{S,t+6}$	(5) $r_{S,t+6}$	(6) $r_{F,t+6}$	(7) $r_{F,t+6}$	(8) $r_{F,t+6}$	(9) $r_{F,t+6}$	(10) $r_{F,t+6}$
$VRP_{t,SPX}$	0.03* (0.019)		0.03 (0.020)		0.01 (0.020)	0.02 (0.012)		0.02 (0.012)		0.00 (0.017)
$FI_{t,SPX}$		0.09** (0.036)	0.05 (0.037)		0.02 (0.049)		0.01 (0.031)	-0.01 (0.030)		-0.02 (0.060)
$RJV_{t,T}^Q$				-14.98* (7.902)	-14.33* (7.729)				-14.88** (7.125)	-13.91* (7.253)
$LJV_{t,T}^Q$				5.72*** (1.792)	4.97*** (1.815)				4.44*** (1.491)	4.35*** (1.731)
Adj. R^2	0.05	0.03	0.05	0.14	0.13	0.01	0.00	0.01	0.06	0.06
Obs.	210	210	210	210	210	210	210	210	210	210

Newey-West Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1.10: Forecasting results for six months oil futures and spot market returns. The dependent variable $r_{S,t+6}$ denotes the six months oil spot return, $r_{F,t+6}$ six months futures excess returns. $RJV_{t,T}^Q$ and $LJV_{t,T}^Q$ are the right tail oil variation measure and left tail oil variation measure. $FI_{t,SPX}$ is the fear index computed from S&P 500 as proxied through the left tail variation measure, suggested in Bollerslev, Todorov, and Xu 2014. $VRP_{t,SPX}$ is the variance risk premium computed from S&P 500 futures and options. The sample period is 1996 - 2013.

Forecasting stock market returns

	(1) 1989 - 2013	(2) 1989 - 07/2008
Variables	$r_{Mkt,t+6}$	$r_{Mkt,t+6}$
L6. $RJV_{t,T}^Q$	-1.50 (3.349)	5.78*** (1.409)
L6. $LJV_{t,T}^Q$	1.36 (1.075)	-3.84*** (1.344)
Observations	294	229

Newey-West standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1.11: Forecasting results for stock market returns. $r_{Mkt,t+6}$ is the six months market excess return using CRSP data.

Forecasting real activity

Variables	(1) $r_{IP,2}$	(2) $r_{IP,2}$	(3) $r_{IP,2}$	(4) $r_{IP,2}$
$L2.RJV_{t,T}^Q$		-0.26*** (0.084)	-0.27*** (0.070)	-0.28** (0.133)
$l2.r_{IP,2}$	0.35*** (0.057)	0.25*** (0.071)	0.25*** (0.066)	0.24*** (0.073)
$L4.r_{IP,2}$	0.30*** (0.115)	0.22** (0.094)	0.22** (0.094)	0.23** (0.093)
$L6.r_{IP,2}$		-0.04 (0.085)	-0.04 (0.078)	-0.04 (0.075)
$L2.r_{SPX,1}$				0.03*** (0.010)
$L2.r_{WTI,1}$				0.00 (0.006)
$L2.OILVIX2$				0.00 (0.012)
$L2.LJV_{t,T}^Q$		-0.01 (0.038)		
Observations	316	316	316	316
Adj. R^2	0.29	0.34	0.35	0.37

Newey-West standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1.12: Forecasting results for industrial production growth. $r_{IP,2}$ is the two month growth rate of US industrial production, $r_{SPX,1}$ is the one month stock market return, $r_{WTI,1}$ is the one month increase in the crude oil spot price, OILVIX2 is the option implied oil market volatility.

References

- Ahoniemi, Katja and Markku Lanne (2013). “Overnight stock returns and realized volatility.” *International Journal of Forecasting* 29.4, pp. 592–604.
- Alquist, Ron and Lutz Kilian (2010). “What do we learn from the price of crude oil futures?” *Journal of Applied Econometrics* 25.4, pp. 539–573.
- Barone-Adesi, Giovanni and Robert E Whaley (1987). “Efficient analytic approximation of American option values.” *The Journal of Finance* 42.2, pp. 301–320.
- Barro, Robert J (2006). “Rare disasters and asset markets in the twentieth century.” *The Quarterly Journal of Economics*, pp. 823–866.
- Baumeister, Christiane and Lutz Kilian (2012). “Real-time forecasts of the real price of oil.” *Journal of Business & Economic Statistics* 30.2, pp. 326–336.
- Baumeister, Christiane and Lutz Kilian (2014). “A general approach to recovering market expectations from futures prices with an application to crude oil.” *CFS Working Paper*.
- Baumeister, Christiane and Gert Peersman (2013). “The role of time-varying price elasticities in accounting for volatility changes in the crude oil market.” *Journal of Applied Econometrics* 28.7, pp. 1087–1109.
- Bernanke, Ben S (1983). “Irreversibility, uncertainty, and cyclical investment.” *The Quarterly Journal of Economics* 98.1, pp. 85–106.
- Bessembinder, Hendrik (1992). “Systematic risk, hedging pressure, and risk premiums in futures markets.” *Review of Financial Studies* 5.4, pp. 637–667.
- Bessembinder, Hendrik and Kalok Chan (1992). “Time-varying risk premia and forecastable returns in futures markets.” *Journal of Financial Economics* 32.2, pp. 169–193.
- Bloom, Nicholas (2009). “The impact of uncertainty shocks.” *econometrica* 77.3, pp. 623–685.
- Bollerslev, Tim, George Tauchen, and Hao Zhou (2009). “Expected stock returns and variance risk premia.” *Review of Financial Studies* 22.11, pp. 4463–4492.
- Bollerslev, Tim and Viktor Todorov (2011a). “Estimation of jump tails.” *Econometrica* 79.6, pp. 1727–1783.
- Bollerslev, Tim and Viktor Todorov (2011b). “Tails, fears, and risk premia.” *The Journal of Finance* 66.6, pp. 2165–2211.
- Bollerslev, Tim and Viktor Todorov (2013). “Time-Varying Jump Tails.” *Journal of Econometrics, forthcoming*.
- Bollerslev, Tim, Viktor Todorov, and Lai Xu (2014). “Tail Risk Premia and Return Predictability.” *Working Paper*.
- Carr, Peter and Liuren Wu (2009). “Variance risk premiums.” *Review of Financial Studies* 22.3, pp. 1311–1341.

- Christoffersen, Peter and Xuhui Nick Pan (2014). "Oil Risk Exposure and Expected Stock Returns." *SSRN Working Paper*.
- Doran, James S and Ehud I Ronn (2008). "Computing the market price of volatility risk in the energy commodity markets." *Journal of Banking & Finance* 32.12, pp. 2541–2552.
- Elder, John and Apostolos Serletis (2010). "Oil price uncertainty." *Journal of Money, Credit and Banking* 42.6, pp. 1137–1159.
- Erb, Claude B and Campbell R Harvey (2006). "The strategic and tactical value of commodity futures." *Financial Analysts Journal*, pp. 69–97.
- Fama, Eugene F and Kenneth R French (1987). "Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage." *Journal of Business*, pp. 55–73.
- Gorton, G.B. and K.G. Rouwenhorst (2004). "Facts and fantasies about commodity futures." *NBER Working Paper*.
- Gospodinov, Nikolay and Serena Ng (2013). "Commodity prices, convenience yields, and inflation." *The Review of Economics and Statistics* 95.1, pp. 206–219.
- Hamilton, James D and Jing Cynthia Wu (2014). "Risk premia in crude oil futures prices." *Journal of International Money and Finance* 42, pp. 9–37.
- Hong, Harrison and Motohiro Yogo (2012). "What does futures market interest tell us about the macroeconomy and asset prices?" *Journal of Financial Economics* 105.3, pp. 473–490.
- Jo, Soojin (2014). "The Effects of Oil Price Uncertainty on Global Real Economic Activity." *Journal of Money, Credit and Banking* 46.6, pp. 1113–1135.
- Kaldor, Nicholas (1939). "Speculation and economic stability." *The Review of Economic Studies* 7.1, pp. 1–27.
- Kang, Sang Baum and Xuhui Nick Pan (2013). "Commodity variance risk premia and expected futures returns: evidence from the crude oil market." *SSRN Working Paper*.
- Keynes, J.M. (1976). *A treatise on money*. New York; AMS Press.
- Kilian, Lutz (2009). "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market." *The American Economic Review*, pp. 1053–1069.
- Kilian, Lutz and Daniel P Murphy (2014). "The role of inventories and speculative trading in the global market for crude oil." *Journal of Applied Econometrics* 29.3, pp. 454–478.
- Mancini, Cecilia (2001). "Disentangling the jumps of the diffusion in a geometric jumping Brownian motion." *Giornale dell'Istituto Italiano degli Attuari* 64, pp. 19–47.
- Milonas, Nikolaos T and Stavros B Thomadakis (1997). "Convenience yields as call options: an empirical analysis." *Journal of Futures Markets* 17.1, pp. 1–15.
- Singleton, Kenneth J (2013). "Investor flows and the 2008 boom/bust in oil prices." *Management Science* 60.2, pp. 300–318.

- Szymanowska, Marta et al. (2014). “An anatomy of commodity futures risk premia.” *The Journal of Finance* 69.1, pp. 453–482.
- Tang, Ke and Wei Xiong (2012). “Index investment and the financialization of commodities.” *Financial Analysts Journal* 68.5, pp. 54–74.
- Trolle, Anders B and Eduardo S Schwartz (2010). “Variance risk premia in energy commodities.” *The Journal of Derivatives* 17.3, pp. 15–32.
- Wachter, Jessica A (2013). “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?” *The Journal of Finance* 68.3, pp. 987–1035.
- Working, Holbrook (1949). “The theory of price of storage.” *The American Economic Review*, pp. 1254–1262.

Chapter 2

The Role of Commodity Index Investment in Commodity and Asset Price Comovement

2.1 Introduction

Over the last decade, large financial entities such as pension, mutual and hedge funds have assumed an increasingly important roles in commodity futures markets, and the connection between commodity price movements and financial investor participation has spurred numerous debates among academics and policy makers (Masters 2008; Hamilton 2009; Irwin, Sanders, and Merrin 2009; Fattouh, Kilian, and Mahadeva 2013). Besides price volatility, another recent feature has drawn attention: the cross-market correlation between stock and commodity price changes has grown significantly. Figure (2.1) in Appendix B displays the rolling window estimation of the S&P 500 returns and two major commodity index returns, and indicates the increased comovement, which seems to start prior to 2008 but is most noticeable after the outbreak of the financial crises. Interestingly, this correlation has not declined over the recent years. Similarly, UNCTAD (2011), Tang and Xiong (2010), Maystre and Bicchetti (2012) and Lombardi and Ravazzolo (2013) document steady increases in the correlations between equities and commodities at a variety of frequencies, including intrasecond intervals, since 2004, with a particularly sharp increase after the collapse of Lehman Brother in September 2008. Several of these authors suggest that this change is related to the “financialization” of commodity markets, and that financial investors that are active in both markets play a particular role in transmitting shocks from equity across

different commodity markets during times of crises.¹ Still, their analysis leaves unanswered questions regarding the future persistence of the co-movements, and the particular role of economic fundamentals and trading strategies by various market participants in explaining this comovement.

The goal of this paper is to study the effects of the financialization of commodity future markets on the joint distribution of commodity and equity prices. Along this line, this paper makes two contributions to the existing literature: First, I formalize suggestive pieces of argumentation put forward by several of the aforementioned authors and derive a stylized model depicting a mechanism through which financial investment in “paper commodities” impacts the joint distribution of commodity and stock prices. Second, using data from the US Commodity Futures Trading Commission (CFTC) on futures positions held by financial traders in agricultural commodity markets, I provide empirical evidence for spillover shocks from equity to commodity markets through commodity index trading, a passive investment style that holds and rolls over long futures contracts according to fixed rules and that has experienced increased popularity over the recent decade.

The mechanism presented in the theoretical section relies on the importance of financial investors for time varying risk premiums determining futures prices. The “Theory of Normal Backwardation”, going back to Keynes (1976), proposes that commodity producers and inventory holders that are hedging their output through the futures market must pay a risk premium to speculators on the long side of the contracts in order to compensate them for taking on non-diversifiable price risks. This implies that the futures price tends to be below the corresponding expected spot price. The “Theory of Normal Backwardation” has found empirical support by Carter, Rausser, and Schmitz (1983), Chang (1985) and Bessembinder (1992), among others. When a sufficiently large number of financial investors takes on long positions in commodity futures in order to diversify their portfolio, or to hedge inflation risks, this long hedging pressure can induce a diminishing effect on the risk premium.²

If diversifying investors represent a large fraction of futures market participants, their portfolio choices establish direct links between futures risk premiums and other asset market conditions. Hence, their demand for futures will generally vary with their perceived risks of the outside portfolio. Importantly, the equilibrium futures price is a major determinant for the optimal storage decision of risk averse inventory holders, who are naturally short on commodity futures (Kawai 1983). A reduction of the financial demand of long futures will increase the risk premium and result in lower equilibrium prices and therefore increase the

¹Their analysis also indicates that the changes in correlation were structural and not a pure artifact of heteroscedasticity that influences changes in correlations, as warned by Forbes and Rigobon (2002).

²For example, Hamilton and Wu (2014) documents empirical evidence for significant changes in the level and volatility of the risk premium in the oil futures market that coincide with an increasing importance of commodity index investment in this market.

cost of hedging for inventory holders, who in turn will reduce storage activity and increase spot market supply. In order to understand the implications of the model, consider the situation where the diversifying investors are affected by adverse shocks in the central market, i.e. the equity market. Such a shock will affect the wealth of their portfolio, decrease the demand for futures, and decrease the equilibrium futures price. Inventory holders will react by decreasing inventory, current supply rises and spot prices fall, even if demand is not affected by shocks. This mechanism relies on limits to perfect arbitrage in futures markets, which have been well documented in recent studies, and has played a particularly important role during the recent financial crisis (Hong and Yogo 2012; Acharya, Lochstoer, and Ramadorai 2013).

The theoretical model indicates that the effect of shocks to the central asset market on commodity prices through the depicted mechanism is particularly important when diversifying investors play a large role in the futures market. In fact, the last decade has seen a phenomenal increase in open interests of commodity futures and options, as depicted in figure (2.2) in Appendix B for the case of oil. The figure also suggests that non-commercial traders, i.e. traders who are not exposed to the physical commodities, were the main driver of this change. As also noticed by Hamilton and Wu (2014), an important part of this trend was caused by increased participation of commodity index traders (CITs), financial investors that invest in so-called “index funds”. These funds intent to replicate movements of the spot price of the commodity by taking on long position in a near futures contract and rolling over the contract before its maturity in order to take a new long position in the next contract. CITs are typically not interested in handling the commodities physically, which gives index investment a speculative component. Irwin and Sanders (2011) indicate that the important motives for index investment are portfolio diversification, along with hedging of inflation risks and absolute returns. The portfolio diversification motive was popularized by studies by investors and academics showing that commodity futures that are rolled over on a regular basis could earn yields similar to those of stock returns, while their (historical) negative correlation with stock market returns provided additional benefits through strategic portfolio diversification (Greer 2000; Gorton and Rouwenhorst 2004; Ambroseno and Simpson 2008). Relaxation of regulation and the quest for returns after the dot-com bubble made commodity derivatives, most of all individual futures and futures indexes, an increasingly popular asset class. Parsons (2010), Vansteenkiste (2011) and Stoll and Whaley (2010), among others, document the increased importance of CITs and other derivative traders. Liu, Qiu, and Tang (2011) shows that investment in commodity futures indexes in the US grew from about \$ 15 billion in 2003 to over \$ 200 billion in 2008, while Büyüksahin and Robe (2014) suggests that the trading volume of the commodity futures market now vastly outsizes the market for physical commodities in terms of of traded contracts and their

value. Commodity index investing is often described as a “passive” investment strategy due to the utilization of long-only and fully collateralized futures, and predetermined investment rules that are invariant to fundamental market changes.

Using public data from the Commodity Index Trader Supplement of the U.S Commodity Trading Futures Commission (CFTC), my analysis focuses on the role of CITs in 12 agricultural markets for which the data is published. Under the CFTC classification, diversified investors such as index funds, pension funds, and mutual funds represent a major fraction of CITs, making this group of traders a particularly suitable candidate for the empirical implications of my model. As seen in table (2.1) in Appendix C, index investments made up to 40 % of the open interest in agricultural markets during the 2006 - 2011 period. My empirical analysis focuses on the relationship between asset market returns as measured by returns to the S&P 500, and position changes of CITs in these markets. I show that CITs react consistently by increasing positions after positive stock market shocks. Consistent with the implications of my theoretical model, the effect is particularly strong during the recent financial crisis, and statistically significant. The effects are robust after controlling for market specific and a variety of macroeconomic factors, and the empirical results also indicate that these position changes are forecastable. This indicates that CITs might have played an important part in transmitting stock market shocks to a variety of commodity markets. Moreover, when the reaction of CITs to stock market changes is public knowledge, commodity futures and spot prices might adjust fast due to (even limited) arbitrage; hence passive financial investment might help to explain part of the increased comovement between stock and commodity prices.

The remainder of this paper evolves as follows: section 2 reviews the related literature; section 3 presents an equilibrium model of a commodity futures market that relates increased financial investment to the joint distribution of commodity and equity prices. Sections 4 presents the empirical evidence for the cross-market linkages through CITs, and section 5 concludes.

2.2 Related Literature

As noticed by Hirshleifer (1990), derivative speculation can only have an effect on equilibrium cash prices and quantities when there are financial frictions. Shleifer (1986) and De Long et al. (1990) provide early evidence for downward-sloping demand in a variety of financial markets, and the arguments have been extended to the futures markets in more recent studies (see e.g. Danielsson, Shin, and Zigrand 2009; Gromb and Vayanos 2010, and the literature cited therein). Etula (2013) presents an equilibrium model in which limits to arbitrage arise through Value-at-Risk constraints of risk neutral broker. His results indicate that the broker

balance sheets' exposure to risk forecast commodity future returns in the pre-2007 period. In my model, investor's arbitrage constraints are introduced in a similar fashion as in Acharya, Lochstoer, and Ramadorai (2013), whose model setup I follow closely. Consistent with the theoretical model, these authors show that producers' risk aversion as measured by producers' default risk forecasts futures returns, spot prices, and inventories in oil and gas market data from 1980-2006, and the component of the commodity futures risk premium associated with producer hedging demand rises when speculative activity reduces. An important extension of my work is that I explicitly take into account cross-market linkages, and show how these links affect the joint distribution of commodity and asset prices. My work is also closely related to Liu, Qiu, and Tang (2011), who use a model of heterogeneous financial investors that hold a diversified portfolio to show how the demand for financial assets affects futures markets and commodity cash prices. Their model differs from mine as it focuses on the (exogenous) convenience yield, and does not model the demand for physical commodities explicitly.

A variety of studies have arose in the connection to the effects of commodity index investors and other financial players. The main focus of existing studies relates to the impacts the futures risk premium, which is also an important component of my theoretical model. However, my model includes also the the relationship futures and spot returns, which constitutes an extension to these studies. Singleton (2013) finds that a variety of measures, including a 13-week change in index-fund holdings imputed from the Supplemental Commitments of Traders report, could help predict weekly and monthly returns on crude oil futures contracts over September 2006 to January 2010. His results are consistent with limited arbitrage that arises through the interplay between imperfect information about real economic activity of commercials and speculators.³ Sanders and Irwin (2011b) finds no discernible effect of weekly net positions of swap dealers on the forecastability of returns on 14 different commodity futures contracts over 2006-2009. Sanders and Irwin (2011a) investigates grain commodities with proprietary CFTC data on commodity index traders and finds some predictability for soybeans, but not for other grains. Stoll and Whaley (2010) use the public Supplement for Index Traders and document small effects from changes in the long positions of commodity index traders for weekly commodity returns predictability during 2006-2009. Büyüksahin and Robe (2014), using a non-public dataset, do not find significant price impacts from index investors, but show that the correlations between the returns on investable commodity and equity indices increase amid greater participation by hedge funds. They also argue that hedge funds activity could strengthen cross market linkages, and that "their exit from satellite markets (such as emerging markets or commodity

³However, these results are not robust to extending the sample period by only two more years (Hamilton and Wu 2015).

markets) after a major shock in a central asset market (such as the U.S. equity market) could in theory bring about cross-market contagion” (p.2, Büyüksahin and Robe 2014). Lombardi and Van Robays (2011) argue that financial agents such as index investors might not base their futures demand on their expectation about future fundamentals, but rather for reasons of broader portfolio diversification. An index fund, for example, might want to offer their customers an exposure to oil price risk, and place themselves in a long position of the futures market irrespective of their current expectations. The prevalence of uncertainty among too many traders can overwhelm the arbitrageurs’ willingness to take on risky positions and impact the future price. In a more recent study, Buyuksahin and Robe (2011) provide empirical evidence that the correlation between stock market and energy futures returns has increased due to financial investment in “paper” commodities. In my theoretical model, I formalize these ideas by looking at the hedging and output decisions of risk averse firms. From standard hedging literature the firms’ decision about future commodity output is closely related to the futures price (Kawai 1983; Danthine 1978). By modeling an speculator’s investment strategy, I combine Büyüksahin and Robe (2014)’s and Lombardi and Van Robays (2011)’s argumentation with the firm’s output decision and show that under certain assumptions, futures market speculation will change current and future spot prices based to the non-speculation benchmark. Moreover, when commodity speculation is part of a broader portfolio strategy, investment decisions can link commodity and traditional asset markets systematically. My empirical results are also nicely consistent with those of Gao and Suess (2011), who suggest that their indicator of investor sentiments, that includes measures of stock market volatility, among others, is able to forecast commodity future returns.

Since the first draft of this study in the beginning of 2012, several studies have investigated the relationship between index investment and commodity future returns. Hamilton and Wu (2015) find that CIT positions in agricultural futures markets do not help to predict futures returns for these commodities, a results that casts some doubts on the mechanism portrait in the theoretical section. These results stand in contrast to the suggestion by Hamilton and Wu (2014) who document diminishing but increasingly volatile futures risk premia in the crude oil futures market, a finding “consistent with the claim that index-fund investing has become more important relative to commerical hedging in determining the structure of crude oil futures risk premia over time.” While an implication the theoretical model presented here is that the (relative) number of positions, and not the notional exposure used by Hamilton and Wu (2015) should be a predicting variable, these results warrant further investigation. Further, Cheng, Kirilenko, and Xiong (2014), based on non-public data, show that “financial traders reduced their net long positions during the crisis in response to market distress”. Their empirical framework is similar to the one proposed in this paper; instead of using stock market returns in order to predict position changes of different trader groups,

these authors use volatility as a proxy for market distress.

2.3 A Model of Commodity Futures Market with Financial Investment

I present a two-period model of commodity spot and futures price determination that includes optimal inventory management and hedging demand, similar to the models of Kawai (1983) and Acharya, Lochstoer, and Ramadorai (2013). Their setup features (partial) output hedging of commodity producers and speculators that are compensated for taking on commodity price risk. The novelty of the model is the consideration of a second type of speculators that invest in commodity futures in order to diversify their portfolio, and their effect on the equilibrium outcomes. These speculators can be thought of large pension or mutual funds that have entered commodity index investment over the last decade.

Commodity demand and asset prices

In my model of the economy I consider a two periods, denoted by $t = 0$ and $t = 1$, and three markets: a commodity spot market for each period, and a futures market that opens in period zero. Consumers' behavior is captured by the inverse demand function for commodities:

$$S_t = \omega \frac{A_t}{Q_t}, \quad (2.1)$$

where S_t is the commodity spot price in period t , Q_t is the consumption of the commodity good and ω is a preference parameter. A_t is exogenous and stochastic, and the shocks to A_t summarize the various factors that affect the demand for commodities (e.g. technological changes in production). In particular, I assume that A_t is composed of a deterministic part, denoted \bar{A} , and a stochastic shock, $\epsilon_{D,t}$:

$$A_t = \bar{A} + \epsilon_{D,t}. \quad (2.2)$$

A further source of randomness in this economy are asset prices (also referred to as stock prices), denoted P_t , that are assumed to follow a random walk:

$$P_t = P_{t-1} + \epsilon_{P,t}. \quad (2.3)$$

I assume the existence of a joint distribution for the random innovations $\epsilon_{D,t}$ and $\epsilon_{P,t}$ which are both mean zero and a generic covariance matrix that defines σ_D^2 , the variance

of the shocks to the demand curve, σ_P^2 , the variance of asset prices and the covariance of the shocks $\sigma_{D,P}$. This distribution can be viewed as a reduced form expression of an underlying structural shock (e.g. productivity) that affects asset prices and commodity demand simultaneously.

In the market for commodity futures, agents can trade an unrestricted quantity x of commodity futures in period 0. Futures are contracts that specify a reference price F , which is then compared to the spot price in period 1, S_1 . At maturity, the holder of a long (short) contract, characterised by $x > 0$ ($x < 0$) is then paid the difference between the contracted futures price and the actual spot price, $S_1 - F$ ($F - S_1$). Initially, future trading is costless, but the contracts have to be honoured in period 1.

Producers

Producing firms are competitive with a mass normalized to N_P . Output per individual firm in both periods is fixed and given by g_0 and g_1 , respectively, but in period 0, managers can decide to store a certain amount i .⁴ I assume that storage is costless. Then, using capital letters to denote aggregate quantities, aggregate supply in period 0 is given by $G_0 - I$, while aggregate supply in period 1 is $G_1 + I$. Additionally, the managers can also agree on a number h_p of futures contracts, that specify the delivery/acceptance of the relevant quantity of commodities in period 1 for a price F that is agreed on in period 0 (if h_p is negative, the firm promises to deliver, i.e. it goes short; vice versa if h_p is positive, the firm goes long). Hence the firm's expected profits are given by:

$$\Pi_P = S_0(g_0 - i) + S_1(g_1 + i) + h_p(S_1 - F). \quad (2.4)$$

When the managers are risk averse with a mean-variance utility function, the maximization of their wealth is equal to the maximization of $E(\Pi_P) - \frac{\gamma_P}{2} Var(\Pi_P)$, where γ_P is the coefficient of constant risk aversion and Var refers to the conditional variance based on period 0 information. In the absence of stock outs ($i > 0$), from the first order conditions with respect to the inventory holdings and the futures contracts, we get:⁵

$$i = -g_1 - h_p + \frac{E_0(S_1 - F)}{\gamma_P \sigma^2} \quad (2.5)$$

$$h_p = -g_1 - i + \frac{E_0(S_1 - S_0)}{\gamma_P \sigma^2}, \quad (2.6)$$

⁴Acharya, Lochstoer, and Ramadorai (2013) argue that supply might be fixed in the short run because it is very expensive to adjust production.

⁵Note that I do not take into account the possibility of stock-outs in my model. A suitable extension would related the model to the "Theory of Storage", as noted by Acharya, Lochstoer, and Ramadorai (2013).

where σ^2 is the variance of the spot price in period 1 given $t = 0$ information. Equation (2.5) indicates that inventory is positively related to short hedging, $-h_p$ and decreasing in the next periods' output g_1 . The term $\frac{E(S_1 - F)}{\gamma_P \sigma^2}$ represents a speculative component: when the risk premium, $E(S_1 - F)$, is large, inventory holders are more likely to increase unhedged inventory. Likewise equation (2.6) indicates that short hedging is equal to the quantity sold in period 1, and a speculative component that depends on the risk premium. We also can see that $F = S_0$ in equilibrium. Intuitively, both the current spot price and the futures price provide the inventory holder with riskless profits, and riskless arbitrage would be possible if the equality does not hold. This also implies that changes in the futures price will be reflected by simultaneous changes in the first periods spot price.

Traditional Speculators

The risk averseness of producers induces activities by arbitrage speculators who assume open positions in futures and thus provide the necessary liquidity in the futures market. Without loss of generality, their mass is normalized to one. Following Danielsson, Shin, and Zigrand (2009) and Etula (2013), investors are risk neutral, but face Value-at-Risk and/or margin constraints. In the case of the traditional speculators, such constraints can be motivated by limited collateral, that is necessary to make their contracting credible. As in Acharya, Lochstoer, and Ramadorai (2013), I use a reduced form approach to model the constraints and assume that the constraint is proportional to the variance of the speculator's position h_A . When speculators act competitively and maximize the expected income in period 1 subject to the constraint, the objective function to be maximized is

$$h_A E_0(S_1 - F) - \frac{\gamma_A}{2} \text{Var}_0[h_A(S_1 - F)], \quad (2.7)$$

where h_A is the number of contracted futures and γ_A indicates the severity of the margin constraint. The optimal number of contracts is then given by

$$h_A = \frac{E_0(S_1 - F)}{\gamma_A \sigma^2}, \quad (2.8)$$

where σ^2 is the conditional variance of the spot price in period 1. Equation (2.8) indicates that in the case of normal backwardation, i.e. when $F < E_0(S_1)$, the speculators assume long positions. The optimal amount of contracts is increasing in the risk premium $E(S_1 - F)$, and negatively related to the severity of the arbitrage constraint and the volatility of the premium.

Diversifying Speculators

Diversifying speculators are agents with a fixed investment in the asset market, and who use the commodity future market as an additional channel of speculation and portfolio diversification. They act competitively with a mass normalized to N_D , which is given exogenously.⁶ I assume that diversifiers are endowed with a number of long contracts in the asset market, which can be thought of as a bond/stock portfolio. Further, diversifiers are born with a fraction N of the long asset, such that their initial wealth W_0 is given by $W_0 = NP_0$. Similar to the case of the traditional speculators, diversifying speculators constraints are modeled in reduced form. Thus they maximize $E_0(W_1) - \frac{\gamma_D}{2} Var_0(W_1)$, where $W_1 = N \cdot P_1 + h_D(S_1 - F)$, with respect to h_D , the number of commodity futures contracts. The FOC is

$$h_D = \frac{E(S_1 - F)}{\gamma_D \sigma^2} - \frac{W_0 \cdot Cov_0(P_1, S_1)}{\sigma^2}, \quad (2.9)$$

where γ_D is the severity of the D-speculators Value-at-Risk constraint, and $Cov_0(P_1, S_1)$ is the conditional covariance between period 1 asset and commodity prices based on $t = 0$ information. The first term on the right hand side of equation (2.9) is the speculative component, with a similar interpretation as for the traditional speculators. The second term constitutes the hedging demand of the diversifiers: when the return to futures covaries negatively with the remaining asset returns, diversifying speculators will increase their exposure in long commodity futures. In the following, I will maintain the assumption that this is the case. Note that this requires that $\sigma_{D,P} < 0$, since $Cov_0(P_1, S_1) = \frac{\omega}{g_1 + i} \sigma_{D,P}$. There are two important reasons that justify this assumption. The first is through historical observation: the most significant changes in the futures markets have been the massive increase in index investment, a passive long only strategy that was popularized through large financial institutions as a portfolio diversifying strategy (Irwin and Sanders 2011). The second reason is that long futures are important instruments to hedge against non-diversifiable risks, in particular expected and unexpected inflation (see e.g. Boons, De Roon, and Szymanowska 2011, and the literature referred to therein). For expositional purposes, it is thus convenient to assume that $\sigma_{D,P} < 0$, which has a similar effect on the portfolio choice and does not change the main results of this paper. Also worth pointing out is that the above assumptions imply that the hedging component is proportional to the value of the stock market portfolio $NP_0 = W_0$.

⁶This mimics the increase in CIT positions over the last decade, which, as previously argued, was to a large extent exogenous to current market conditions.

Equilibrium and Comparative Statics

The rational expectations equilibrium is given by a set of prices S_0, S_1, F and expectations, such that given these, commodity spot market and futures markets clear:

$$G_0 - I = Q_0 \quad (2.10)$$

$$G_1 + I = Q_1 \quad (2.11)$$

$$h_P + h_A + h_D = 0. \quad (2.12)$$

$$(2.13)$$

Rational expectations imply that $E_0(A_1) = \bar{A}$ and $\sigma^2 = (\frac{\omega}{g_1+i})^2 \sigma_D^2$, and $Cov_0(P_1, S_1) = \frac{\omega}{g_1+i} \sigma_{D,P}$. From equations (2.5) and (2.6), we have that $S_0 = F$, and, by clearing for futures market:

$$\begin{aligned} & -i \cdot N_P - g_1 \cdot N_P + N_P \cdot \frac{E(S_1 - S_0)}{\gamma_P \sigma^2} + \frac{E(S_1 - S_0)}{\gamma_A \sigma^2} + \\ & N_D \cdot \left(\frac{E(S_1 - S_0)}{\gamma_D \sigma^2} - \frac{W_0 \cdot Cov_0(P_1, S_1)}{\sigma^2} \right) = 0. \end{aligned} \quad (2.14)$$

Equation (2.14) defines an implicit solution to the equilibrium amount of storage, since the conditional moments are non-linear functions of i . Under the maintained assumption that $\sigma_{D,P} < 0$, the following Propositions are shown in the Appendix A via the implicit function theorem.

Proposition 1 *Under normal backwardation, the risk premium $E(S_1 - F)$ is decreasing in the number of diversifying speculators N_D .*

The intuition behind the proposition is straightforward. If diversifying speculators take on long positions, the demand for long futures raises with the number of speculators and puts upward pressure on the equilibrium futures price. This reduces hedging costs for producers, who increase inventory accordingly. Supply in $t = 1$ rises and the expected spot price fall. The result is consistent with empirical evidence provided by Hamilton and Wu (2014) and Brunetti and Reiffen (2014), which indicates that financial investors had an impact on decline in the risk premiums in various commodity markets during the recent years.

Proposition 2 *In the presence of diversifying speculators, positive (negative) asset market shocks have an increasing (decreasing) effect on the commodity spot price.*

This effect arises through a change in the risk premium. A positive shock that increases the value of the diversifying speculator's portfolio also raises his long hedging demand

through equation (2.9). This reduces the risk premium and increases the equilibrium futures price. Because hedging future output becomes less costly for producers, a larger fraction of the commodity is sold in the future. The current spot price increases along with storage.

Proposition 3 *The size of the effect of asset market shocks on commodity spot prices is increasing in the constraints of the traditional speculators.*

The intuition behind this result is that *ceteris paribus*, an increase in the traditional speculators' potential to arbitrage reduces the risk premium, i.e. the returns to long futures. A further reduction of expected returns through the effect described in Proposition 2 has a more than proportional negative influence on the expected returns and hedging potential of long futures. In other words, increased portfolio hedging demand by diversifiers cannot reduce the risk premium by too much if they do not want to forfeit the benefits from hedging. Since the risk premium is already relatively small, the additional demand from portfolio hedging is also smaller. Proposition 3 relates nicely to the empirical analysis in the following chapters, since it has been suggested that the the arbitrage potential was particularly low after the outbreak of the financial crisis, when many institutions were forced to reorganize their balance sheets and withdraw positions. A direct implication of the proposition is that the spillover effects from the stock to commodity markets should have had a stronger effect during this period.

2.4 The Empirical Effect of Stock Market Movements on CITs' Positions

Data Description

The empirical analysis focuses on the role of the CITs, which most closely resembles the diversifying speculators presented in the theoretical section. The data on positions in commodity futures markets can be obtained through the Commitments of Traders (COT) report of the US Commodity Futures Trading Commission (CFTC).⁷ The publicly available version of this report contains the aggregate futures positions of different groups of traders as of the Tuesday of each week. The CFTC collects this information via questionnaires and interviews for all traders meeting the reporting levels set by the Commission, where the non-reportable traders are small producers, hedgers or speculators, whose holdings fall below a determined threshold. The positions of reportable traders typically cover 70 to 90 percent of the total open interest for any particular commodity. An important breakdown

⁷See <http://www.cftc.gov/MarketReports/CommitmentsofTraders>.

of the positions of the various groups of traders has been provided retrospectively until 2006 through the Commodity Index Trader Supplement for 12 agricultural and livestock markets. In addition to the standard commercial (hedgers) and non-commercial (speculators) classification of the COT reports, the supplement reports the positions of commodity index traders (CIT). Index investors use two principal means to engage in index investment activity: direct investment in futures markets and indirect investment through over-the-counter (OTC) swap agreements with financial firms. Via personal interview and special questionnaires the CFTC computes daily information of the index activity of index funds, swap dealers, pension funds, hedge funds and mutual funds investments in exchange traded funds (ETFs), exchange traded notes (ETNs) and similar exchange-traded products that have a fiduciary or other obligation to track the value of a commodity or basket of commodities in an essentially passive manner. The remaining positions are classified as commercials (Producers/Merchants/Processors/Users) and non-commercials, without taking into account the CIT positions that are displayed separately. The Commodity Index Trader Supplement is the best publicly available data on CIT positions, but not without problems. Particularly non-commercials have incentives to misreport their true positions and activities due to tighter regulation for this group, and the CFTC classification is not uncontroversial (Stoll and Whaley 2010).

An overview over the 12 agricultural commodity markets considered in this study is given in table (2.1). These commodities can be grouped into grains (wheat, corn, soybeans, and soybean oil), fibers (cotton), livestock (live cattle, feeder cattle and lean hogs) and softs (cocoa, coffee, sugar). The cross-sectional variation of open interest is considerable, with an 2006- 2011 average from 1.7 million in the Corn market to an average of 35 thousand outstanding contracts for Feeder cattle.⁸ Likewise, the relative importance of CITs differs across markets. In most of the markets, CIT held an average of about 25% to 30% over the considered time period. Noticeable exceptions are cocoa (15%), Chicago wheat (42%) and Lean Hogs (41%). Figure (2.3) describes the evolution of CIT long positions and commodity prices in the different markets over time. While some CIT positions exhibit a slight upward trend in some markets, these increases are much less pronounced than in the pre-2006 period. Most noticeable is the simultaneous decline in CIT positions in the end of 2008 in all markets.

The other variables included in the regression contain the Baltic Dry Cargo Index, the secondary market T-Bill rate and lagged price changes. The Baltic Dry Cargo Index replicates current shipping freight costs, and is thus an important indicator of global economy activity and demand (Kilian 2009). The measure of the current interest rate is an impor-

⁸However, it has to be taken into account that the contract units are not easily comparable across commodity groups. Part of the literature suggests to weight the open interest with the dollar value of contracts, but since I am interested in relative position changes, this is not necessary here.

tant determinant of asset prices, and also connected to the financing and opportunity costs of storage. Lagged changes in the price of the commodity are included to account for the possibility that the CIT's strategy conditions on them.⁹ Due to lack of availability, the first nearby futures price is used to proxy the spot price of the commodity. As the futures price converges towards the spot price with approaching maturity, and the first contract is often used to settle nearby physical delivery, the approximation error should be small. All variables are adjusted to a weekly frequency and match the Tuesday's CFTC Report. All series are available from 03/Jan/2006 until 27/Dec/2011, comprising 313 observations over time.

Hypothesis and Empirical Framework

The hypothesis tested in this section provide empirical evidence for the mechanisms reflected in Proposition 2 as well as Proposition 3 of the theoretical section. Regarding Proposition 2, the key theoretical mechanism suggests that diversifying speculators, i.e. CITs, should react consistently to positive stock market shocks by increasing their long futures position. Also, I exploit the fact that during the financial crisis, traditional speculators were more constraint (Cheng, Kirilenko, and Xiong 2014) to provide evidence for Proposition 3. The implied effects on the risk premium and prices are only presented descriptively. One of the reasons is that when the fundamental correlation between stock and commodity markets is time-varying, these effects are empirically not identified. This is because stock market returns have demand effects through changes in wealth, and might also contain signals about future economic activity (Beaudry and Portier 2004), thereby affecting current demand or expectations. Speaking in terms of my model, suppose that period $t = 0$ asset price shocks have a permanent impact on the stochastic demand parameter A . In this case, the expected spot price in period 1 rises after a stock market shock, storers have incentives to hedge further output and diversifiers increase positions alongside. Moreover, this effect would induce an increased correlation between asset and commodity prices through reduced demand in period 1, which is observationally equivalent to a financial demand induced shock. By focusing on the reaction of different types of market participants this problem can be somewhat alleviated, as they help to discriminate between fundamental shocks, that affect asset prices and commodity demand simultaneously, and financial demand induced shocks. A crucial difference to discriminate between the the two mechanisms is the reaction of commercial speculators that are buyers of commodities and thus naturally long. While not considered explicitly in the model, the fundamental shock should lead to an increased demand of long futures of the latter, while the financial demand shock would should have a negligible or

⁹Futures prices for agricultural and livestock commodities are obtained from wikiposit.org. Sample data from the CME website was used to verify their correctness.

negative impact on their demand due to an increased cost of futures hedging.¹⁰

The baseline model is a fixed effects panel regression of the form:

$$\Delta Pos_t^{j,i} = \alpha^{j,i} + \sum_{k=0}^L \beta_j r_{t-k}^{S\&P} + Controls' \gamma + \epsilon_t^{i,j}, \quad (2.15)$$

where $\Delta Pos_t^{j,i}$, is the position change of trader group j in commodity market i , $r_t^{S\&P}$ is the return to the S&P 500 stock market index, and the controls include the percentage changes in the Baltic Dry Index and changes in the interest rate as common variables, as well as a set of market specific variables: lagged percentage changes in the spot price, in order to account for price induced position changes, lagged position changes of the specific trader group in order to account for momentum effects and a set of commodity specific month dummies in order to account for seasonal effects. The estimations are performed for each trader group j separately. The panel framework suggested by (2.15) is particularly relevant for the analysis of CITs, since they are typically diversified over a wide range of markets and the corresponding investment decisions are typically proportional according to fixed rules. Finally, a fixed effects regression is employed in order to account for market specific time invariant factors.

Regression results

Table (2.3) displays the results for the baseline specification for CIT position changes, which includes contemporaneous changes in macroeconomic variables on the right hand side. An overview over the variables used in the regressions is given in table (2.2). Contemporaneous market specific changes, in particular the current spot price returns are excluded, since my model strongly suggests that it is determined endogenously by changes in CIT positions. CIT positions are divided by a constant measuring the size of the market (average open interest over the 2006 - 2011 period), in order to account for the fact that the position changes will potentially be larger in a big market. Thus, the dependent variable, CIT long position variable is the same as displayed in figure (2.3), varying between 0.2 and 0.5 over markets and time. The interpretation is somewhat similar to percentage terms, but assuming that other traders' positions stay equal. The estimates for the most parsimonious model outlined previously are presented in column 2 of table (2.3). Effects from contemporaneous and lagged stock market returns are statistically significant at a 1% significance level (5% for the second lag). A one unit increase in contemporaneous stock returns rises the relative long positions by about 0.02 units, and lagged stock returns increase these positions by about 0.012 units. The second lag of the returns of the Baltic Dry Index also have statistically significant impacts,

¹⁰A similar argument is forwarded by Cheng, Kirilenko, and Xiong (2014).

along with percentage the lag 2 % changes of the interest rate, although the coefficient is substantially smaller. There seems to be some inertia in CIT position changes, with the lagged change in the CIT position influencing the current change significantly. However, the overall R^2 is very small with just over 0.07.

Column 2 in table (2.4), shows that the results also hold for the forecasting equation. Here, I leave out all contemporaneous variables in order to avoid a potential bias that might arise through endogeneity, as outlined in the previous section. The results are very similar to those of the baseline baseline and the impact is statistically very significant.

Robustness checks and Crisis Period

As robustness checks I include three further variables that might be important for the determination of CIT positions. The first variable is a measure of the exchange rate, which is an important for internationally traded commodities, and which also has shown increasing, albeit negative, comovement with commodity prices over the recent years (UNCTAD 2011). I control also for changes in the oil price, since oil is an important input for agricultural production and thus might induce supply shocks in these markets (see e.g. Baumeister and Kilian 2014). Finally, I also include a measure of the changes in the future curve by computing the difference of the spot and the 4th month futures contract. The reason is that the larger the market is in backwardation, the larger the roll yield for CITs becomes, which might induce additional speculative activity from this group of traders. The empirical results for the baseline and the forecasting model are depicted in the first columns of table (2.3) and (2.4). Of the additional variables, especially changes in the slope of the futures curve seems to be of importance. While the size of the coefficients for current and lagged stock prices decreases, they are still statistically significant at all conventional significance levels.

The third column in table (2.4) and (2.3) present the estimates from the post-crises period only. Again the effect of stock market changes is statistically significant, and the size of the coefficients increases substantially in comparison with the relevant column 1. This is in line with the hypothesis that stock market shocks had a particular strong influence in and after the crises period.

Reaction of Commercials

This section presents the responses of the other trading participants to stock market shocks. If the effect of stock market shocks on commodity prices would work primarily through the demand or expectation channel, we would expect merchants - hedgers classified as commercials such as airlines - to react systematically to these shocks as well. The results in table (2.5) provide no evidence for a systematic reaction of this group. If anything, the

contemporaneous regression indicate a negative relationship between stock market returns and positions. This supports my hypothesis that position changes that are attributable to stock market returns are significant for non-commercials, in particular CITs, but not to other groups.

2.5 Conclusion

This paper discusses the recent increase in asset and commodity price comovement and increased participation of financial investors in commodity future markets. I develop a theoretical model that shows demand for futures by financial investors can transmit stock market shocks into commodity prices via a time varying risk premium. Based on data from the Commodity Trading Futures Commission on futures position of different market participants, I introduce an empirical framework for empirical evaluation of my model. The results suggest that commodity index investors react systematically to negative stock market shocks by reducing their commodity risk exposure. This increases the risk premium for commercial producers, and thereby transmits into commodity prices. This mechanism is consistent with my theoretical model, and helps to explain the increased comovement of stock market and commodity prices.

2.A Appendix: Proof of Proposition 1 - 3

This sections sketches the proofs for the Propositions in the theoretical section.

Proof of Proposition 1

Plugging in the conditional moments, equation (2.14) can be written as

$$\begin{aligned}
& -iN_P(\gamma_A\gamma_P\gamma_D)\left(\frac{\omega}{g_1+i}\right)^2\sigma_D^2 - g_1N_P(\gamma_A\gamma_P\gamma_D)\left(\frac{\omega}{g_1+i}\right)^2\sigma_D^2 \\
& + \left(\frac{\omega\bar{A}}{g_1+i} - \frac{\omega A_0}{g_0-i}\right)\left(\frac{\gamma_D}{N_D}\gamma_A + \frac{\gamma_D}{N_D}\frac{\gamma_P}{N_P} + \frac{\gamma_P}{N_P}\gamma_A\right) \\
& - \frac{\gamma_P}{N_P}\gamma_D\gamma_A W_0\left(\frac{\omega}{g_1+i}\right)\sigma_{D,P} = 0 \equiv F(i, \Theta),
\end{aligned} \tag{2.16}$$

where Θ is a vector of parameters. The result follows from implicitly differentiating

$$\frac{\partial i}{\partial N_D} = -\frac{F(\partial\cdot)}{\partial N_D} / \frac{\partial F(\partial\cdot)}{\partial i} > 0 \tag{2.17}$$

where the inequality holds if $\left(\frac{\omega\bar{A}}{g_1+i} - \frac{\omega A_0}{g_0-i}\right)$, i.e. the market is in normal backwardation. Moreover, $E(S_1)$ is decreasing in i and $F = S_0$ is increasing in i . The result follows.

Proof of Proposition 2

I am comparing a situation with diversifying speculators to the situation without them. In the case no diversifying speculators are present, equation (2.14) collapses to

$$-i \cdot N_P - g_1 \cdot N_P + N_P \cdot \frac{E(S_1 - S_0)}{\gamma_P \sigma^2} + \frac{E(S_1 - S_0)}{\gamma_A \sigma^2}, \tag{2.18}$$

and asset market shocks have only an indirect effect on the spot market price through the covariance $\sigma_{D,P}$. If diversifying speculators are present, equation (2.16) applies. Moreover, the diversifiers' wealth $W_0 = N \cdot P_{-1} + \epsilon_{P,0}$ is increasing in asset market shocks. Implicitly differentiating

$$\frac{\partial i}{\partial W_0} = -\frac{F(\partial\cdot)}{\partial W_0} / \frac{\partial F(\partial\cdot)}{\partial i} > 0 \tag{2.19}$$

proves the result.

Proof of Proposition 3

The direct effect of asset market shocks can be calculated by the chain rule $\partial S_0 / \partial \epsilon_{P,0} = \partial S_0 / \partial i \cdot \partial i / \partial \epsilon_{P,0}$. Differentiating this expression with respect to γ_S proves the result.

2.B Appendix: Figures

Figure 2.1: This figure depicts the one year rolling window return correlations between the S&P 500 and two major commodity indexes based on daily data. Data source: Datastream

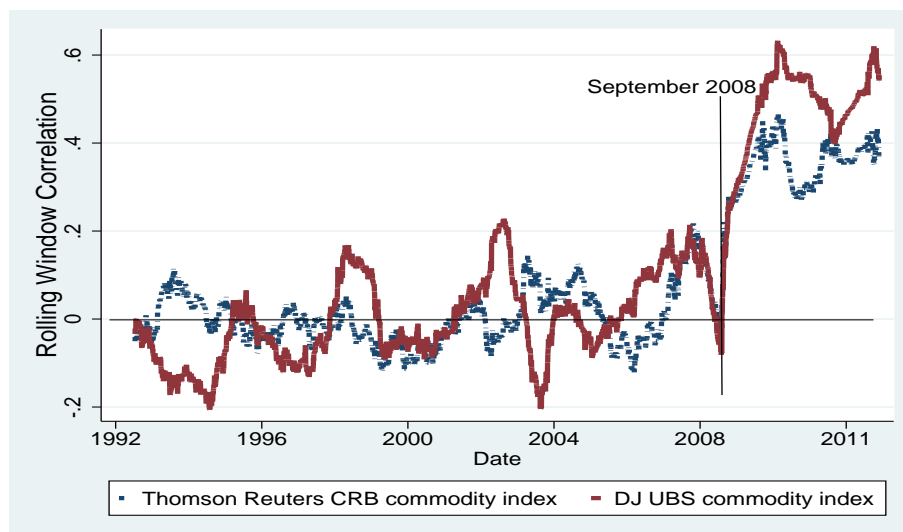


Figure 2.2: This figure depicts the positions of futures traders classified as commercials and the total open interest (futures and options) for the NYMEX Light Sweet Crude Oil relative to total open interest from 2000 to 2006. Commercial positions are computed as the sum of short, long and spread positions. Data source: CFTC

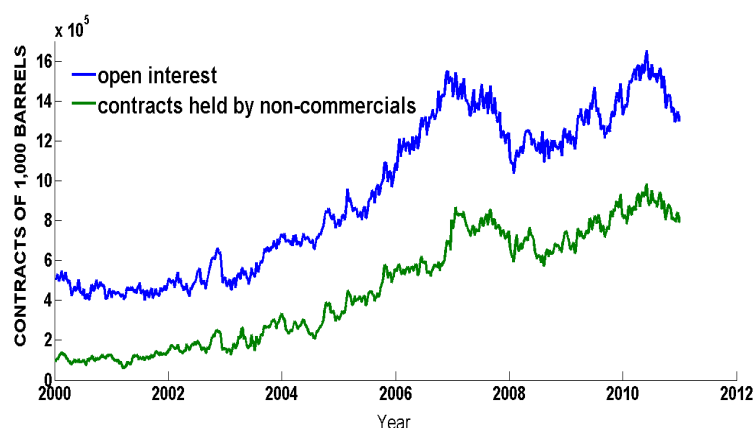
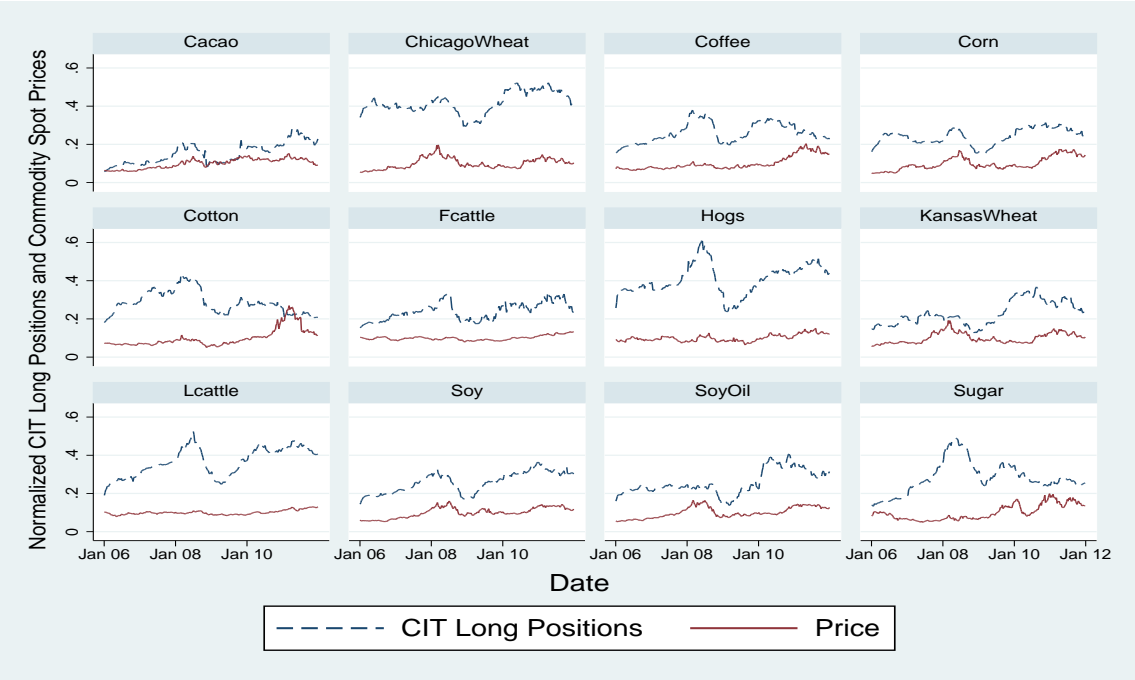


Figure 2.3: This figure depicts changes in CIT long futures positions and price changes of the different commodities. In order to make the different markets comparable, positions (prices) are normalized by dividing by through the average total open interest (10 times the average price) over the 2006 - 2011 period. Data source: CFTC



2.C Appendix: Tables

Table 2.1: Overview over agricultural commodities investigated in the empirical section.

Commodity	Category	Contract Unit	Aver. Price per Unit	Exchange	Aver. OI	% CIT of OI
Wheat Chicago	Grains	5,000 Bushels	620 \$	CBOT	502,481	42%
Wheat Kansas	Grains	5,000 Bushels	660 \$	KBOT	148,552	24%
Corn	Grains	5,000 Bushels	440 \$	CBOT	1,729,850	25%
Soybeans	Grains	5,000 Bushels	1,013 \$	CBOT	631,619	27%
Soybean Oil	Grains	60,000 Pounds	41 \$	CBOT	321,630	26%
Cotton No.2	Fibers	50,000 Pounds	0.77 \$	NYBOT/ICE	90,459	30%
Lean Hogs	Livestock	40,000 Pounds	70 \$	CME	214,830	41%
Live Cattle	Livestock	40,000 Pounds	95 \$	CME	326,902	37%
Feeder Cattle	Livestock	50,000 Pounds	110 \$	CME	35,113	25%
Cocoa	Softs	10 Metric Tons	2,432 \$	NYBOT/ICE	155,229	15%
Sugar	Softs	112,000 Pounds	0.17 \$	NYBOT/ICE	934,001	28%
Coffee	Softs	37,500 Pounds	14,986 \$	NYBOT/ICE	178,545	24%

Aver. Price per Unit = Average Price per Unit; Aver. OI= Average Open Interest in Contracts;
CIT Long = Pct. of Open Interest represented by CIT Long Contracts;
Averages computed through 03/01/2006 - 27/12/2011

Table 2.2: Variables used in the regression

Name	Underlying variable	Transformation
sp500ret	S&P 500	returns
CIT_long_relative	CIT Long positions	long positions / marketsize (average OI)
D.Com.Long_relative	Commercial Long positions	long positions / marketsize (average OI)
BDYret	Balt Dry Index	returns
fpriceret	Spot price	returns
tbsecret	3M T-bill	returns
usg6wic	Exchange rate	changes
oilret	Oil price	returns
relchange	Spot and Futures price	% change spot - % change futures

	(1)	(2)	(3)
	D.CIT_Long_relative	D.CIT_Long_relative	D.CIT_Long_relative
sp500ret	0.0162** (3.22)	0.0338*** (7.86)	0.0348*** (5.58)
L.sp500ret	0.0167** (4.04)	0.0241*** (5.79)	0.0254*** (6.01)
L2.sp500ret	0.0132* (2.88)	0.0147** (3.40)	0.0186* (2.38)
LD.CIT_Long_relative	0.109*** (4.69)	0.113*** (5.04)	0.0936** (3.12)
L2D.CIT_Long_relative	-0.0221 (-0.75)	-0.0218 (-0.76)	-0.0580 (-1.21)
BDYret	0.000980 (0.46)	0.00202 (0.98)	0.000938 (0.44)
L.BDYret	0.000314 (0.23)	0.000550 (0.45)	0.00205 (1.33)
L2.BDYret	0.00540** (4.34)	0.00658*** (5.11)	0.00704*** (6.13)
L.fpriceret	-0.00293 (-0.78)	0.00251 (0.76)	-0.00184 (-0.37)
L2.fpriceret	0.000770 (0.29)	0.00223 (0.88)	0.00494 (1.31)
tbsecret	-0.000604 (-1.65)	-0.000345 (-1.31)	-0.000510 (-1.20)
L.tbsecret	0.000202 (0.40)	0.000228 (0.54)	0.000205 (0.38)
L2.tbsecret	-0.00210* (-2.87)	-0.00210* (-2.83)	-0.00184* (-2.63)
D.usg6wic	-0.000329*** (-5.26)		0.000101 (0.75)
LD.usg6wic	-0.000327* (-2.33)		-0.000230 (-1.56)
L2D.usg6wic	-0.0000136 (-0.10)		0.0000320 (0.22)
oilret	0.0132*** (7.04)		0.00964** (4.29)
L.oilret	0.00509 (1.56)		0.000434 (0.10)
L2.oilret	0.00278 (1.43)		0.00133 (0.38)
L.relchange22	-0.0166** (-4.20)		-0.0180* (-2.87)
L2.relchange22	0.00196 (0.56)		0.00627 (1.14)
time fixed effects	yes	yes	yes
market specific			
seasonal dummies	yes	yes	yes
constant	yes	yes	yes
N	3720	3720	2088
R^2	0.087	0.073	0.107

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.3: Baseline estimation including contemporaneous variables. The dependent variable is (relative) CIT position changes. Variables are described in table (2.2). T-statistics based on clustered standard errors reported.

	(1)	(2)	(3)
	D.CIT_Long_relative	D.CIT_Long_relative	D.CIT_Long_relative
L.sp500ret	0.0151** (3.29)	0.0219*** (4.57)	0.0233*** (6.11)
L2.sp500ret	0.0145** (3.16)	0.0136** (3.14)	0.0149 (1.84)
LD.CIT_Long_relative	0.118*** (4.95)	0.119*** (5.21)	0.105** (3.37)
L2D.CIT_Long_relative	-0.0166 (-0.56)	-0.0166 (-0.58)	-0.0472 (-0.98)
L.BDYret	0.000801 (0.66)	0.00155 (1.37)	0.00249 (1.77)
L2.BDYret	0.00709*** (6.14)	0.00689*** (6.16)	0.00849*** (7.82)
L.fpriceret	-0.00302 (-0.83)	0.00270 (0.85)	-0.00203 (-0.44)
L2.fpriceret	0.00247 (0.91)	0.00163 (0.60)	0.00719 (1.84)
L.tbsecret	0.000323 (0.72)	0.000384 (0.97)	0.000551 (1.12)
L2.tbsecret	-0.00217* (-2.98)	-0.00223* (-2.99)	-0.00193* (-2.72)
LD.usg6wic	-0.000292 (-2.15)		-0.000238 (-1.72)
L2D.usg6wic	0.0000764 (0.58)		0.0000927 (0.67)
L.oilret	0.00441 (1.48)		-0.000410 (-0.10)
L2.oilret	0.000540 (0.29)		-0.00137 (-0.39)
L.relchange22	-0.0173** (-4.17)		-0.0183* (-2.83)
L2.relchange22	0.00322 (0.83)		0.00815 (1.26)
time fixed effects	yes	yes	yes
market specific			
seasonal dummies	yes	yes	yes
constant	yes	yes	yes
<i>N</i>	3720	3720	2088
<i>R</i> ²	0.064	0.061	0.081

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.4: Baseline forecasting regression. The dependent variable is (relative) CIT position changes. Variables are described in table (2.2). T-statistics based on clustered standard errors reported.

	(1)	(2)
	D.Com_Long_relative	D.Com_Long_relative
sp500ret	-0.0517*	
	(0.020)	
L.sp500ret,	-0.0026	0.0008
	(0.010)	(0.009)
L2.sp500ret,		0.0137
		(0.009)
L.Com_Long_relative,	0.1754***	0.1743***
	(0.026)	(0.026)
L2.Com_Long_relative,	-0.0502*	-0.0513*
	(0.021)	(0.021)
L.BDYret,	0.0011	0.0001
	(0.003)	(0.003)
L2.BDYret,	0.0057	0.0056
	(0.003)	(0.003)
L.fpriceret,	0.0086	0.0077
	(0.010)	(0.010)
L2.fpriceret,	0.0033	0.0020
	(0.011)	(0.010)
L.tbsecret,	-0.0012	-0.0013
	(0.001)	(0.001)
L2.tbsecret,	0.0009	0.0012
	(0.001)	(0.001)
L.relchange22,	-0.0031	-0.0039
	(0.016)	(0.016)
L2.relchange22,	0.0021	0.0007
	(0.020)	(0.019)
time fixed effects	yes	yes
market specific		
seasonal dummies	yes	yes
Constant	yes	yes
Observations	3,720	3,720
R-squared	0.04	0.04
Robust standard errors in parentheses		
*** p<0.01, ** p<0.02, * p<0.1		

Table 2.5: Regressions for commercial long positions. Column (1) contains the contemporaneous, column (2) the forecasting regression. The dependent variable is (relative) Commercial Long position changes. Variables are described in table (2.2). Clustered standard errors reported.

References

- Acharya, Viral V, Lars A Lochstoer, and Tarun Ramadorai (2013). “Limits to arbitrage and hedging: Evidence from commodity markets.” *Journal of Financial Economics* 109.2, pp. 441–465.
- Ambroseno, L. and J. Simpson (2008). “Commodities as a Strategic Component of an Investment Portfolio.” *Issues of Interest - Morgan Stanley Investment Management report*.
- Baumeister, Christiane and Lutz Kilian (2014). “Do oil price increases cause higher food prices?” *Economic Policy* 29.80, pp. 691–747.
- Beaudry, Paul and Franck Portier (2004). “Stock prices, news and economic fluctuations.”
- Bessembinder, H. (1992). “Systematic risk, hedging pressure, and risk premiums in futures markets.” *Review of Financial Studies* 5.4, pp. 637–667.
- Boons, M., F. De Roon, and M. Szymanowska (2011). “The stock market price of commodity risk.” *Working Paper*.
- Brunetti, Celso and David Reiffen (2014). “Commodity index trading and hedging costs.” *Journal of Financial Markets* 21, pp. 153–180.
- Buyuksahin, B. and M. Robe (2011). “Does ”paper oil” matter? Energy markets’ financialization and equity-commodity co-movements.” *Working paper*.
- Büyüksahin, Bahattin and Michel A Robe (2014). “Speculators, commodities and cross-market linkages.” *Journal of International Money and Finance* 42, pp. 38–70.
- Carter, C.A., G.C. Rausser, and A. Schmitz (1983). “Efficient asset portfolios and the theory of normal backwardation.” *The Journal of Political Economy* 91.2, pp. 319–331.
- Chang, E.C. (1985). “Returns to speculators and the theory of normal backwardation.” *Journal of Finance*, pp. 193–208.
- Cheng, Haw, Andrei Kirilenko, and Wei Xiong (2014). “Convective risk flows in commodity futures markets.” *Review of Finance*.
- Danielsson, J., H. Shin, and J.P. Zigrand (2009). “Risk appetite and endogenous risk.” *London School of Economics, working paper*.
- Danthine, J.P. (1978). “Information, Futures Prices, and Stabilizing Speculation.” *Journal of Economic Theory* 17, pp. 79–98.
- De Long, J.B. et al. (1990). “Noise Trader Risk in Financial Markets.” *The Journal of Political Economy* 98.4, pp. 703–738.
- Etula, Erkki (2013). “Broker-dealer risk appetite and commodity returns.” *Journal of Financial Econometrics*, nbs024.
- Fattouh, Bassam, Lutz Kilian, and Lavan Mahadeva (2013). “The Role of Speculation in Oil Markets: What Have We Learned So Far?” *The Energy Journal* 34.3.

- Forbes, K.J. and R. Rigobon (2002). “No contagion, only interdependence: measuring stock market comovements.” *The Journal of Finance* 57.5, pp. 2223–2261.
- Gao, L. and S. Suess (2011). “Market Sentiment in Commodity Futures Returns.” *Working paper*.
- Gorton, G.B. and K.G. Rouwenhorst (2004). “Facts and fantasies about commodity futures.” *NBER Working Paper*.
- Greer, R.J. (2000). “The nature of commodity index returns.” *The Journal of Alternative Investments* 3.1, pp. 45–52.
- Gromb, D. and D. Vayanos (2010). *Limits of arbitrage: The state of the theory*. Tech. rep. National Bureau of Economic Research.
- Hamilton, James D and Jing Cynthia Wu (2014). “Risk premia in crude oil futures prices.” *Journal of International Money and Finance* 42, pp. 9–37.
- Hamilton, James D and Jing Cynthia Wu (2015). “Effects of Index-Fund Investing on Commodity Futures Prices.” *International Economic Review* 56.1, pp. 187–205.
- Hamilton, J.D. (2009). “Causes and Consequences of the Oil Shock of 2007–08.” *Brookings Papers on Economic Activity* 1, pp. 215–261.
- Hirshleifer, D. (1990). “Hedging pressure and futures price movements in a general equilibrium model.” *Econometrica: Journal of the Econometric Society*, pp. 411–428.
- Hong, Harrison and Motohiro Yogo (2012). “What does futures market interest tell us about the macroeconomy and asset prices?” *Journal of Financial Economics* 105.3, pp. 473–490.
- Irwin, S.H. and D.R. Sanders (2011). “Index funds, financialization, and commodity futures markets.” *Applied Economic Perspectives and Policy* 33.1, pp. 1–31.
- Irwin, S.H., D.R. Sanders, and R.P. Merrin (2009). “Devil or angel? The role of speculation in the recent commodity price boom (and bust).” *Journal of Agricultural and Applied Economics* 41.2, pp. 377–91.
- Kawai, M. (1983). “Price volatility of storable commodities under rational expectations in spot and futures markets.” *International Economic Review* 24.2, pp. 435–459.
- Keynes, J.M. (1976). *A treatise on money*. New York; AMS Press.
- Kilian, Lutz (2009). “Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market.” *The American Economic Review*, pp. 1053–1069.
- Liu, P., Z. Qiu, and K. Tang (2011). “Financial-Demand Based Commodity Pricing.” *Working paper*.
- Lombardi, Marco Jacopo and Francesco Ravazzolo (2013). “On the correlation between commodity and equity returns: implications for portfolio allocation.”
- Lombardi, M.J. and I. Van Robays (2011). “Do Financial Investors destabilize the oil price?” *ECB Working Paper*.

- Masters, M.W. (2008). "Testimony before the Committee on Homeland Security and Governmental Affairs." *US Senate, Washington, May 20*.
- Maystre, Nicolas and David Bicchetti (2012). "The synchronized and long-lasting structural change on commodity markets: evidence from high frequency data." *Algorithmic Finance (2013)* 2, pp. 3–4.
- Parsons, J.E. (2010). "Black gold and fool's gold: speculation in the oil futures market." *Economía* 10.2, pp. 81–116.
- Sanders, D.R. and S.H. Irwin (2011a). "New evidence on the impact of index funds in US grain futures markets." *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*.
- Sanders, D.R. and S.H. Irwin (2011b). "The impact of index funds in commodity futures markets: A systems approach." *The Journal of Alternative Investments* 14.1, pp. 40–49.
- Shleifer, A. (1986). "Do demand curves for stocks slope down?" *Journal of Finance*, pp. 579–590.
- Singleton, Kenneth J (2013). "Investor flows and the 2008 boom/bust in oil prices." *Management Science* 60.2, pp. 300–318.
- Stoll, H.R. and R.E. Whaley (2010). "Commodity index investing and commodity futures prices." *Journal of Applied Finance* 20.1, pp. 7–46.
- Tang, K. and W. Xiong (2010). *Index investment and financialization of commodities*. Tech. rep. National Bureau of Economic Research.
- UNCTAD (2011). "Price Formation in Financialized Commodity Markets: The Role of Information." *United Nations Report*.
- Vansteenkiste, Isabel (2011). "What is driving oil futures prices? Fundamentals versus speculation." *ECB Working Paper*.

Chapter 3

Regional Decentralization and Fiscal Policy Effects - International and Intranational Evidence

Joint with Abián García Rodríguez

3.1 Introduction

This paper investigates the relationship between the government structure and the effectiveness of fiscal policy. The effects of fiscal policy on economic outcomes is of high importance for policy makers and academics alike; also, it is a striking feature that countries vary substantially in their organization of taxing and spending competences between different levels of government. We refer to the latter as the degree of “regionalization” and provide evidence that more decentralized countries, in other words, countries in which the regional governments have more fiscal competences, are associated with more effective fiscal policies.

The organization of taxing and spending competences between different levels of government has been shown to be important for government deficit reduction (Schaltegger and Feld 2009), the size of the government (Feld, Kirchgässner, and Schaltegger 2004), economic growth (Thießen 2003) and fiscal discipline (Rodden 2002). At the same time, the academic literature on the determinants and transmission mechanisms of fiscal policy is still inconclusive (see e.g. Perotti 2007; Corsetti, Meier, and Müller 2012). While determinants of fiscal policy effectiveness such as the level of development, exchange rate regime, openness to trade and public indebtedness have been addressed in the literature before (see e.g. Ilzetzi, Mendoza, and Végh 2013) this is to our best knowledge the first study to investigate the relationship between decentralization and fiscal policy effectiveness in terms of multipliers.

Using two distinct empirical approaches we show that fiscal decentralization is associated positively with the effectiveness of fiscal policy for stimulating economic growth. The first approach exploits cross country variation in the degree of relative spending competences between central and regional governments. In none of the various European countries considered in our study, government spending is completely carried out by the central government, and, in some countries, regional governments are responsible for more than half of the total government spending. In a first step, we construct a “decentralization index” based on the relative spending of the regional governments and classify countries accordingly. We then estimate the response of GDP to unexpected government spending using the SVAR approach suggested by Blanchard and Perotti (2002) for each country separately. The results indicate that a higher degree of decentralization in terms of government spending and taxation is, on average, associated with a larger impact of government spending and revenue shocks on economic growth. In particular, the corresponding fiscal multipliers tend to be larger in countries that are more decentralized. While the empirical approach is similar to studies including cross country evidence such as Corsetti, Meier, and Müller (2012) or Ilzetzki, Mendoza, and Végh (2013), we acknowledge two potential shortcomings of our approach. First, the countries and their regions used in our study vary importantly in size of the regional and central governments, which makes the comparison of the different type of governments across countries difficult. Second, our the SVAR approach is based on total government spending, i.e. the sum of central and regional government spending; thus it does not identify if in any point of time, the government spending shocks resulted from the regional or central government.

We address these shortcomings by contrasting our results with a case study from Spain, which is one of the few countries in our sample displaying important time variation in the degree of decentralization. Spain offers a unique example of fiscal decentralization due to historical reasons. The dictatorship that ruled the country for almost forty years until 1975 imposed a very centralized fiscal system on a country with very heterogeneous regions. An important pillar of the Spanish transition to democracy involved transferring fiscal autonomy to the regions. We exploit the time variation as well as the fact that this decentralization process was not implemented in all regions at the same time. We argue that the timing of the implementations is a reaction to political rather than economic forces, therefore yielding an identification for the effects of decentralization on economic growth. The exact timing and sizes of these shocks are pinpointed via the “narrative approach” to fiscal policy evaluation (Romer and Romer 2010). We use data on fiscal spending and revenues at the regional level, which also allows for a more thorough decomposition of fiscal policy instruments into three series: direct taxation income, indirect taxation income and spending. Our results provide evidence for significant positive effects of decentralization on regional GDP growth with the

size of the effect being particularly large for the direct taxation series. Contrary to the cases of the decentralization of spending or indirect taxation, the regional governments were allowed to modify important legislative aspects concerning direct taxation in the latter case. Hence the increased decision power of the regions to employ fiscal instruments in general, and direct taxation in particular, leads to the positive impact of fiscal decentralization on regional output growth documented in this paper.

Our results are consistent with standard theories of fiscal federalism (Oates et al. 1972; Bordignon, Manasse, and Tabellini 2001) postulating that local governments have an informational advantage when implementing fiscal policy. While central governments will tend to make rather homogeneous allocations, regional fiscal policy can be tailored and will therefore be more effective if there is a large degree of heterogeneity in preferences and / or economic conditions within a single country. The case study also shows that when it comes to fiscal decentralization, the “how” is important relative to the “how much”: decentralization of direct taxation, which can be designed and implemented relatively freely by the regions, seems to be more effective than the decentralization of other fiscal instruments such as indirect taxation and spending.

The remainder of this paper is structured as followed: section 2 presents the cross country evidence starting with a discussion of the decentralization measure followed by the empirical implementations and results; section 3 presents evidence from Spain, while section 4 concludes.

3.2 International Evidence

3.2.1 Measuring Decentralization

This section investigates the cross-country relationship between decentralization and fiscal policy effectiveness in the form of government spending and revenue multiplier. Measures of decentralization typically fall in one of the two categories: the first focuses on fiscal policy, and the relation between expenditures and allocations, while the second focuses on the nature of the intergovernmental relations and their regulation (see e.g the survey Sharma 2006). We draw on the former, since it provides a clearer quantitative measure and the focus of this study is fiscal policy. Indeed, Sharma (2006) concludes that when it comes to the measurement of *fiscal* decentralization, the share of sub-national expenditures and revenues is considered to be the best indicator. Following this idea, the measure of decentralization we consider in the subsequent analysis is subnational (regional) spending as a percentage of total public spending of a respective country. Since we are interested in the effect of direct government spending, we exclude transfers (“social protection”) from both the regional

and the total government spending. However, as shown below, alternative decentralization measures based on total spending (i.e. including transfers) and relative tax rather than spending competences lead to a similar classification of “centralized” and “decentralized” countries.

One caveat of our approach to measuring decentralization is that it does not account for potentially delegated spending, i.e. regional spending that was not carried out in an autonomous manner but rather part of a central government’s mandate. However, similar decentralization measures are common in the literature and a good proxy for decentralization (Davoodi and Zou 1998; Oates 1985; De Mello 2001). Ebel and Yilmaz (2002) suggest that the measure should perform reasonably well for the developed countries used in this study.

We use yearly data from 1996 to 2011 on regional central government spending. Our decentralization index, presented in figure (3.1), is then computed as the average regional government spending net of transfers over the total government spending net of transfers:¹

$$D^i = \frac{1}{T} \sum_t \frac{\text{government spending by non-central government}_t^i}{\text{total government spending}_t^i}, \quad (3.1)$$

where D^i is the value of the index for country i .

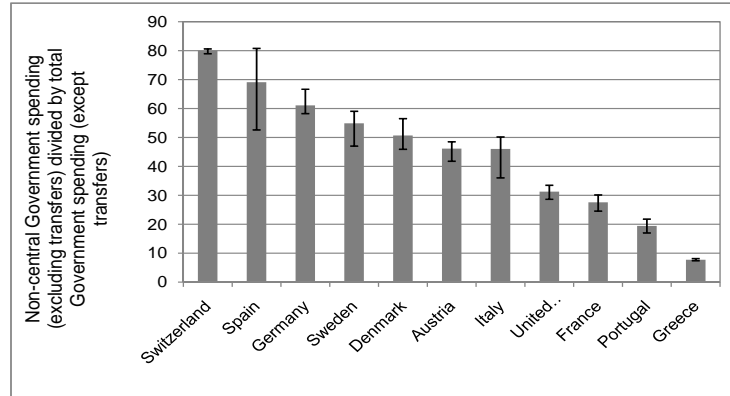


Figure 3.1: Decentralization Index: Regional share of spending net of transfers. Averages along with highest and lowest observation. Calculation based on yearly data 1996 - 2011.

The average of the decentralization measure for the period considered about 45 %, implying that on average, less than half of total non-transfer spending is carried out at the

¹The particular selection of the countries is based on data availability (excluding for example Eastern European countries) and variation in decentralization. Countries such as Belgium and Netherlands, for example, have a decentralization index very close to the median and where hence excluded.

sub-national level. However, there is a high degree of variability between the extremes of the index, with values as low as 8 % for Greece and as high as 82 % for Switzerland. In the medium of the spectrum, between 45 % and 60%, there is a group of five countries from Germany to Italy that are very close in terms of the index.

As shown in figure (3.13) in Appendix A, the measure of decentralization is very similar to the one that is computed including social spending. Similarly reassuring is that the decentralization index based on taxes, (figure 3.14 in Appendix A), suggests an almost identical grouping of the countries. The main difference is that Denmark appears to be the most decentralized country. Based on these observations, we group countries into decentralized countries (Group 1: Switzerland, Spain and Denmark), centralized countries (Group 3: UK, Greece, Portugal and France), and a range of medium centralized countries (Group 2: Sweden, Germany, Italy and Austria) for our later analysis.

A potential caveat is that there might be important time variation in our index that is “averaged away” through our computation of the index. In fact, due to data availability the data we use for the index construction starts in 1996, while the VAR analysis is based on data starting from 1980. However, empirically we find the index to be relatively stable over time. The graphs (3.1) and (3.13) and (3.14) in Appendix A also depict the maximum and the minimum observation for each country (indicated by the end of the whiskers). The range appears relatively small, except in the case of Spain. It is exactly this time variation that we exploit for the second part of this study.

Decentralization and Other Determinants of Fiscal Multipliers

The structure of government is not the only source of variability in the effect of government spending shocks. In order to identify effects that might arise through the degree of decentralization, we would like the index not to covary systematically with other determinants of government spending and taxation effectiveness. Indeed, the extensive literature on fiscal multipliers has identified several factors that can determine the size of fiscal multipliers. For example, Ilzetzki, Mendoza, and Végh (2013) show that for a large set of countries, openness to trade, exchange rate flexibility and outstanding government debt influence the size of multipliers. Auerbach and Gorodnichenko (2012) findings suggest that for a given country, multipliers depend on the current state of the economy. In particular, multipliers appear to be significantly larger in recessions than in expansions. Corsetti, Meier, and Müller (2012) find fiscal multipliers to be larger during financial crisis and fixed exchange rate regimes. Although the evidence concerning the determinants of fiscal multipliers is far from conclusive, we discuss below how the specific countries and groups might be affected differently by the

various factors.²

	Spending	Cons. Spending	Inv. Spending	Revenues
Group 1	0.211 (0.0595)	0.175 (0.0571)	0.0361 (0.00722)	0.410 (0.0975)
Group 2	0.262 (0.0645)	0.233 (0.0557)	0.0292 (0.00994)	0.479 (0.0742)
Group 3	0.246 (0.0263)	0.213 (0.0254)	0.0339 (0.0116)	0.389 (0.0817)

Table 3.1: Spending and Revenues over GDP by decentralization group. Spending denotes the sum of government consumption and investment expenditure; Cons. and Inv. Spending stand for consumption and investment spending, respectively. Standard deviations in parenthesis. Quarterly observations, 1980 - 2007.

Table (3.1) displays the average government spending, its disaggregated consumption and investment components and revenues, divided by GDP, for the various groups.³ Government spending, in particular consumption, appears to be somewhat smaller in decentralized countries (Group 1), but as seen in figure (3.15) in Appendix A, this relationship is rather weak. Contrary, the average share of revenues tend to be larger in medium decentralized countries (Group 2).⁴ Also, there does not seem to be a systematic relationship between decentralization and the level of government debt. The average government debt per GDP is around 50% for both centralized and decentralized countries, but larger for the medium group (table 3.2). As seen in figure (3.16) in Appendix A, this is mainly driven by the high debt-to-GDP ratio in Italy. Moreover, the numbers presented in table (3.2) document that the decentralized and centralized groups of countries appear to be quite similar along the crucial dimensions of GDP growth (capturing boom vs. recessions) and debt to GDP ratio and the sum of imports and exports relative to GDP as a measure of openness.⁵ Only the medium group displays a considerably higher average debt to GDP ratio, and also the highest openness indicator as measured by import plus exports to GDP.⁶

²In contrast to Auerbach and Gorodnichenko (2012), for example, Ramey and Zubairy (2014) do not find evidence for elevated fiscal multipliers during economic slacks and the recent financial crisis.

³Table (3.6) in Appendix A provides a more detailed overview over these variables by country.

⁴Note that table displays gross revenues (including transfers), which means that spending and revenues do not necessarily have to be equal to imply a balanced budget.

⁵Table (3.7) in Appendix A provides a more detailed overview over these variables by country.

⁶To the extent that decentralization is associated with factors such as fiscal discipline, as argued by Rodden (2002), for example, the effect from decentralization to fiscal multipliers might be indirect but can still be traced back to the former.

	GDP growth (pc)	(X+IM)/GDP	Population	Gross debt/GDP (pct.)
Group 1	1.847 (1.788)	2.228 (0.532)	17.33 (16.10)	48.59 (13.93)
Group 2	1.853 (1.430)	2.342 (0.614)	38.39 (31.38)	71.05 (21.26)
Group 3	2.082 (2.031)	1.843 (0.379)	34.01 (23.86)	53.91 (21.00)

Table 3.2: Descriptive statistics for additional determinants of fiscal multipliers by decentralization group, yearly observations 1980 - 2007. Population is in million. Standard deviation in parenthesis. Source: International Monetary Fund, World Economic Outlook Database and eLibrary.

Corsetti, Meier, and Müller (2012) provide definitions for fixed exchange rates regimes and financial crisis episodes for most countries in our sample.⁷ Table (3.3) shows that most countries, with the exception of UK, had a pegged exchange rate for most of the sample period. Financial crisis, however, seemed to be more frequent in the decentralized countries. Taken together though, we find only little evidence for major overlaps between decentralization and other determinants of fiscal multipliers that might be driving our results.

Country	Group	Currency Peg	Financial Crisis
Denmark	decentralized	1980 - 2007	1987 - 1992
Spain	decentralized	1984 - 2007	1984-1985
Austria	medium	1980 - 2007	-
Italy	medium	1983 - 1991, 1998 - 2007	1990-91
Sweden	medium	1980 -92	1991-94
France	centralized	1982 - 2007	1994 - 1995
Portugal	centralized	1990 - 2007	-
UK	centralized	-	-

Table 3.3: Overview of exchange rate characteristics and financial crisis periods according to Corsetti, Meier, and Müller (2012) for available countries.

3.2.2 Econometric Framework and Data Description

Our cross-country analysis considers unexpected changes in government spending and revenue in order to evaluate the impact of fiscal policy on output growth. Following Fatás, Mihov, et al. (2001) and Blanchard and Perotti (2002) we employ Structural VARs (SVAR) to quantify this impact. Their approaches have in common that they exploit decision lags in fiscal policy-making which allow to identify fiscal shocks. Since we are using quarterly

⁷Not covered are Greece, Germany and Switzerland.

data in our study, the assumption that discretionary government purchases and revenues are not going to be made effective law and implemented within the same observation period as a GDP shock is likely to be met; hence they can be predetermined with respect to the macroeconomic variables.

Proceeding with the description of the reduced form model, for our baseline specification, we consider the following vector

$$X_t = (g_t, y_t, r_t)', \quad (3.2)$$

where g_t is the growth rate of real government consumption and investment spending, y_t is the growth rate of real GDP and r_t is real government revenue growth.⁸

The following reduced form model is then estimated individually for each country:

$$X_t = c_0 + \sum_{i=1}^k \phi X_{t-i} + e_t, \quad (3.3)$$

where c_0 is a constant and $e_t \sim \text{WN}(0, \Sigma)$ represent the reduced form error shocks. In order to ensure comparability between the countries, we choose a lag length of $k = 5$ in each estimation. This appears to be a reasonable compromise between the four lags proposed by Blanchard and Perotti (2002) and the 6 lags employed by Mountford and Uhlig (2009).

The data for government investment and consumption spending, revenues and GDP for 11 European countries are obtained from Oxford Economics. The spending and the GDP variables are already obtained in real terms, while we transform nominal revenues using the GDP deflator.⁹ Growth rates are computed via log changes. For all countries, we use data from 1985 - 2007, yielding $T = 85$ observations. Using pre-2008 data ensures that our results are not affected by the financial crisis.

3.2.3 Identification of Structural Innovations

We assume that the reduced form errors are related to their structural counterparts via the representation $Ae_t = Bu_t$, where $u_t \sim (0, I)$ are the structural shocks. Without restrictions on the parameters in A and B the structural model is not identified. Hence additional assumptions will be necessary to disentangle A and B from the estimates of the variance-covariance matrix of the reduced form errors. Our model is similar to the one of Blanchard and Perotti (2002) and Perotti (2005), who applied this approach to estimate the effect of government spending and tax shocks for the US and several other OECD countries. In particular, let the structural relationship between the reduced form and the structural errors

⁸Ideally, the government revenue should be net of transfers. However, we take the gross series as a proxy as we did not find consistent data regarding net transfers for all countries.

⁹The series not already seasonally adjusted were adjusted using the Census X-13 methodology.

take the form

$$\begin{pmatrix} 1 & 0 & 0 \\ \alpha_{21} & 1 & \alpha_{23} \\ 0 & \alpha_{32} & 1 \end{pmatrix} \begin{pmatrix} e_t^g \\ e_t^y \\ e_t^t \end{pmatrix} = \begin{pmatrix} \sigma_g^2 & 0 & \beta_{13} \\ 0 & \sigma_y^2 & 0 \\ \beta_{31} & 0 & \sigma_t^2 \end{pmatrix} \begin{pmatrix} u_t^g \\ u_t^y \\ u_t^t \end{pmatrix}, \quad (3.4)$$

where α_{21} , α_{23} and α_{32} are, respectively, the value of the elasticities of output relative to government spending and taxes, and the elasticity of taxes relative to output. The parameter β_{13} captures the response of government spending to unexpected (structural) shocks in revenue, while conversely, β_{31} presents the response of government revenues to unexpected (structural) shocks in spending. Since the reduced form variance-covariance matrix has six distinct elements, additional assumptions are necessary to identify the parameters α_{21} , α_{23} , α_{32} , β_{13} and β_{31} . They are derived from both exclusion restrictions and outside information. On a quarterly frequency, fiscal policy is plausibly subject to decision lags, i.e. the time needed for fiscal policy makers to respond to changes in output is at least one quarter. Then any remaining correlation between the unpredicted components of government spending and output is due to the impact of government spending on output. Similarly, α_{32} , the output elasticity of taxes can then be obtained by regressing revenue on the tax base; and the corresponding estimate can be imposed directly in equation (3.4). One caveat is that the tax multipliers obtained this way are quite sensitive to the particular estimate of α_{32} (Caldara and Kamps 2012). However, for the countries considered in this study, the values appear quite similar and close to 1 according to recent OECD estimates (Price, Dang, and Guillemette 2014). This is considerably lower than the estimate of 1.85 obtained by Blanchard and Perotti (2002) for the United States, but close to the one of 0.95 obtained by Tenhofen, Wolff, and Heppke-Falk (2010) for Germany. Moreover, since the main purpose of this study is to derive *relative* values of fiscal multipliers across countries, and the elasticities appear indeed similar for these countries, we expect the uncertainty regarding the output elasticity of taxes not to weight into our results significantly.¹⁰ Finally, we follow Blanchard and Perotti (2002) in setting $\beta_{13} = 0$, which implies that spending decisions come before tax decisions.

One caveat of our specification is that given that g_t presents aggregate government spending, we have insufficient information if, in each given point in time, spending came from local or central government.¹¹ Our VAR results therefore identify the average effect of government spending on growth, without discriminating between the regional and central government

¹⁰In fact, we performed robustness analysis using different values for α_{32} , which changed the size of the multipliers somewhat, but not the relative ordering.

¹¹This of course would allow us to evaluate the relative effect of central and regional government shocks, respectively more directly. Unfortunately historical time series for quarterly series of regional expenditure is not available for most countries.

spending more directly.

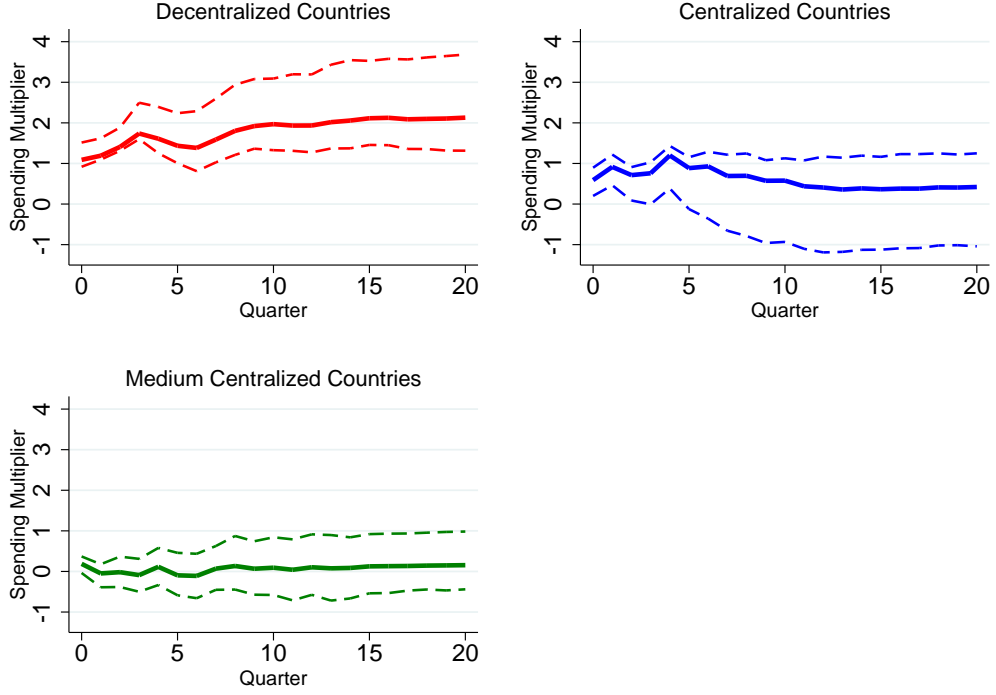


Figure 3.2: Cumulative Spending Multipliers, average across country group. Point estimates with 84% and 16% bootstrap percentiles.

3.2.4 Results

This section describes the estimation results of the model presented in the previous section. We present the results in the following form. First, each countries' output cumulative response to a fiscal shock is standardized by dividing the cumulative GDP response to fiscal shocks by the ratio of GDP relative to the respective fiscal variable and the standard deviation of the fiscal shocks:

$$\text{Dynamic Multiplier} = \frac{\text{Output response}}{\text{Initial Fiscal Shock}} * (\text{Average fiscal variable share of GDP}).$$

Second, we present average results for the three groups of countries. Averages are taken over the country specific multipliers in the respective group at each point of time. This approach is similar to the one employed by Ilzetki, Mendoza, and Végh (2013), who classify countries according to a certain characteristic and then estimate fiscal multipliers separately for each group using panel VARs. Instead, our averaging method does not restrict the

dynamics for each country in the group to be the same. We employ a bootstrap procedure, which is outlined in Appendix A in more detail, in order to compute the corresponding confidence intervals.

The Effect of Government Spending Shocks

Figure (3.2) depicts the multipliers for total government spending. The multiplier for the most decentralized countries is about one on impact, and slightly increasing to a value of around two after ten quarters. The multiplier for the centralized countries is noticeably lower and around 0.5 on impact, and reaches its maximum of one after about one year before declining afterward. In contrast, the multiplier for the medium countries is only slightly positive on impact but indistinguishable from zero thereafter. One explanation for this results is that the group of medium countries also exhibits the highest average debt to GDP and openness measure, displayed in table (3.2), both of which are associated with lower multipliers (Ilzetzi, Mendoza, and Végh 2013).

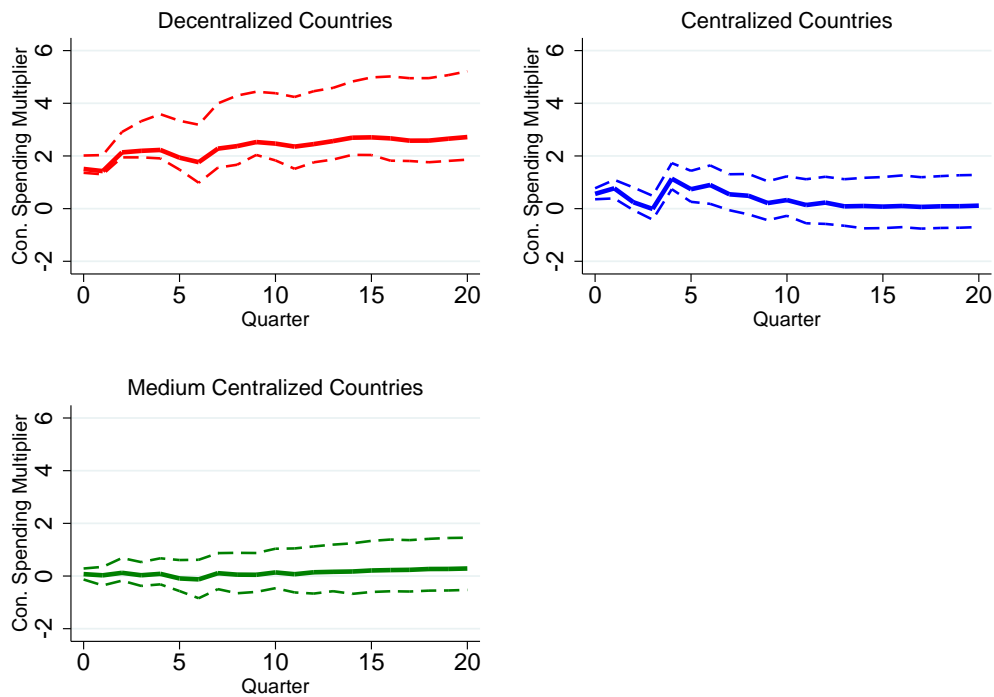


Figure 3.3: Cumulative Government Consumption Multipliers, average across country group. Point estimates with 84% and 16% bootstrap percentiles.

Disaggregated Government Spending

This section considers the effect of government consumption and government investment spending separately. Hemming, Kell, and Mahfouz (2002) note that fiscal multipliers can vary across different policy instruments, so we expect to obtain additional insights from a disaggregation of the different components of government spending. The estimation follows the baseline model described above, where g_t contains either consumption or investment spending. Figure (3.3) depicts the multipliers for government consumption spending only, with a pattern strikingly similar to the aggregate spending multiplier. The only differences appear to be a dip in the multiplier for the centralized countries after 3 quarters, and a slightly larger long-run response of output growth for the medium countries.

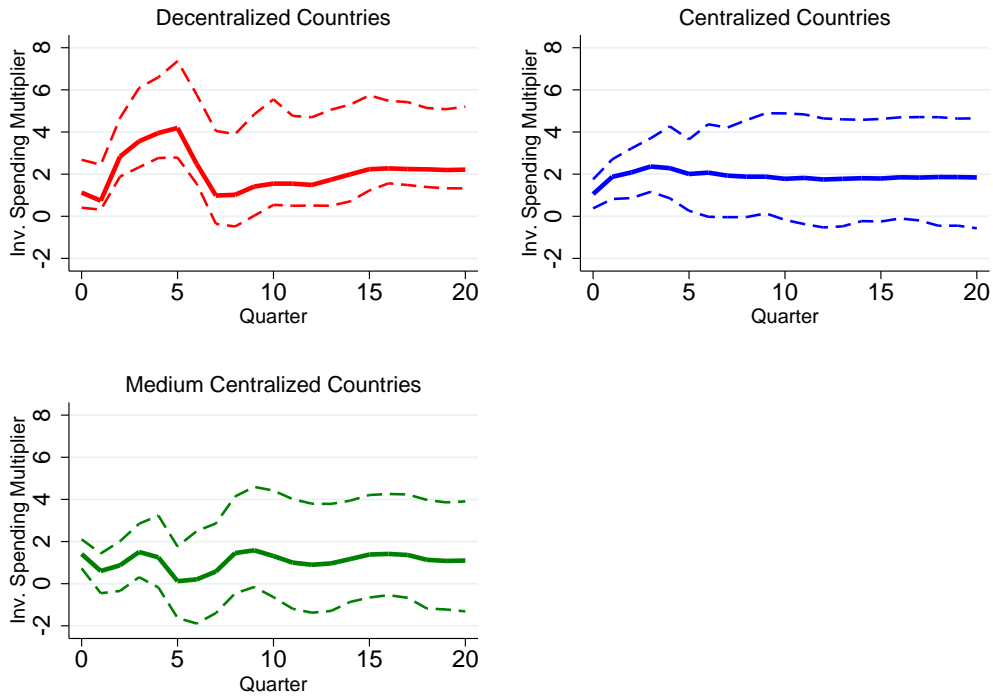


Figure 3.4: Cumulative Government Investment Multipliers, average across country group. Point estimates with 84% and 16% bootstrap percentiles.

In contrast, the results for the government investment spending only, presented in figure (3.4), indicate much larger multipliers around two for both the centralized and decentralized countries. Here we notice little difference according to decentralization. Moreover, the multiplier for the medium group turns negative after several quarters. Obtaining negative estimates of the multipliers is not uncommon (Perotti 2005), and can occur when distortionary taxes are imposed following debt financed spending (Baxter and King 1993).

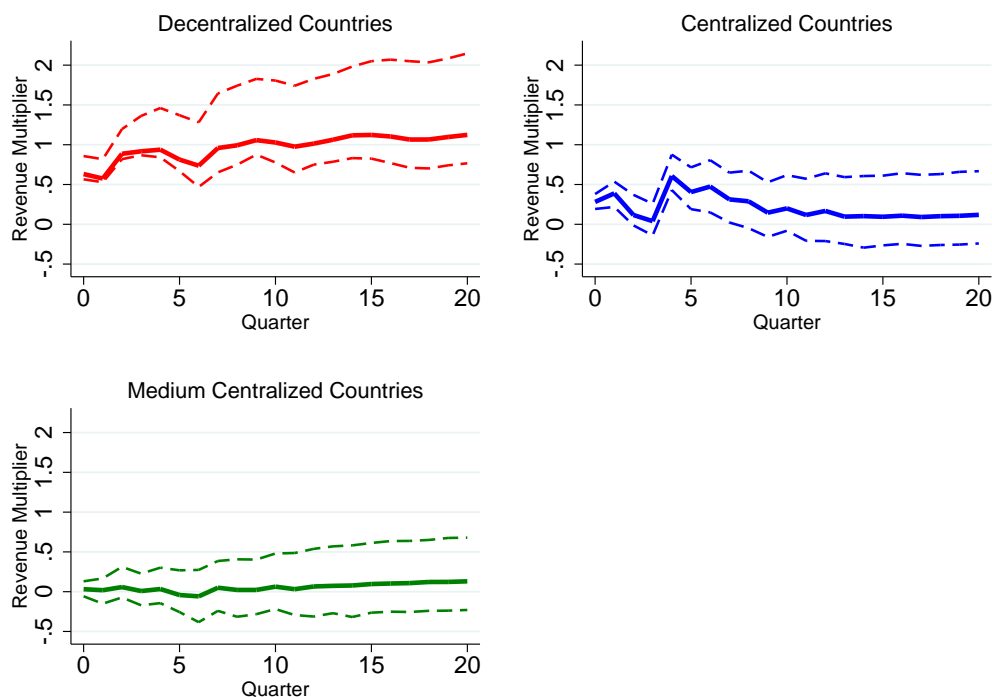


Figure 3.5: Cumulative Revenue Multipliers, average across country group. The multipliers describe the output response to a negative tax shock. Point estimates with 84% and 16% bootstrap percentiles.

The Effect of Government Revenue Shocks

Figure (3.5) displays the estimated revenue multipliers, which appear to be of a smaller magnitude. On impact all multipliers are similar and around 0.5, but while the multipliers for the decentralized and medium countries increase over time to a value around 1, the multiplier for the centralized countries slowly declines.

Taken together, for all cases considered, the decentralized countries exhibit relatively large multipliers. Only in the case of government investment spending, the group of centralized countries exhibits multipliers of a similar magnitude as for the decentralized countries: for government consumption spending and revenues, they are substantially lower and appear to be less persistent.

3.2.5 Robustness

This section discusses two robustness checks for the results presented in the previous subsection. The first is based on estimations of bivariate VARs that identify government spending and government revenue shocks separately; the second contrasts our results from those re-

ported in Dellas, Neusser, and Wälti (2005), who provide point estimates for multipliers for the majority of the countries in our sample.

Estimates based on Bivariate VARs

A further robustness check involves contrasting our results from a series of bivariate VARs that identify government spending and government revenue shocks separately. An advantage of using bivariate models is the calculation of the cumulative multipliers from the VAR system described below are not sensitive to the persistence of the fiscal shocks, whereas for larger system, this is generally not the case (Giordano et al. 2007). We consider the vectors X_t^1 , X_t^2 , referring to the models with government spending and with government revenue, respectively,

$$X_t^1 = (g_t, y_t)', \quad X_t^2 = (y_t, r_t)', \quad (3.5)$$

and estimate the reduced form models with the same specification as the baseline case. Similar to the baseline model, the identifying assumptions are

$$\begin{pmatrix} 1 & 0 \\ \alpha_{21} & 1 \end{pmatrix} \begin{pmatrix} e_t^g \\ e_t^y \end{pmatrix} = \begin{pmatrix} \sigma_g^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \begin{pmatrix} u_t^g \\ u_t^y \end{pmatrix}, \quad (3.6)$$

in the case of the model including government spending and

$$\begin{pmatrix} 1 & \alpha_{12} \\ \alpha_{21} & 1 \end{pmatrix} \begin{pmatrix} e_t^y \\ e_t^r \end{pmatrix} = \begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix} \begin{pmatrix} u_t^y \\ u_t^r \end{pmatrix}, \quad (3.7)$$

in the case of the model including revenues, where α_{21} is imposed by the output elasticity of revenues. As for the three variable VARs, we choose $-\alpha_{21} = 1$ for all countries.

Figures (3.18), (3.17), (3.19) and (3.20) depicts the multipliers for total government spending, government revenues, and the disaggregated consumption and investment multipliers respectively, showing almost identical patterns as for the VARs with three variables.

Comparison with Results from Other Cross-country Studies

Our results regarding spending multipliers are supported by the estimates presented in Dellas, Neusser, and Wälti (2005), who investigate the effect economic openness has on the size of fiscal multipliers. The authors' estimates of government consumption multipliers rely on a similar yet substantially larger SVAR system, that includes inflation and interest rates, among others.

Figure (3.6) plots, for the countries available, the decentralization index against the estimates obtained by Dellas, Neusser, and Wälti (2005). Supportive of our findings, there is

a small, positive relationship between decentralization and government consumption multipliers on impact, that becomes substantial after 4 quarters.

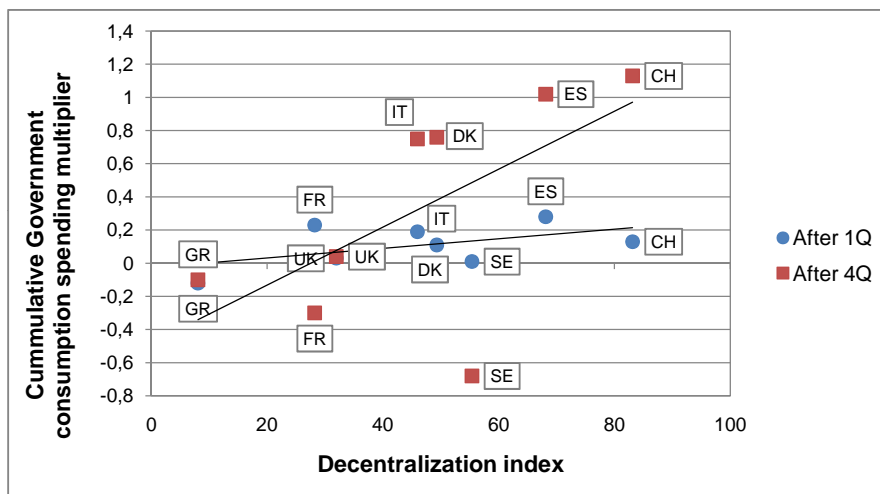


Figure 3.6: Decentralization Index vs. cumulative government consumption spending multipliers reported in Dellas, Neusser, and Wälti (2005).

3.3 Intrnational Evidence from Spain

3.3.1 The Decentralization Process in Spain (1975-2007)

As apparent from the previous section, the regional decentralization of a country can have a noticeable impact on the effectiveness of fiscal policy. Allowing the regional governments to carry out a larger share of the total public spending or collect a bigger proportion of taxes appears to affect how the economy in general - and economic growth, in particular - react with respect to shocks to the fiscal instruments. In this section, we address some weaknesses of the cross-country approach, namely the inability to differentiate between central and regional government spending in each point of time, and a potential difficulty to compare international regions that differ vastly in its size.¹² The countries studied on the previous section exhibit a relatively stable regional configuration, as shown by the rather small variation of their decentralization indexes, with one notable exception: Spain. It is exactly this time variation that we exploit in order to address the following question: if decentralization affects fiscal policy, what are the direct effects from further decentralizing it fiscal policy on economic growth?

¹²For example, Germany's largest region Nordrhein-Westfalen is about three times larger than Denmark in terms of population.

Spain offers a unique example of fiscal decentralization due to historical reasons. The dictatorship that ruled the country for almost forty years until 1975 imposed a very centralized fiscal system on a very heterogeneous country. One of the main pillars of the Spanish transition to democracy involved transferring autonomy to the regions. The gradual and asymmetrical nature of the process can be naturally exploited to test the effects of increasing the fiscal competence of the regional governments.

The Spanish Constitution of 1978, on its Title VII, allowed for the concession of extensive prerogatives to the regional governments, called *Comunidades Autónomas*.¹³ The process was not immediate, however, to the point that it can still be considered an ongoing. Some regions with stronger regional identities moved quickly to approve their regional Constitutions (*Estatutos de Autonomía*) and started the transfer of prerogatives while other regions lagged behind and only received these prerogatives after nationwide agreements. The heterogeneity on the timeline and its predominantly political nature offers a natural experiment of fiscal decentralization that we exploit to measure its impact on economic growth.

The analysis of this section is based on the data of the *Comunidades Autónomas*' Budget Series from the General Secretary for Local and Regional Coordination (*Secretaría General de Coordinación Local y Autonómica*). The database offers yearly consolidated series of 9 income categories and 9 expenditure categories from 1984 to 2013 for the 17 main Spanish regions. As in the previous section, we will focus on the 1984-2007 period, as the depth of the current economic crises complicates greatly any analysis of the data from 2008 onwards.

Our analysis will focus on 3 series: direct taxation income, indirect taxation income and spending. The last series is a composite of expenditure on public wages and public consumption. The analysis of these series allows us to identify major transfers of autonomy to the regional governments, its timing and its size as a percentage of regional GDP. Combining the analysis of the series and a narrative approach, we were able to identify five major episodes of fiscal decentralization, where a prerogative was transferred to the regional governments.

Two important points characterize this analysis. First, even though in every episode regional tax income or spending increases, these changes are rather interpreted as decentralization shocks than “classical” tax or spending shocks. We use this interpretation because, in principle, the increase in regional spending and revenue merely offsets the fiscal activity previously carried out by the central government and does not necessarily have to lead to a change in total (local and central) government spending or taxes in the region. To the extent that it does change total government spending or taxes, our analysis provides a measure for the joint effect of a (decentralization induced) change in actual spending or taxes and a change in efficiency of spending and tax collection. In either case the results provide evidence for the effects arising from decentralization through increased fiscal independence of

¹³A detailed report on the beginning of the process can be found in Molero (2001).

the regions. Second, we provide evidence that the timing of the implementation of regional fiscal policy was largely exogenous to economic conditions in a particular region (or Spain) and primarily politically motivated.¹⁴ This exogenous variation reinforces our identification of the effects arising from decentralization shocks.¹⁵

3.3.2 Episodes of Decentralization

Based on a narrative approach to identify discretionary policy measures, we observe five major episodes in the decentralization process of Spain:

1. *The decentralization of health services.* This process spanned over 20 years, with some regions, like Catalonia, gaining the prerogatives on health services as early as 1981 whereas the majority of the regions finally gained the competence in 2001. On average, this transfer of competence resulted on a permanent increase of regional spending of over 3% of regional GDP.
2. *The decentralization of non-tertiary education.* As with the previous case, regions as Catalonia and the Basque Country started handling non-tertiary education as early as 1980. The process of decentralization concluded in 1999. The result was an average permanent increase of regional spending of around 2.5% of regional GDP.
3. *First transfer of the Income Tax.* In 1996, regions were allowed to keep up to 15% of all the Income Tax collected on their territory. Five regions stood out of this agreement: on one hand, Navarra and the Basque Country already handled most of their own taxes; on the other hand, Andalucía, Extremadura and Castilla-La Mancha argued that the cession broke the principle of regional solidarity and was therefore unconstitutional. These regions appealed the law on the Constitutional Court. The appeal was overturned and these regions finally complied and started collecting their allotted share when the Income Tax was transferred for the second time, as discussed below. On average, this transfer created a permanent increase on direct taxation income of the regions of around 1% of regional GDP.
4. *Second transfer of the Income Tax.* In 2001, the proportion of the Income Tax collected on their territory that regions were allowed to keep raised up to 33 %. This time, only Navarra and the Basque Country were not directly affected. The size of the shock was similar to the first transfer of Income Tax.

¹⁴See, for example, the analysis of Suárez-Pandiello (1999).

¹⁵ One drawback of our approach (which is also the case for the cross-country section) is that we cannot control for anticipation of these effects.

5. *Transfer of the Value Added Tax (VAT)*. Also in 2001, regions were allowed to keep up to 35% of all the VAT collected on their territory. As before, Navarra and the Basque Country were not directly affected, but neither were the Canary Islands, which have a different indirect taxation regime. The transfer resulted on a permanent increase of indirect taxation income of around 3.7% of regional GDP.

The nature of these series allows us to use an empirical strategy similar to Romer and Romer (2010). The authors introduce what has become known as the “narrative approach” to identify fiscal policy shocks by analyzing a series of exogenous tax changes based on historical records in the US.¹⁶ Similarly, we construct series of exogenous decentralization changes, valued 0 for every t except where we have identified a decentralization change, in which case the series takes the value of the size of the change in terms of % of regional GDP. In this case, the year of the shock will be the year in which we observe the actual change in the series of the respective fiscal measure, not the year in which the legislation was introduced. Further, we combine the first two episodes into a single “Spending Decentralization” shocks series; episodes three and four are combined into a single “Direct Taxation” shocks series, while the last episode defines “Indirect Taxation” shocks. We also define a “Decentralization” shock series, created by combining all shocks. With this definition, we aim at capturing the effect of the change when it is effectively introduced. Given that our series on decentralization changes are the reflection of a political process, we expect no systematic correlation between these changes and other determinants of output growth. Evidence in favor of this hypothesis is presented in the next section.

We explain our approach to identifying decentralization shocks via the example of the region of Aragón. Figure (3.7) shows the first difference of the three aforementioned series measured as percentage of regional GDP. It shows clearly the five episodes of decentralization. The transfer of non-tertiary education appears as an increase on spending of almost 2% of regional GDP in 1999, whereas the transfer of health services is captured by the jump on spending of more than a 3% of regional GDP during 2004. The increased cession of the Income Tax appears as spikes in Direct Taxation of around 1% of GDP during 1998 and then 2002. Finally, the cession of part of the Value Added Tax creates a spike on Indirect Taxation of 3.5% of regional GDP during 2004. Notice also that figure (3.7) presents the first difference of the relative fiscal variable and hence all episodes represent permanent shocks as the prerogative is transferred to the region on a permanent basis.

¹⁶Subsequently these series are employed by the authors to quantify the effect of tax changes on GDP growth.

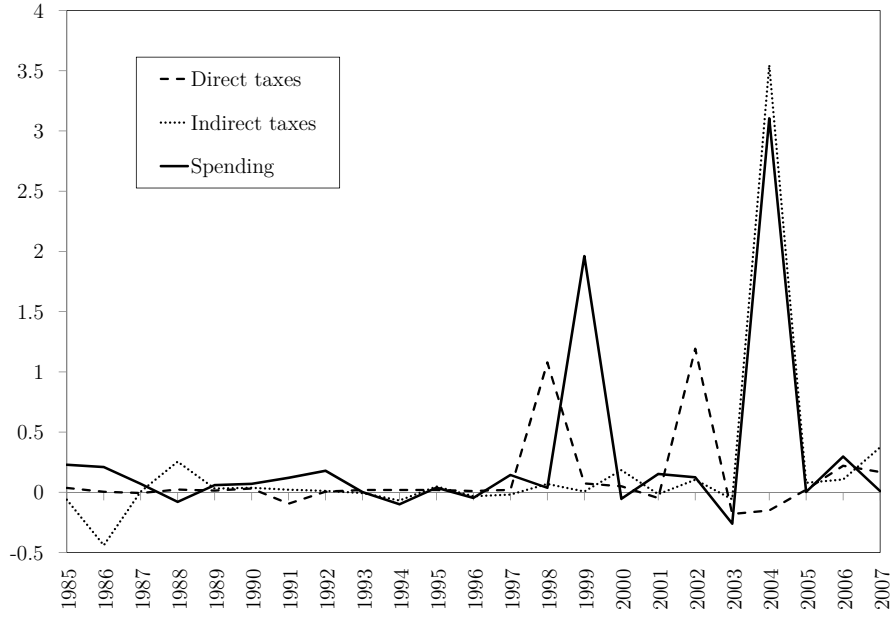


Figure 3.7: First difference of the Spending, Indirect Taxation and Direct Taxation series for Aragón, measured as % of regional GDP, 1985-2007.

	Full sample		Early adopters		Late adopters	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Avg Population	2335.5	(2071.7)	3467.6	(2432.4)	1542.9	(1410.9)
Avg. GDP per capita	12821.7	(2639.7)	13172.8	(2585.0)	12575.9	(2787.5)
Avg. real GDP growth	3.71	(0.58)	3.70	(0.56)	3.74	(0.63)

Table 3.4: Descriptive statistics for Early adopters and Late adopters. Population is measured in thousands, GDP in Euros and growth rates are computed year-on-year. Standard deviation in parenthesis. Years 1984-2007.

3.3.3 Justification of the Narrative Approach

In order to identify the effect of decentralization on output, it is crucial that the timing of the policy measures is not driven systematically by economic conditions.¹⁷ As we have discussed previously, the decentralization process in Spain is interesting in this sense: because the timing of the implementation was brought about by the political process, the shocks can be reasonably thought of as variations in fiscal policy that are exogenous to economic conditions. In this section, we investigate the claim of exogeneity in two ways.

As a first approximation, we divide the regions into early and late adopters to see if we observe systematic differences between these groups. We classify as early adopters those regions that took the initiative to decentralize competences and, therefore, got these competences early; and as late adopters the regions that only received competences on the

¹⁷If, for example, the implementation is carried out during episodes of (non-) favorable economic forecasts, we might find a spurious positive (negative) effect from decentralization to output growth.

	Direct taxation	Indirect taxation	Spending	Any shock
ΔY	1.533 (0.93)	0.877 (0.69)	0.956 (0.86)	2.703*
$\Delta Y(-1)$	0.607 (0.92)	-0.338 (0.68)	-0.072 (0.85)	-0.032 (1.38)
$\Delta Y(-2)$	-0.931 (0.87)	-0.077 (0.64)	0.351 (0.80)	-0.381 (1.30)
$\Delta Y(-3)$	0.402 (0.75)	0.874 (0.56)	0.316 (0.70)	1.164 (1.13)
$\Delta Y(-4)$	-0.041 (0.67)	0.537 (0.50)	-0.297 (0.62)	-0.017 (1.01)
Constant	0.031 (0.05)	-0.021 (0.03)	0.028 (0.04)	0.049 (0.07)
R-squared	0.018	0.021	0.010	0.020
No. of obs.	323	323	323	323

Table 3.5: Regression results for the effects of past GDP growth on the probability of implementing a decentralization measure. The dependent variable is a dummy that takes on the value 1 in the years when reforms were implemented.

Standard errors are reported in parenthesis

*, ** and *** indicate significance at the .90, .95 and .99 level, respectively

framework of nationwide agreements, where the central government took the initiative. The sorting - motivated by the historical accounts of the decentralization process provided in the previous section - is displayed in table (3.8) in Appendix B. The differences in terms of the timing of decentralization is indeed substantial: for example, the average year in which the decentralization of health services occurred was 1988 for the early adopters and 2000 for the late adopters. Table (3.4) shows that while early adopters are larger regions in terms of population, there appear to be no systematic differences in GDP per capita or growth rates for our sample period.

To further justify our narrative approach - in particular to exclude economic conditions as a cause for the introduction of decentralization - we perform four predictive regressions of the following form: as independent variable, we defined dummy variables (one for each of the three types of decentralization shocks and another one picking up any decentralization shock) with value 1 on the years where we have identified a decentralization shock in our series and 0 otherwise. As dependent variables we use lags of real regional GDP annual growth. Due to the nature of our data, we use a panel data regression with fixed effects in a linear probability framework. The results of this exercise can be seen in Table (3.5).

As can be seen in table (3.5), we fail to observe any consistent and significant impact from past GDP growth on the timing of the decentralization shocks. The largest and only significant coefficient corresponds to the *contemporaneous* relationship between decentraliza-

tion and growth and plausibly - in particular due to the absence of any effects from lagged output growth - captures the effect from decentralization on growth. Both pieces of evidence reinforce our interpretation that the timing of the decentralization episodes were mostly politically motivated and hence exogenous to current output growth. This justifies the use of the narrative approach as outlined in the next section.

3.3.4 Empirical Framework

The empirical specification extends the regression framework proposed by Romer and Romer 2010 to panel data: we employ fixed effects regressions (i.e. including region-specific intercepts) with real regional output growth as the dependent variable.

The regressors include the series of decentralization shocks, including lags, as well as lags of output growth. More specifically, denoting by dir_t , ind_t and spe_t our series of Direct Taxation shocks, Indirect Taxation shocks and Spending shocks respectively, the regression framework is

$$\Delta Y_{i,t} = a_i + \sum_{j=0}^M b_j D_{i,t-j} + \sum_{k=1}^N c_k \Delta Y_{i,t-k} + e_{i,t}, \quad (3.8)$$

where $Y_{i,t}$ is the logarithm of real regional output in region i at time t and $D = \{dir, ind, spe\}$ is our measure of decentralization. Estimations are carried out for the joint set of fiscal policy shocks series, as well as for the individual and the aggregate (sum of the three single series) series. In order to allow for a lag in the output reaction to decentralization, we include five lags of the fiscal policy shocks, i.e. $M = 5$, and one lag of GDP growth, to control for the usual autoregressive dynamics of GDP growth.

We measure output using nominal series of regional GDP deflated by an annual average price index for every region. $D = \{dir, ind, spe\}$ is our measure of changes in decentralization as a % of regional GDP. The data is yearly, the period of analysis is 1985-2007 and the cross-sectional units are the 17 Comunidades Autónomas.

3.3.5 Results

This section presents and discusses the estimation results for the effects of the different type of decentralization shocks. Table (3.9) in Appendix B presents the results for the regressions of Equation (3.8). In addition, we compute the cumulative impact after m periods as the sum of the regression coefficients corresponding to respective lags of the fiscal policy shocks, to obtain the corresponding point estimates,

$$\text{Cumulative Effect}(m) = \sum_{j=0}^m b_j. \quad (3.9)$$

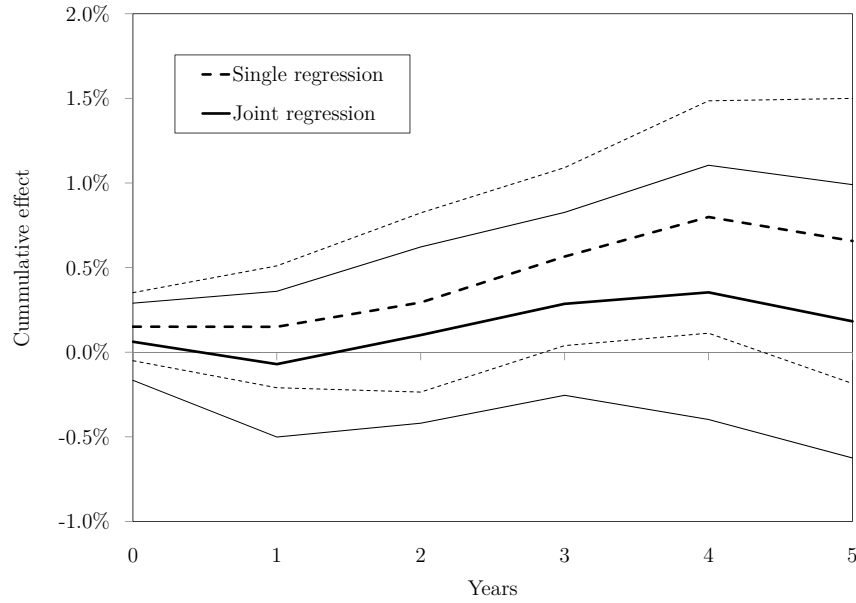


Figure 3.8: Accumulated response of an spending decentralization shock, single and joint regressions. Thinner lines represent the 95% confidence interval.

Spending Decentralization

The decentralization of spending categories like health care and non-tertiary education were some of the largest, in terms of regional GDP, transfers of competences registered in our database. However, we find only a small effect of these transfers on regional growth.

On the individual regression, without the other two shocks included, a transfer of spending competences equivalent to 1% of GDP achieves would achieve its maximum effect 4 years after the shock, adding less than 0.8 percentage points to regional GDP after that period, although for most periods the effect is not significant at the 95 % confidence level. When we introduce the other two shocks in the regression, the spending decentralization shocks are not longer significant and the size of their effect is smaller.

Indirect taxation decentralization

In the case of the decentralization of indirect taxation competences, we find a somewhat larger effect that is, however, only significant in the long term (after 4 years) in the joint regression. On the single regression, the decentralization of indirect taxation equivalent to 1% of regional GDP would add a maximum of around 1.28 percentage points to regional growth after 5 years. However, after controlling for the other shocks, this decentralization shock only gains significance after 4 years. The size of the effect is smaller and just 0.94 percentage points after 5 years.

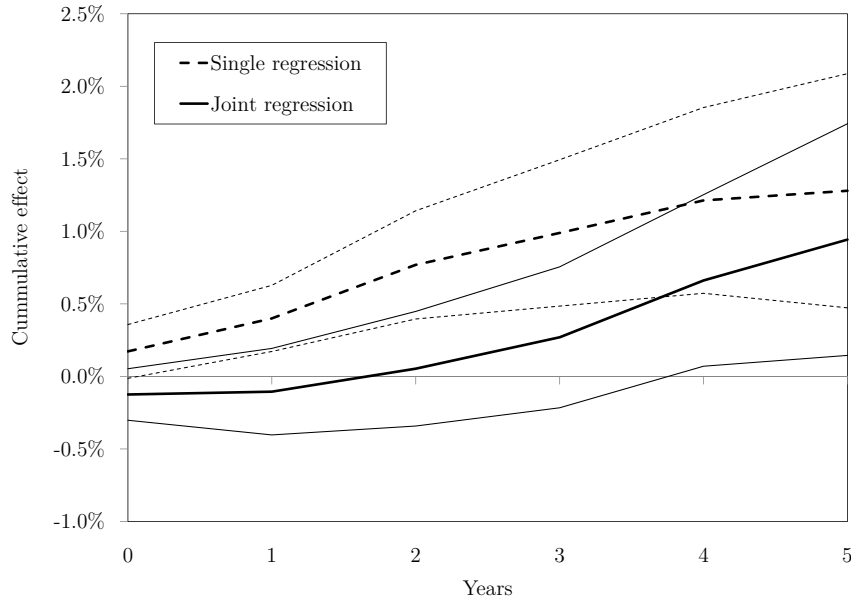


Figure 3.9: Accumulated response of an indirect taxation decentralization shock, single and joint regressions. Thinner lines represent the 95% confidence interval.

To put this result into perspective, during the sample period of 1984-2007, the Spanish regions grew an average of 3.71% yearly. On a six year period, this growth rate implies a total growth of a 22.26 per cent. The average indirect taxation decentralization shock amounted to 3.70% of regional GDP which, multiplied by the accumulated effect computed before, implies an increase of 3.49 percentage points over the next 6 years (to include contemporaneous effects) or, in other words, an increase of around 15.7% over the average registered regional GDP growth.

Direct taxation decentralization

The decentralization of direct taxation registers the largest effects on regional GDP growth despite being, on average, the smaller of the three shocks considered. Furthermore, of the three shocks considered, the decentralization of direct taxation is the only one whose effect is positive and statistically significant for all number of lags considered. When the direct taxation shocks are included as regressors individually, the estimated accumulated effect of decentralizing direct taxation amounts to 1% of regional GDP would be to add 5.26 percentage points to regional GDP after 5 years. When all decentralization shocks series are included in the regression, the maximum effect peaks after 3 years at around 4.4 percentage points.

Similar to the case of the decentralization of indirect taxation, we can put this number into perspective. In this case, the average direct decentralization shock amounted to around

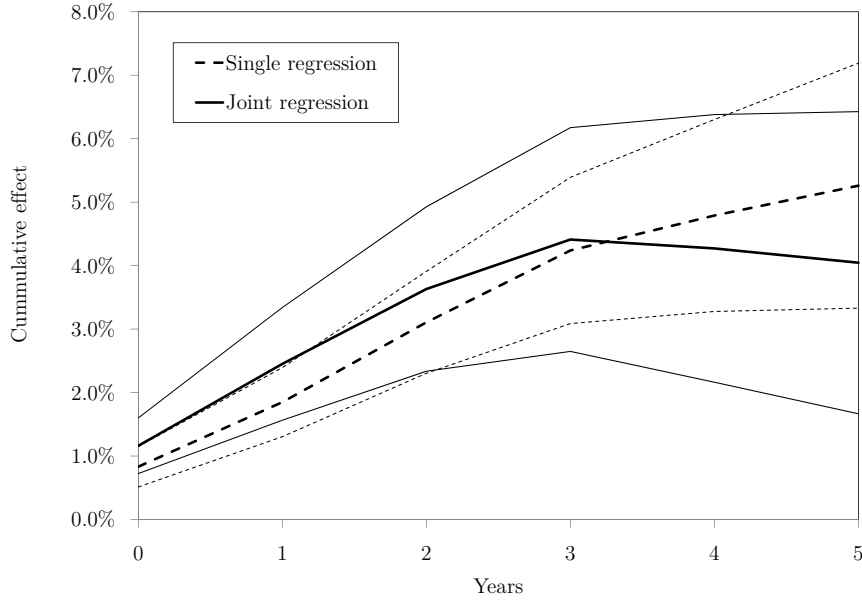


Figure 3.10: Accumulated response of a direct taxation decentralization shock, single and joint regressions. Thinner lines represent the 95% confidence interval.

1.06% of regional GDP, so the comparison is more straightforward. With an average growth of 3.71% yearly, the average region would have grown 14.48% over a 4 year span (again, to include contemporaneous effects), so that the added growth due to the decentralization shock amounts to 31.4% of the average registered regional GDP growth.

Aggregate decentralization shock

Finally, for reference, we consider an aggregate decentralization shock, constructed simply by adding together the series of the three individual shocks. Doing so, we observe how this aggregate shock is positive and statistically significant from impact, peaking after 4 years with an cumulative effect of around 0.74% of regional GDP.

In our data, the average aggregate decentralization shock amounts to 2.90% of regional GDP which, multiplied by the accumulated effect computed before, implies adding 2.14 percentage points to GDP growth over the following 4 years. Given that the average region would have grown 18.55% over a 5 year span, the shock would be equivalent to an increase of 11.6% of the total registered GDP growth.

3.3.6 Discussion

For a interpretation of these results, it is useful to place them in their historical context. Furthermore, the results we have obtained in the previous section allow us to connect to

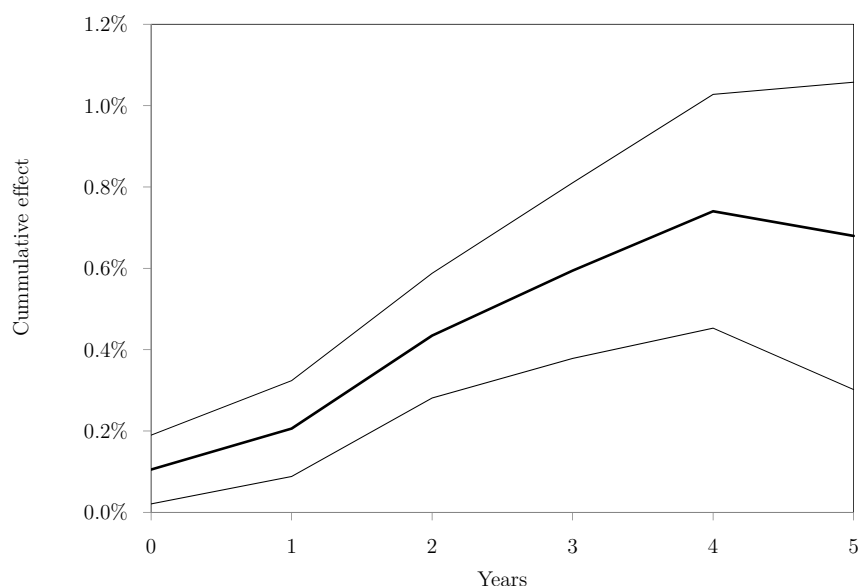


Figure 3.11: Accumulated response of an aggregate decentralization shock. Thinner lines represent the 95% confidence interval.

prevailing theories of fiscal federalism.

First, the changes on spending decentralization involved mainly the transfer of the competences on Health Care and Education to the regions. Since all regions already had functioning sanitary and educative sectors, the margin for fiscal maneuver was limited: collective agreements with the workers had to be honored and standards were set so as to ensure that every citizen in the country had access to a similar level of services, for example. Therefore, it is to be expected that the spending decentralization shocks studied here will have a small effect on the economic output of the regions.

Second, the changes in indirect taxation consisted typically in an increase of the proportion of taxes collected on the region that the regional governments could hold on to. On the one hand, these changes were usually matched with corresponding reductions on the amount of transfers received from the central government. Furthermore, the tax rates for the VAT are decided at the central level and are the same for all regions. On the other hand, decentralizing the collection of indirect taxation can in principle induce the regional governments to foster economic growth, as it expands the tax base and so increases their income. This idea follows the arguments developed in the so-called “Second generation fiscal federalism”, that emphasizes the importance of fiscal incentives for producing local economic prosperity.¹⁸ The lower reliance on transfers and higher reliance on own resources could nudge the regions into introducing measures destined to expand their tax base and create economic

¹⁸For a survey in the topic, see for example Weingast (2009).

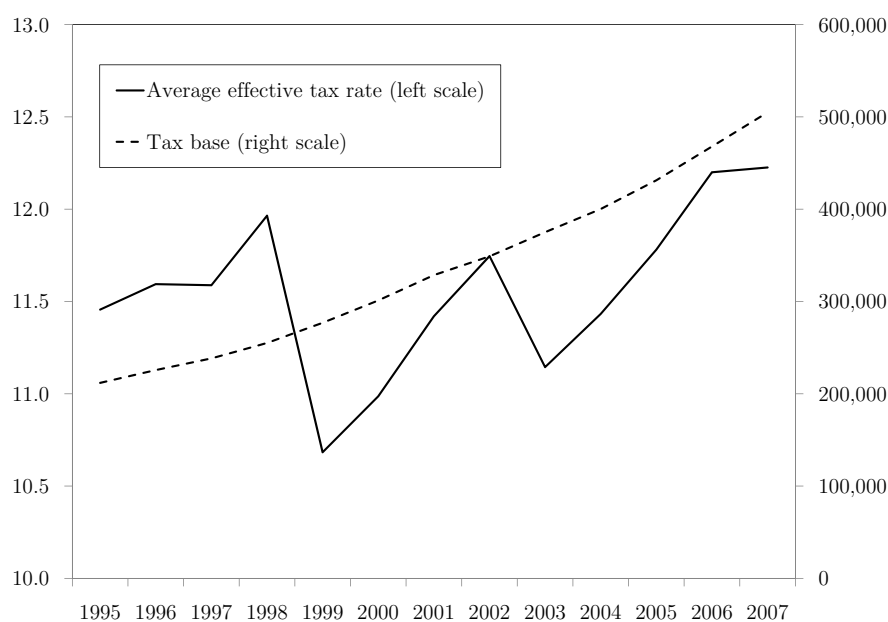


Figure 3.12: Average effective tax rate and tax base (in millions of euros) for the labor income tax. Spain 1995-2007. Source: Agencia Tributaria.

growth. In principle, this mechanism would explain the fact that the decentralization of indirect taxation only turns significant after some years, as the potential changes introduced by the regions to foster growth and so reap the benefits of an increased tax base would take time to materialize.

Finally, the decentralization of direct taxation registers a significant, positive effect on regional income growth. The key in this case is that the decentralization of the income tax included the possibility for regional government to modify legislative aspects, such as tax rates, brackets and deductions, affecting their tax scheme design as well as total taxes collected. Therefore, the decentralization of direct taxation implied not only a change in the way a region finances itself, as the decentralization of indirect taxation, but also opened up the possibilities of the regions to conduct fiscal policy.

On top of the possible effect on the incentives to increase the tax base, the decentralization of direct taxation plus giving legislative powers on taxation to the regions opens a new channel of influence of decentralization on growth through tax competition. Allowing regional governments to (partially) set the tax rate can promote tax competition between jurisdictions, resulting in lower tax rate and promoting growth. A modified version of this theory can be traced back to the work of Brennan and Buchanan (1980), where the authors used the tax competition argument to partly construct the “Leviathan hypothesis”: more-decentralized government structures should be smaller, in terms of government spending, relative to the size of the economy.

In fact, we find some evidence of tax competition, as can be seen in figure (3.12). The average effective tax rate naturally grows over time as the nominal tax base grows and workers move into higher brackets. There are two big drops in the average effective tax rate that match exactly the years were the income tax was reformed to allow the regions to keep part of the tax collected on their territory. Unfortunately, the available data does not allow to disentangle the part of the drop in the average effective tax rate due to the decrease in the tax rate induced by the central government from the part induced by the regional governments.

The results discussed here also tie nicely with the analysis in the previous section. In the first part we saw that more decentralized countries tend to have larger multipliers corresponding to their fiscal policy. Similarly, the second part showed how moving into more decentralized political structures produces a positive effect on the economic growth of the regions involved. The main conclusion from the paper is, therefore, in line with the classical theory of fiscal decentralization: fiscal policy becomes more effective when it allows regions to manage policies tailored to their citizens. At the same time, it is important to notice that the focus of this study were the short term dynamics and transition mechanism for fiscal policy. Potential long run effects that include a more detailed investigation of regional debt dynamics, for example, are left for further research.

3.4 Conclusion

This paper evaluates empirically the effects from decentralizing fiscal policy. The first section comprises a cross-country analysis of 11 European countries that differ substantially in their degree of decentralization. We find that more decentralized countries, i.e. countries in which regions have larger fiscal competences relative to the central government, tend to have larger fiscal multipliers. We interpret this as evidence in favorable gains from regionally tailored fiscal policy. We also provide evidence from a case study of Spain, where we exploit time variation in decentralization. These decentralization changes were orthogonal to economic conditions, and can therefore be used to measure the direct impact of decentralization on output growth. We find economically large and statistically significant positive effects from the decentralization of direct taxation, and, to a lesser extent, from indirect taxation and government spending on GDP growth. Part of the positive effect on output growth can be attributed to a reduction in taxes collected, and is likely to be attributed to tax competition between provinces. In line with our cross-country analysis, the results from the case study of Spain reinforce the evidence for efficiency gains through (regionally) tailored fiscal policy.

3.A Appendix: Decentralization Measures and Cross-country Comparison

Alternative Measures of Decentralization

Figure 3.13: Decentralization Index: Regional share of spending including transfers in %. Averages along with highest and lowest observation. Calculation based on yearly data 1996 - 2011.

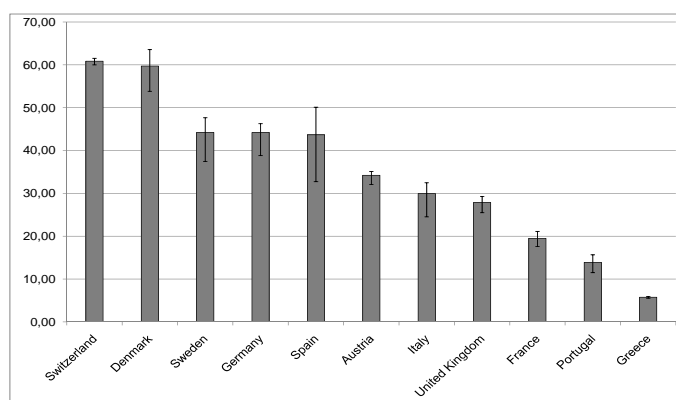
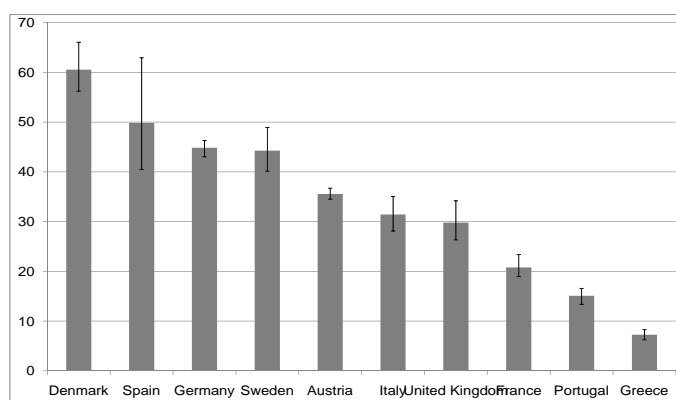


Figure 3.14: Decentralization Index: Regional share of taxation including transfers in %. Averages along with highest and lowest observation. Calculation based on yearly data 1996 - 2011.



Decentralization and key economic variables

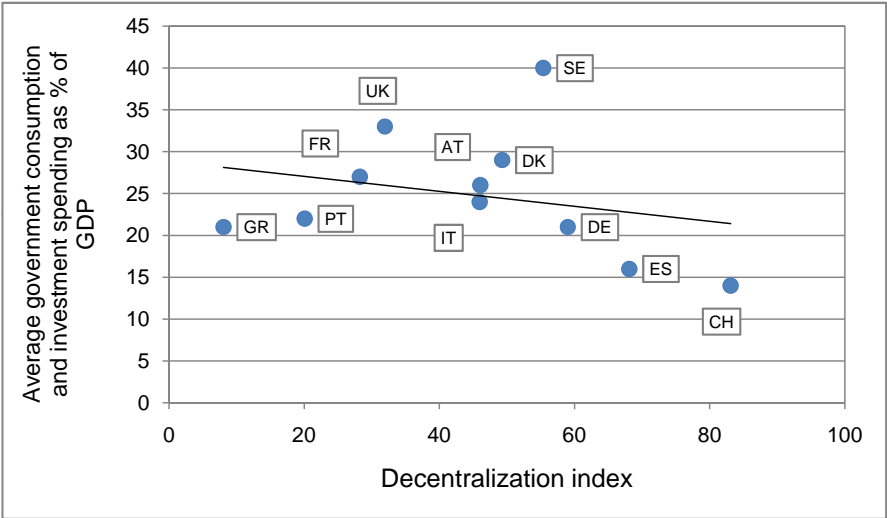


Figure 3.15: Decentralization Index vs. Size of Government Spending.

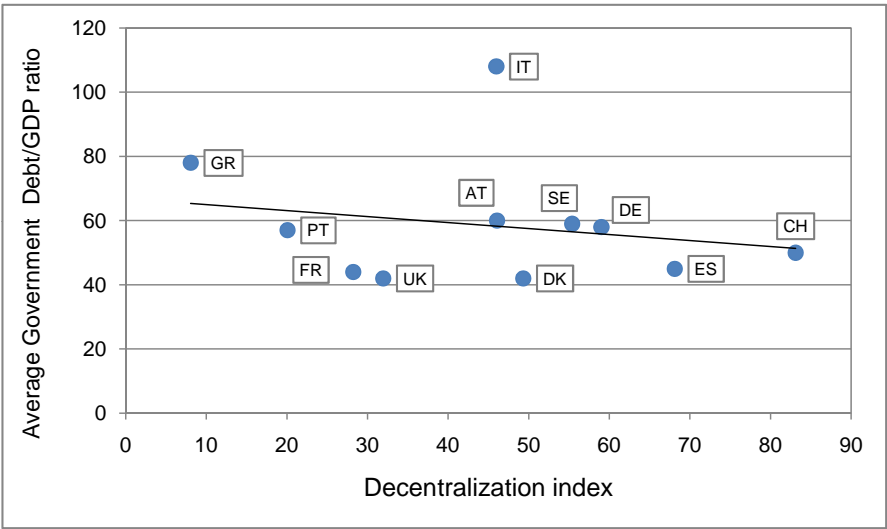


Figure 3.16: Decentralization Index vs. Government Debt

	Spending	Con. Spending	Inv. Spending	Revenues
Austria	0.233 (0.0193)	0.209 (0.0111)	0.0247 (0.00907)	0.500 (0.0268)
Denmark	0.288 (0.0121)	0.251 (0.00754)	0.0371 (0.00792)	0.541 (0.0180)
France	0.282 (0.0109)	0.240 (0.00938)	0.0425 (0.00265)	0.487 (0.0123)
Germany	0.216 (0.0116)	0.192 (0.0103)	0.0241 (0.00304)	0.413 (0.0270)
Greece	0.234 (0.00743)	0.206 (0.0106)	0.0321 (0.00468)	0.352 (0.0669)
Italy	0.232 (0.0161)	0.208 (0.0134)	0.0246 (0.00328)	0.421 (0.0367)
Portugal	0.226 (0.0122)	0.183 (0.0128)	0.0432 (0.00566)	0.345 (0.0901)
Spain	0.197 (0.0190)	0.159 (0.0124)	0.0384 (0.00813)	0.361 (0.0289)
Sweden	0.366 (0.0347)	0.322 (0.0323)	0.0435 (0.00474)	0.581 (0.0248)
Switzerland	0.148 (0.00816)	0.115 (0.00604)	0.0327 (0.00321)	0.327 (0.0290)
UK	0.239 (0.0202)	0.221 (0.0227)	0.0174 (0.00393)	0.374 (0.0309)

Table 3.6: Spending and Revenues over GDP by country. Spending denotes the sum of government consumption and investment expenditure; Cons. and Inv. Spending stand for consumption and investment spending, respectively. Standard deviations in parenthesis. Quarterly observations, 1980 - 2007.

	GDP growth (pct.)	(X+IM)/GDP	Population	Gross debt/GDP (pct.)
Austria	1.946 (1.154)	2.951 (0.626)	7.853 (0.252)	63.27 (4.403)
Denmark	1.961 (1.656)	2.444 (0.261)	5.232 (0.117)	54.88 (15.43)
France	1.656 (1.177)	1.821 (0.258)	57.46 (2.302)	46.07 (15.77)
Germany	1.705 (1.487)	2.304 (0.451)	80.11 (2.450)	56.35 (8.688)
Greece	1.704 (2.528)	1.417 (0.133)	10.46 (0.494)	72.77 (27.80)
Italy	1.665 (1.262)	1.782 (0.252)	56.91 (0.464)	101.9 (9.609)
Portugal	2.557 (2.439)	2.348 (0.195)	10.14 (0.219)	55.73 (5.078)
Spain	2.380 (1.769)	1.620 (0.338)	39.88 (2.046)	44.95 (13.16)
Sweden	2.094 (1.776)	2.332 (0.421)	8.699 (0.276)	56.87 (11.20)
Switzerland	1.199 (1.796)	2.618 (0.315)	6.883 (0.383)	48.63 (12.81)
United Kingdom	2.412 (1.602)	1.785 (0.103)	57.99 (1.404)	41.72 (4.876)

Table 3.7: Descriptive statistics for additional determinants of fiscal multipliers by country, yearly observations 1980 - 2007. Population is in million. Standard deviation in parenthesis. Source: International Monetary Fund, World Economic Outlook Database and eLibrary.

Bootstrap algorithm for Averaged Multipliers

This section describes the bootstrap algorithm used for obtaining the confidence bands for the average multiplier across a specific country group.

1. Draw n nonparametric impulse responses (IRFs) for each country individually. Each of the n IRFs is obtained by bootstrapping the residuals and reestimating the VAR. In each replication, the residuals drawn have the same time index t for each country in a respective group.
2. Standardize each impulse response by dividing by the size of the fiscal shock and the average fiscal variable to GDP ratio.
3. Draw (with replacement) one impulse response function for each country in the respective group. These draws are not independent across countries, but ensure that

the impulse response function of each country are based on the same time index t residuals. Then compute the average impulse response for each horizon.

4. Repeat step 3 m times.
5. Compute Hall (2013)-percentiles from m draws obtained above (as described e.g. in Lütkepohl 2005).

Robustness: Results from 2 Variable VARs

Figure 3.17: Cumulative Spending Multipliers, average across country group. The multipliers describe the output response to a positive spending shock. Estimates are based on a 2-variable VAR including output and government spending.

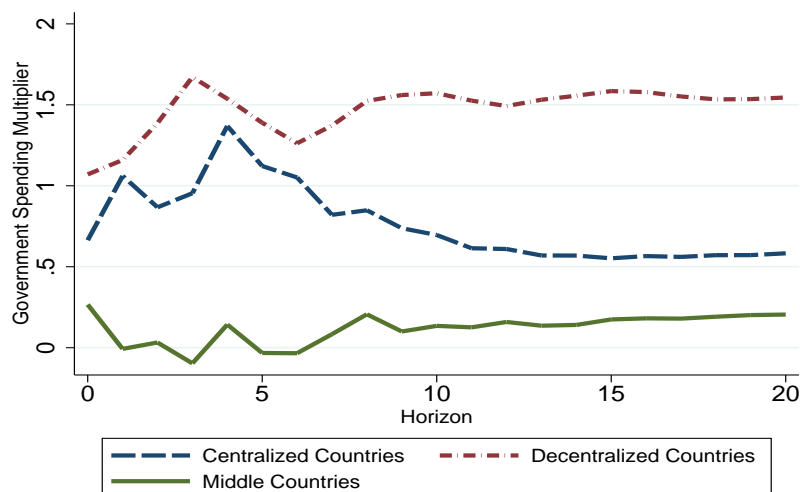


Figure 3.18: Cumulative Revenue Multipliers, average across country group. The multipliers describe the output response to a negative tax shock. Estimates are based on a 2-variable VAR including output and government revenues.

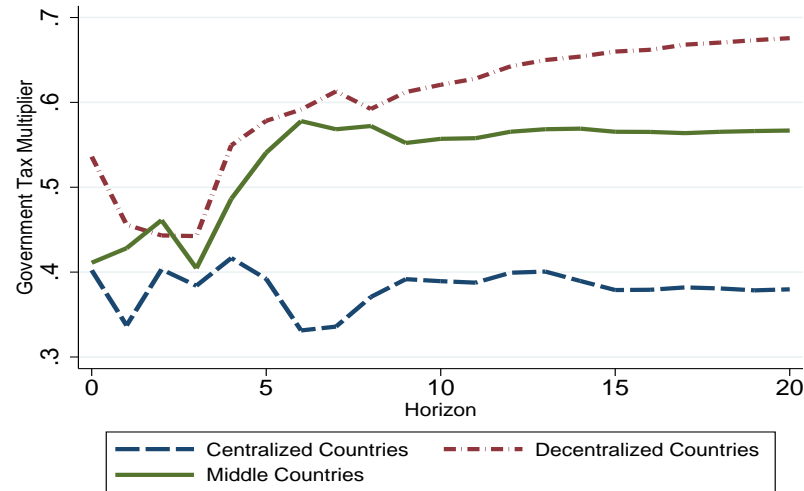


Figure 3.19: Cumulative Revenue Multipliers, average across country group. The multipliers describe the output response to a positive government consumption shock. Estimates are based on a 2-variable VAR including output and government consumption spending.

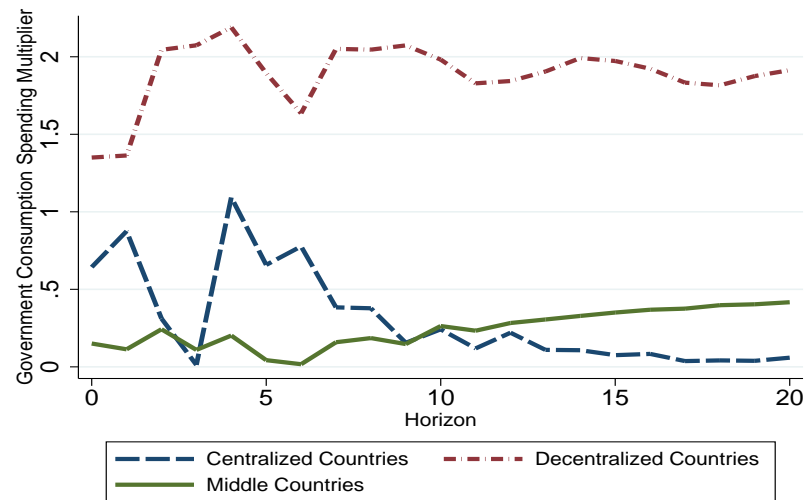
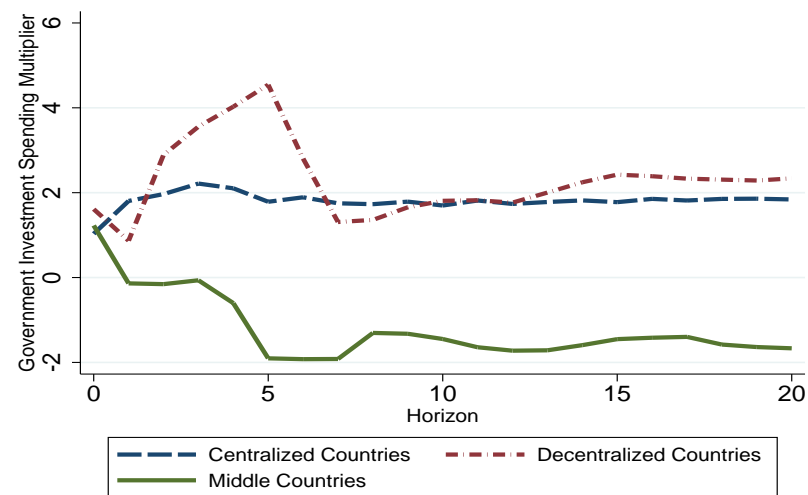


Figure 3.20: Cumulative Revenue Multipliers, average across country group. The multipliers describe the output response to a positive government consumption shock. Estimates are based on a 2-variable VAR including output and government investment spending.



3.B Appendix: Intranational Evidence

Justification of the narrative approach - Late vs. Early Adopters

Early adopters	Late adopters
Andalucía	Aragón
Canarias	Asturias
Catalunya	Islas Baleares
Galicia	Castilla-La Mancha
Navarra	Cantabria
País Vasco	Castilla y León
Comunitat Valenciana	Extremadura
	Comunidad de Madrid
	Murcia
	La Rioja

Table 3.8: Early adopters are regions that tended to implement fiscal reforms prior to the late adopters.

Regression results

	A	B	C	D
$\Delta Y(-1)$	0.3512*** (0.05866)	0.4205*** (0.05685)	0.4315*** (0.04804)	0.3196*** (0.07006)
Direct taxation	0.0083*** (0.00149)			0.0116*** (0.00205)
L.Direct taxation	0.0102*** (0.00228)			0.0129*** (0.00320)
L2.Direct taxation	0.0126*** (0.00280)			0.0118*** (0.00310)
L3.Direct taxation	0.0113*** (0.00218)			0.0078** (0.00273)
L4.Direct taxation	0.0055** (0.00221)			-0.0014 (0.00254)
L5.Direct taxation	0.0047 (0.00290)			-0.0023 (0.00295)
Indirect taxation		0.0017* (0.00086)		-0.0012 (0.00083)
L.Indirect taxation		0.0023*** (0.00049)		0.0002 (0.00103)
L2.Indirect taxation		0.0037*** (0.00078)		0.0016** (0.00073)
L3.Indirect taxation		0.0022** (0.00077)		0.0022** (0.00086)
L4.Indirect taxation		0.0022** (0.00077)		0.0039*** (0.00117)
L5.Indirect taxation		0.0007 (0.00127)		0.0028 (0.00168)
Spending			0.0015 (0.00095)	0.0006 (0.00106)
L.Spending			-0.0000 (0.00125)	-0.0013 (0.00157)
L2.Spending			0.0014 (0.00157)	0.0017* (0.00089)
L3.Spending			0.0027*** (0.00079)	0.0018* (0.00104)
L4.Spending			0.0023 (0.00136)	0.0007 (0.00150)
L5.Spending			-0.0014 (0.00171)	-0.0017 (0.00157)
Constant	0.0163*** (0.00132)	0.0172*** (0.00154)	0.0173*** (0.00137)	0.0166*** (0.00142)
R-squared	0.3151	0.2745	0.2441	0.3443
No. of obs.	267	261	306	260

Table 3.9: Effect of a decentralization shock on ΔY , the (log) growth rate of regional GDP. Standard errors are reported in parenthesis; *, ** and *** indicate significance at the .90, .95 and .99 level, respectively

References

- Auerbach, Alan J and Yuriy Gorodnichenko (2012). “Fiscal multipliers in recession and expansion. In: Fiscal Policy after the Financial crisis.” *University of Chicago press*, pp. 63–98.
- Baxter, Marianne and Robert G King (1993). “Fiscal policy in general equilibrium.” *The American Economic Review*, pp. 315–334.
- Blanchard, Olivier J and Roberto Perotti (2002). “An Empirical Characterization Of The Dynamic Effects Of Changes In Government Spending And Taxes On Output.” *The Quarterly Journal of Economics* 117.4, pp. 1329–1368.
- Bordignon, Massimo, Paolo Manasse, and Guido Tabellini (2001). “Optimal regional redistribution under asymmetric information.” *American Economic Review*, pp. 709–723.
- Brennan, Geoffrey and James M Buchanan (1980). *The power to tax: Analytic foundations of a fiscal constitution*. Cambridge University Press.
- Caldara, Dario and Christophe Kamps (2012). “The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers.” *FEDS Working Paper*.
- Corsetti, Giancarlo, Andre Meier, and Gernot J Müller (2012). “What determines government spending multipliers?” *Economic Policy* 27.72, pp. 521–565.
- Davoodi, Hamid and Heng-fu Zou (1998). “Fiscal decentralization and economic growth: A cross-country study.” *Journal of Urban economics* 43.2, pp. 244–257.
- De Mello, Luiz R (2001). “Fiscal decentralization and borrowing costs: the case of local governments.” *Public Finance Review* 29.2, pp. 108–138.
- Dellas, Harris, Klaus Neusser, and Manuel Wälti (2005). “Fiscal policy in open economies.” *manuscript, University of Bern*.
- Ebel, Robert D and Serdar Yilmaz (2002). *On the measurement and impact of fiscal decentralization*. Vol. 2809. World Bank Publications.
- Fatás, Antonio, Ilian Mihov, et al. (2001). “The effects of fiscal policy on consumption and employment: theory and evidence.” 2760.
- Feld, Lars P, Gebhard Kirchgässner, and Christoph A Schaltegger (2004). “Fiscal federalism and economic performance: evidence from Swiss cantons.” *Marburger volkswirtschaftliche Beiträge*.
- Giordano, Raffaella et al. (2007). “The effects of fiscal policy in Italy: Evidence from a VAR model.” *European Journal of Political Economy* 23.3, pp. 707–733.
- Hall, Peter (2013). *The bootstrap and Edgeworth expansion*. Springer Science & Business Media.
- Hemming, Richard, Michael Kell, and Selma Mahfouz (2002). “The effectiveness of fiscal policy in stimulating economic activity: A review of the literature.” *IMF working paper*.

- Ilzetki, Ethan, Enrique G Mendoza, and Carlos A Végh (2013). “How big (small?) are fiscal multipliers?” *Journal of Monetary Economics* 60.2, pp. 239–254.
- Lütkepohl, Helmut (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- Molero, Juan Carlos (2001). “Analysis of the decentralization of public spending in Spain.”
- Mountford, Andrew and Harald Uhlig (2009). “What are the effects of fiscal policy shocks?” *Journal of applied econometrics* 24.6, pp. 960–992.
- Oates, Wallace E (1985). “Searching for Leviathan: An empirical study.” *The American Economic Review*, pp. 748–757.
- Oates, Wallace E et al. (1972). “Fiscal federalism.” *Books*.
- Perotti, Roberto (2005). “Estimating the effects of fiscal policy in OECD countries.” *CEPR Discussion paper*.
- Perotti, Roberto (2007). “In search of the transmission mechanism of fiscal policy.” *National Bureau of Economic Research Cambridge, Mass., USA*.
- Price, Robert WR, Thai-Thanh Dang, and Yvan Guillemette (2014). “New Tax and Expenditure Elasticity Estimates for EU Budget Surveillance.”
- Ramey, Valerie A and Sarah Zubairy (2014). “Government spending multipliers in good times and in bad: evidence from US historical data.”
- Rodden, Jonathan (2002). “The dilemma of fiscal federalism: grants and fiscal performance around the world.” *American Journal of Political Science*, pp. 670–687.
- 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, pp. 763–801.
- Schaltegger, Christoph A and Lars P Feld (2009). “Are fiscal adjustments less successful in decentralized governments?” *European Journal of Political Economy* 25.1, pp. 115–123.
- Sharma, Chanchal Kumar (2006). “Decentralization dilemma: measuring the degree and evaluating the outcomes.” *The Indian Journal of Political Science*, pp. 49–64.
- Suárez-Pandiello, Javier (1999). “Fiscal federalism in Spain.” *Fiscal federalism in the European Union* 9, p. 219.
- Tenhofen, Jörn, Guntram B Wolff, and Kirsten H Heppke-Falk (2010). “The macroeconomic effects of exogenous fiscal policy shocks in Germany: a disaggregated SVAR analysis.” *Jahrbücher für Nationalökonomie und Statistik*, pp. 328–355.
- Thießen, Ulrich (2003). “Fiscal decentralisation and economic growth in high-income OECD Countries.” *Fiscal studies* 24.3, pp. 237–274.
- Weingast, Barry R (2009). “Second generation fiscal federalism: The implications of fiscal incentives.” *Journal of Urban Economics* 65.3, pp. 279–293.