# Three Essays in Economics

Dissertation zur Erlangung der Würde eines

Doktors der Staatswissenschaften

vorgelegt der Wirtschaftswissenschaftlichen Fakultät

der Universität Basel von

Daniel Müller

von

Langenbruck (Basel-Landschaft) und Deutschland

Winter-Industries GmbH

Berlin, 2013

Genehmigt von der Wirtschaftswissenschaftlichen Fakultät der Universität Basel auf Antrag von Prof. Dr. Aleksander Berentsen und Prof. Dr. Gabriele Camera.

Basel, den 5. April 2013

Der Dekan: Prof. Dr. George Sheldon

# Acknowledgments

Looking back at the time when I wrote my thesis, I somewhat like the metaphor of long and exhausting expeditions of famous explorers sailing to America and other undiscovered places. Stormy periods alternated with calm winds. After days of hard work, where one problem after another appeared, there were times where it was like walking on air. Sticking to my metaphor, I want to say thank you to all the colleagues and friends who have supported and accompanied me during my journey.

First of all, I want to thank my first advisor Aleksander Berentsen for all he has done for me. His profound knowledge and his numerous contributions in the field of monetary theory raised my interest for monetary theory and financial intermediation. Following his lectures on game theory, I learned that perfect markets and complete information are rather the exception than the rule and that incomplete markets and asymmetric information can change agents' actions and behaviors substantially. His paper on the economics of doping initiated my own contribution to the doping literature. Finally, I highly appreciated that he allowed me to conduct a study for the Swiss Federal Office of Energy (SFOE), which I ultimately incorporated as the third chapter of my thesis. It was a great advantage to gain profound insights into different economic fields during my thesis.

I would also like to thank my second advisor Gabriele Camera. His genuine interest in all kinds of economic and social questions makes his comments and advices highly valuable. The discussions we had about my work were always inspiring and I have taken his suggestions and recommendations to heart. I am highly indebted to Lukas Mohler for his friendship and his collaboration in the study on substitution elasticities. I have the impression that we *complement* each other very well. In this context, I also want to thank the SFOE for their financial support for our study on substitution elasticities and Nicole Mathys for her help and advice.

I wish to express my gratitude to Yvan Lengwiler, George Nöldeke and Benedikt von Scarpatetti for all the comments and advise they gave me over the years. The fruitful discussions and the exchange of ideas were very stimulating. A special thank goes to Patrik Ryff for his great support and the mutual exchange of ideas–not restricted to economics. I am grateful to Hermione Miller-Moser for correcting typos and improving the readability of this thesis considerably. Moreover, I want to thank all my colleagues and friends of the WWZ who supported me and made my time here so pleasant.

I am also indebted to my parents, brothers and sister. Having such a big family is advantageous for at least one reason: Someone always has time for you! My final thanks go to my love Gisèle. I deeply appreciate and value all her encouragement, kindness and love during the last years.

# Preface

This thesis contains three essays in different fields of economics. The first essay is about microfinance and inflation in low-income countries. The second essay examines doping behavior of heterogeneous athletes in an environment of private information. The third essay estimates substitution elasticities for Swiss manufacturing industries between production factors. The third essay was conducted for the Swiss Federal Office of Energy (SFOE) and is written in collaboration with Lukas Mohler. The first two essays are theoretical contributions, while the third essay is an empirical study. The essays cover research questions that I am especially interested in. In the following, I give a short introduction to each topic of my thesis and illustrate the importance of my research.

Chapter 1 analyses the effects of microfinance and inflation in low-income countries. The idea behind microfinance is the provision of financial services on a small scale to households that lack access to regular banks. In low-income countries, more than three-quarters of the population have no access to formal financial institutions (World Bank, 2012). The inadequate access that poor households have to financial services is held to be one of the major factors responsible for the serious inequalities and the lack of development that exist in low-income countries. Microfinance is a means for alleviating these problems, as it plays a prominent role in broadening the poor's access to financial services.

The success of microfinance has not gone unnoticed and is considered today to be an important tool for generating access to finance and reducing poverty. A large body of theoretical and applied literature exists on microfinance. Theoretical contributions have closely analyzed the mechanisms utilized in microfinance for reducing transaction costs and mitigating the asymmetric information problem. Numerous field studies have investigated the extent to which access to microfinance institutes increases the wealth of poor households. However, most empirical studies on microfinance neglect general equilibrium effects, as well as the monetary policy dimension in developing countries which is often characterized by high inflation rates. Monetary policy influences not only the inflation rate, but also the terms and conditions of saving and lending.

Our study is the first to analyze the effects of microfinance and inflation in a monetary general equilibrium model. To do so, we introduce a moral hazard problem into a monetary general equilibrium model with credit. Monetary search models are predestined to investigate this kind of research question, since they allow agents' monetary decisions to be explicitly modeled. At the same time, they are well suited for investigating the effects of different monetary policies at a macroeconomic level, taking general equilibrium effects into account. Finally, the prevailing challenges and market frictions that poor households face when applying for loans without possessing collateral can be introduced into the model.

Chapter 2 deals with the doping problem in sport competitions. From the 1920s onwards, individual sport federations began to restrict the use of doping substances. About forty years later, in 1966, the first doping tests were introduced in cycling and football. Regular controls were then adopted in most other professional sports. However, each federation had its own approach to fight doping, and the collaboration with other federations, governments and the International Olympic Committee (IOC) was not always successful. Ongoing doping cases and the Festina scandal during the Tour de France in 1998 provided impetus to improve the anti-doping campaign, and to standardize and coordinate the work of the relevant stakeholders. As a result of this, the World Anti-Doping Agency (WADA) was founded in 1999.

Today, all major sport federations have adopted the World Anti-Doping Code.[1] Every year the WADA releases a list of prohibited substances and methods. National agencies regularly conduct doping controls in the affiliated sport federations. Doping controls are carried out not only during competitions, but also during athletes' preparation periods and holidays. If an athlete is tested positive for a substance on the list or has been convicted of another violation of the anti-doping rules, the athlete is sanctioned and banned from

[1] See WADA (2009) for an extensive description of the World Anti-Doping Code. Information about the anti-doping history stems from the WADA homepage: http://www.wada-ama.org.

competition, usually for a period of two years. The objective of the WADA is to establish a doping-free environment. However, despite regular doping controls and severe sanctions in the case of detection, doping remains present in professional sports.

The main challenge of the doping prevention is that athletes' behavior cannot be directly observed. Moreover, the detection of violations is difficult and expensive. Asymmetric information arises at different levels. The existing doping literature has mainly focused on asymmetric information between athletes and the regulator about the use of performance-enhancing drugs. However, informational asymmetries between athletes about their actual capabilities are also relevant to an athlete's decision to take drugs. Our research question investigates how heterogeneity and an environment of private information affect athletes' doping behavior.

Chapter 3 analyses the substitutability patterns of manufacturing industries in Switzerland. The first oil price shock in 1973 led to major concerns about how firms and industries adapt to energy price shocks. A better understanding of how relative price changes affect the input mix and production costs of firms was required to estimate the overall impact of a price shock on GDP and prospective growth rates of the economy. Today, such questions have become relevant once again. Climate change mitigation policies which aim at increasing the efficiency of industrial production or promoting renewable energies, force firms to undertake adjustments in their production processes or to invest in new technologies. The introduction of carbon taxes will change the absolute and relative prices of different energy sources. Furthermore, Switzerland's decision to phase out nuclear power will entail major adjustments in energy provision and consequently affect energy prices as well.

Energy price changes are especially relevant for manufacturing industries. Production costs of firms with flexible production technologies will only increase marginally with an energy price increase, because they are able to substitute other production factors. If, on the other hand, the production technology is rigid, firms have greater problems in adjusting to energy price increases. In order to measure how industries adjust their use of production factors due to price changes, substitution elasticities are estimated. The larger the elasticities are, the higher is the degree of flexibility of the production technology. Moreover, economic substitution elasticities predict how the input mix changes due to price shocks.

The empirical literature has shown that the magnitude of elasticities varies significantly across countries, industries, and over time. Thus, adequate estimates of substitution elasticities are needed to assess the effects of concrete policy measures. However, no recent sector-specific estimates exist for Switzerland. We close this gap and estimate substitution elasticities for Swiss manufacturing industries between the production factors capital, labor, energy and material. Our focus is on how energy price increases affect the input mix of manufacturing industries and the implications for the production costs of firms.

# Contents

# List of Tables

# List of Figures

# Chapter I

## A General Equilibrium Analysis of Inflation and Microfinance in Developing Countries

**Abstract**

This paper analyses the welfare effects of microfinance and inflation in developing countries. Therefore, we introduce a moral hazard problem into a monetary search model with money and credit. We show how access to basic financial services affects households' decisions to borrow, to save and to hold money balances. The group lending mechanism of the microfinance institution induces peer monitoring, which in turn enables entrepreneurship. Our main result is that there exists an inflation threshold beyond which entrepreneurship collapses. We show that inflation affects the impact of microfinance on social welfare in a nonlinear way. The positive effect of microfinance is largest for moderate rates of inflation and drops substantially for inflation rates above the threshold.

| | |
|---|---|
| Keywords: | Microfinance, Moral Hazard, Group Lending, Peer Monitoring and Monetary Policy. |
| JEL Classification: | D82, E44, G21, O16. |

# 1 Introduction

It is generally accepted that better access to finance reduces poverty. Firstly, credit allows poor households to start small businesses, invest in new production machines, buy livestock, or simply to consume. Access to basic financial services facilitates consumption smoothing, payments for their children's education and wealth accumulation. Secondly, savings accounts pay interest and, thus, mitigate the negative effects of high inflation rates prevalent in developing countries. Finally, well functioning financial institutions have a positive effect on growth (see for instance Levine (2005) for a comprehensive literature survey on the relationship between finance and growth). However, in developing countries, the majority of households have no access to financial institutions. Empirical studies have shown that women, rural populations and, in particular, poor households are most concerned by this issue. In low income countries, 76 percent of adults have no account at a formal financial institution (Demirguc-Kunt and Klapper, 2012), while the access rate lies at 89 percent in high income countries.[1] Poor people have difficulty in general in gaining access to financial services. The World Bank reports that 77 percent of adults earning less than $2 a day are unbanked (World Bank, 2012). Comparing access to finance across countries and regions shows that large differences exist. To make things worse, inflation rates are on average much higher in developing countries than in industrialized countries.[2]

While governments in developing countries are well aware of the benefits of an efficient financial system, the question naturally arises as to why so little effort is made towards improving access to financial services. The reason is that basic banking services are complicated by a number of issues in developing countries. First, poor households have no valuable belongings, nor wealth that they could use as collateral for a loan. Second, transaction costs are especially high for small loans, and enforcement of repayments is difficult in countries with weak legal institutions. Third, asymmetric information between lenders and borrowers leads to principal-agent problems which may result in

[1]The terminology formal financial institutions in Demirguc-Kunt and Klapper (2012) includes banks, credit unions, cooperatives, post offices, and microfinance institutions.

[2]Easterly and Fischer (2001) analyze the effects of inflation on the poor.

a dwindling of the already weak credit market. Microfinance–the provision of financial services on a small scale–has shown that there are ways to overcome these problems and that lending to the poor is not a one-way street. By adopting new approaches such as peer-monitoring schemes or the village banking model, microfinance institutes can overcome the asymmetric information problem. These lending mechanisms show high repayment rates without requesting any collateral (Armendáriz and Morduch, 2010).

Microfinance started in the late 1970s and has expanded quickly over the last three decades. The original idea was to give credit to the poor. Over the last two decades, there has been a paradigm shift from highly subsidized microfinance institutions with limited outreach to a large-scale and financially sustainable microfinance industry (Robinson, 2001). Furthermore, since the early 2000s, many microfinance institutions have broadened their financial services and now also offer clients the possibility to open saving accounts (Matin et al., 2002). A leading example of a microfinance institute is the Grameen Bank in Bangladesh. The Grameen Bank and its charismatic founder Professor Muhammad Yunus were rewarded in 2006 with the Nobel peace price in appreciation of their achievements in poverty reduction and economic development in Bangladesh by providing the poor with access to finance. Today, over 2000 microfinance institutions exist all around the world and serve roughly one billion customers. They are mostly situated in developing countries, but are also to be found in high income countries.

The success of microfinance has not gone unnoticed and is considered today as an important tool for generating access to finance and reducing poverty. A large body of theoretical and applied literature exists on microfinance. Theoretical contributions have thoroughly analyzed the mechanism utilized in microfinance to reduce transaction costs and to mitigate the asymmetric information problem.[3] Numerous field studies have investigated to what extent access

[3] Stiglitz (1990) pioneered the work on group lending. Ghatak and Guinnane (1999) provide an extensive analysis of group lending extend the model to study four different agency problems and also discuss practical issues. Armendáriz (1999) analyses the problem of ex-post moral hazard.

to microfinance institutes increases the wealth of poor households.[4] However, most empirical studies that have analyzed the impact of microfinance neglect the monetary policy dimension in developing countries which is often characterized by high inflation rates. Moreover, general equilibrium effects on prices, caused by the financial intermediation of microfinance institutions, are often neglected.[5] In this paper, we intend to fill this gap by analyzing the effects of inflation on microfinance in a general equilibrium model. Using a model where money and credit are essential, allows us to derive the total welfare of an economy depending on the government's respective monetary policy and the outreach and efficiency of its microfinance institutes. Monetary policy plays an important role, as it directly determines the inflation rate and indirectly determines the market rates of borrowing and lending. Therefore, we use a monetary search model similar to Berentsen et al. (2007) to study the welfare effects of establishing a large-scale and sustainable microfinance institution in developing countries. To represent the agency problem between borrower and lender, we introduce a moral hazard problem in the style of Holmstrom and Tirole (1997). Moreover, we analyze the welfare implications for individual households.

We show that establishing sustainable microfinance institutions in developing countries allows poor households to increase their standards of living above the subsistence level. The reason is that former credit-constrained households are afterwards able to take out consumer loans or to invest in small businesses. Moreover, we show that the actual magnitude of the welfare impact of microfinance crucially depends on the prevailing monetary policy regime.

Our model discloses the relationship between the lending terms of microfinance and the monetary policy of the government. Higher money growth rates increase inflation and this in turn affects deposit and lending terms of the microfinance institution: On the one hand, depositors have to be compensated by a higher interest rate to encourage saving. The higher refinancing costs of

[4] Two studies of particular interest are Kaboski and Townsend (2012) and Banerjee and Duflo (2010). The first study evaluates the impact of the Million Baht Village Fund program in Thailand, and the second study runs a random field experiment, conducted in collaboration with an Indian microfinance institution.

[5] An exception is the paper by Kaboski and Townsend (2011), they develop a structural model to evaluate the impact of large-scale microcredit policy interventions.

the microfinance institution lead, in turn, to an increase in the lending rate. Moreover, higher inflation rates decrease real prices and output, which reduce the gains from trade (real balance effect). Entrepreneurs, who rely on external funding, are more affected by inflation than subsistence producers. Above a specific inflation threshold, entrepreneurship collapses and is displaced by subsistence production. Our numerical example shows that the positive impact of microfinance on social welfare is largest for moderate inflation rates, where entrepreneurship exists. However, for inflation rates above the threshold, the positive impact of microfinance drops substantially.

The structure of the article is as follows. Section 2 introduces the agents and describes the framework of the general equilibrium model with moral hazard and group lending. Section 3 presents the maximization problem of households in the two markets. In Section 4, the market outcome of the equilibrium and the optimal group lending contract are presented. In Section 5, we give a numerical example to present the impact and the welfare effect of microfinance in developing countries. Section 6 concludes.

## 2 The Basic Model

The model is based on Berentsen et al. (2007). It uses the standard Lagos-Wright structure, where time is discrete and every period consists of two subperiods. There exists a continuum [0,1] of infinitely living households and a single microfinance institution (MFI). In each period, households trade their produced goods at two sequentially opening markets. In subperiod A, households produce and trade the *production good*, and in subperiod B, the *general good*. Both goods are perishable and cannot be stored. We assume that the two markets are competitive and that no trading frictions exist. We will proceed by illustrating the structure of the economy and the characteristics of households. Then, we introduce fiat money and show how households can deposit money balances, as well as take out consumption loans from the microfinance institution. Subsequently, we illustrate in Sections 2.1 how entrepreneurs start a business and address the moral hazard problem with external funding. Section 2.2 shows the social planner problem.

Figure 1: Timeline

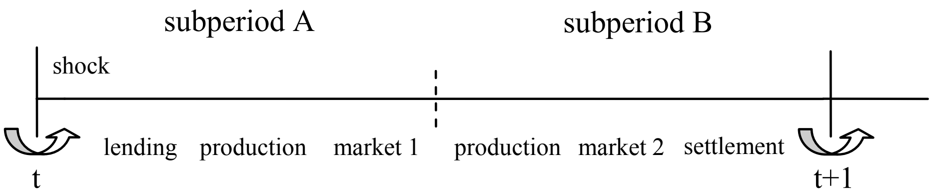



Figure 1 displays the timeline of the model for a representative period $t$. At the beginning of subperiod A, households are hit by a temporary preference and technology shock. With probability $1-n$, the household is a *buyer*; with probability $n\theta$, he is an *entrepreneur*; and with probability $n(1-\theta)$, he is a *producer*. Buyers can consume but cannot produce in the first market. In contrast, producers can produce but cannot consume in the first market. Producers may either produce in home production or work for entrepreneurs. Finally, entrepreneurs have the possibility to start a small enterprise with one employee. In contrast to producers, entrepreneurs cannot produce in home production. We assume that the production technology of the enterprise is superior to the home production technology. From now on, we will use the term *subsistence production* for the inferior home production technology. Buyers, entrepreneurs and subsistence producers trade in the first market, and subsequently the market closes. In subperiod B, all households can consume, produce and trade the general good in the second market.

Households have quasi-linear preferences, where $q_b$ ($q_s$) is the amount of the production good consumed (produced) in subperiod A.[6] In subperiod B, $x$ ($h$) is the amount of the general good consumed (produced), and all households have the same productivity. Equation (1) displays the utility function of a household. To account for the preference shock, the utility function of households is modeled with an indicator function. When a household is hit by the preference shock, then the indicator is one, otherwise it is zero. The utility

[6]The subscript $b$ stands for buyer and the subscript $s$ for subsistence producers. The production of an enterprise will be denoted by subscript $e$.

function of a representative household is

$$\mathcal{U}(q_b, q_s, x, h) = \mathbb{1}u(q_b) - (1 - \mathbb{1})c(q_s) + U(x) - h, \tag{1}$$

where $\mathbb{1}$ is the indicator, and $c(q_s)$ are the utility costs of producing the amount $q_s$ of good $q$ in subperiod A. The cost function $c(\cdot)$ is a convex function with respect to $q$, where $c(0) = 0$, $c'(\cdot) > 0$ and $c''(\cdot) > 0$. The utility function $u(\cdot)$ is a concave function with respect to $q$, where $u(0) = 0$, $u'(0) = \infty$, and $u''(\cdot) < 0$. Good $x$ can be consumed and produced in home production in subperiod B by every household, where $h$ are the utility costs. To ease the calculation, we assume that the production of the general good is linear.[7] The utility function $U(x)$ of the general good is a convex function.

We assume that households trade in anonymous goods markets. Households are not able to recognize former trading partners in future meetings. Hence, a role for a medium of exchange emerges. As the medium of exchange, we introduce fiat money. Money is essential in the markets of the production good and the general good, because there is no commitment and no record-keeping in the two markets. Access to financial services is solely feasible over the MFI. Households deposit money balances at the end of subperiod B and receive interest in the next subperiod B, if they do not withdraw their balances. Buyers use their deposits and can additionally take out small loans for consumption in subperiod A.[8] Entrepreneurs can issue risky debt with a special group lending contract. We will describe the group lending contract below when we show how entrepreneurs start businesses.

The central bank directly influences the amount of fiat currency by means of lump-sum transfers at the beginning of subperiod B. We assume that money grows at a specific but constant rate. The stock of money is indicated by $M$, and the money growth rate is $\gamma$, where $\gamma \geq 1$ ($M = \gamma M_{-1}$). Variables referring

[7]This is the standard assumption in Lagos-Wright models, that makes the model tractable. Actually, we could also assume that the cost function is nonlinear and that $U(x)$ is linear to find a solution. For further discussions, see e.g., Lagos and Wright (2005).

[8]For the saving accounts of the Grameen Bank, it was initially only possible to withdraw savings at an assigned time. In 2004, Grameen allowed customers to withdraw money at will. Thus, today, saving accounts are utilized like current accounts. This has led to large increases of Grameen's deposit portfolio. Since the end of 2004, the deposits of the Grameen Bank exceed their outstanding loans (Rutherford, 2006).

to the previous (subsequent) period are indexed by $-1$ $(+1)$. Households receive lump-sum transfers of $\tau M_{-1}$ from the central bank. To meet the targeted growth rate, it has to be the case that $\tau = \gamma - 1$. The real price of money in subperiod B is indicated by $\phi$. The assumption of a constant money growth rate implies that real money balances are time-invariant. Therefore, it is the case that $\phi/\phi_{+1} = M_{+1}/M = \gamma$. This is the standard way to model monetary policy in the Lagos-Wright framework.

## 2.1 Starting a Business and External Funding

Entrepreneurs have a business idea and start a small business with one employee. Every entrepreneur is matched with one producer and makes a take-it-or-leave-it wage offer that the producer accepts or declines. If the producer declines the wage offer, he produces in home production at the subsistence level. We assume that entrepreneurs have to pay wages in advance of the production (cash-in-advance). This assumption can be motivated through a lack of commitment. Thus, households are only willing to work if they are compensated for their utility costs beforehand. The production technology of the enterprise is superior to the home production technology. We assume that the employee's utility costs of producing the amount $q_e$ are $K + \tilde{c}(q_e)$. The cost function consists of a fixed setup cost term and a variable cost term. The variable cost term $\tilde{c}(\cdot)$ has the same features as $c(\cdot)$.

Entrepreneurship is subject to risk. Production is successful with a probability $\{\mu_h, \mu_l\} < 1$, depending on the behavior of the entrepreneur. If he shirks, the production is successful with the lower probability $\mu_l$, but the entrepreneur receives real, private benefit $B$. The difference between the two probabilities of success is denoted by $\Delta\mu$.[9] Agents can verify whether production was successful, but the behavior of entrepreneurs is private information and can only be revealed by monitoring. If the investment is financed externally, then the entrepreneur has an incentive to shirk. This is the standard moral hazard problem similar to (Holmstrom and Tirole, 1997). We assume

[9] We assume that the expected profit of a shirking entrepreneur, who could finance the production internally, is smaller than zero. The expected profit consists of sales if the production is successful (with probability $\mu_l$) plus the private benefit minus the wage costs for the employee. Moreover, we assume that $\tilde{c}(q)/\mu_h < c(q)$, for all $q > 0$.

that monitoring costs are proportional to the loan size and are denoted by $\delta_m$. For simplicity, we assume that monitoring entrepreneurs will always detect a shirking peer and that strong enough means exist to sanction shirking peers. Hence, if entrepreneurs monitor, shirking can be ruled out.

Let us shortly recapitulate the three obstacles that entrepreneurship faces: First, the already discussed cash-in-advance constraint for enterprises. Second, entrepreneurs are capital-constrained and need external funding. More precisely, their savings are insufficient for self-financing the business and they have no collateral. Finally, asymmetric information leads to a moral hazard problem. The MFI solves the agency problem through group-lending contracts with joint liability and enables entrepreneurship.

**Group Lending Contract**

We suppose that commercial loans are only available as group-lending contracts.[10] After the shock has been revealed, all entrepreneurs meet with the local branch of the MFI to contract for loans. They are divided into small groups of two.[11] A representative group consists of entrepreneurs $i$ and $j$. We assume that they are protected by limited liability. Thus, the MFI can only claim the returns of the project. Moreover, we assume joint liability, which means that entrepreneurs have to take responsibility for the repayment if the peer defaults. In our group lending contract, this means that borrower $i$ pays interest rate $i_s$ if borrower $j$ repays his loan, and $i_f$ if the peer defaults (The subscript $s$ stands for success and $f$ for failure). Typically, in this kind of contract the interest rate $i_f$ is greater than $i_s$.[12]

## 2.2 The Social Planner Allocation

The mission of the planner is to choose the quantities buyers consume and producers produce of the production good and the general good. We assume

[10] MFIs offer larger loans that households can use to finance a marriage, a funeral and, especially, to start small enterprises. These loans can only be used for the stated purpose and the requirements are higher. For example, the Grameen Bank offers different loan contracts, ranging from housing loans to special investment loans intended for entrepreneurs (Rutherford et al., 2004).

[11] For instance, the Grameen Bank lends money to groups of 5 (called kendras).

[12] In practice, institutions not only use group lending mechanisms with joint liability. Armendáriz and Morduch (2000) describe other mechanisms such as regular repayment schedules or non-refinancing threats.

that the planner sees whether entrepreneurs behave or shirk and is able to force them to behave. The social planner maximizes the aggregated lifetime utility of the households subject to the feasibility constraints. It is obvious that the planner decides that all entrepreneurs start a business, since the returns are higher than under subsistence production. The optimization problem of the social planner is

$$\mathcal{W} = \frac{1}{1-\beta}\{(1-n)u(q_b) - n\theta K - n\theta\tilde{c}(q_e) - n(1-2\theta)c(q_s) + U(x) - h\}, \quad (2)$$

where welfare consists of net utility of subperiod A plus net utility of subperiod B. The planner chooses the quantities produced and consumed in the two subperiods. Equation (3) displays the optimal amount consumed ($q_b^*$) and produced ($q_e^*$, $q_s^*$) in subperiod A.

$$u'(q_b^*) = \tilde{c}'(q_e^*) = c'(q_s^*). \quad (3)$$

In subperiod B, every household produces and consumes the amount $x^*$ of the general good such that $U'(x^*) = 1$. In the optimal allocation, aggregate production has to equal aggregate demand for both goods ($q$, $x$). Furthermore, the social planner forces entrepreneurs to behave. Thus, the production is successful with probability $\mu_h$. The production market in subperiod A clears if Equation (4) holds.

$$(1-n)q_b^* = n\theta\mu_h q_e^* + n(1-2\theta)q_s^*. \quad (4)$$

## 3 The Goods Markets

The expected lifetime utility of a household can be specified in a recursive way by value functions. In particular, $V(\cdot)$ denotes the value function at the beginning of subperiod A, and $W(\cdot)$ denotes the value function at the beginning of subperiod B. The ex-ante value function $V(d)$ of a representative household at the beginning of period $t$ with deposits $d$ is given by

$$V(d) = (1-n)\underbrace{\{u(q_b) + W(0, l_b, d + l_b - pq_b)\}}_{\text{buyer}} + n\theta\underbrace{\{-\phi\delta_m l_j + E\{W\}\}}_{\text{entrepreneur}} \quad (5)$$
$$+n\theta\underbrace{\{-K - \tilde{c}(q_e) + W(d, 0, w)\}}_{\text{employee}} + n(1-2\theta)\underbrace{\{-c(q_s) + W(d, 0, pq_s)\}}_{\text{subsistence producer}}.$$

To fully understand the value function $V(d)$, we give a short description of the four states. First, with probability $1-n$, the household is a buyer. The term in the curly brackets indicates the utility of a buyer consuming the amount $q_b$ of the production good plus the continuation value of a buyer in subperiod B. Second, with probability $n\theta$, the household is an entrepreneur. The term in curly brackets indicates the monitoring costs (which depend on the peer's loan size) plus the ex-ante expected continuation value of an entrepreneur. The continuation value depends on the outcome of his own production and the peer's production. Third, with probability $n\theta$, the household is an employee. The term in curly brackets indicates the disutility of producing the amount $q_e$ plus the continuation value of an employee in subperiod B. Finally, with probability $n(1-2\theta)$, the household is a subsistence producer. The term in curly brackets indicates the disutility of producing the amount $q_s$ plus the value function of a producer with deposits $d$ and income $pq_s$ in subperiod B.

Equation (6) shows the expected value of entering subperiod B of entrepreneur $i$, given that $i$ and $j$ behave. The expected continuation value $W(\cdot)$ is the weighted sum of three possible outcomes: In the first outcome, the production is a failure. In this case, entrepreneur $i$ defaults and enters subperiod B with no income. In the second outcome, the production of entrepreneur $i$ is successful and at the same time entrepreneur $j$ repays his loan. In this case, entrepreneur $i$ only pays for his own obligations and makes a large profit. In the third outcome, entrepreneur $i$ is successful, but entrepreneur $j$ defaults. In this case, entrepreneur $i$ not only has to come up for his own obligation, he has also to repay part of borrower $j$'s loan. Entrepreneur $i$ enters subperiod B with a lower net profit than in the second case, which is indicated by $\Pi_l < \Pi_h$. The business of entrepreneur $i$ ($j$) is successful with probability $\mu_{h,i}$ ($\mu_{h,j}$).

$$E\{W\} = \underbrace{\mu_{h,i}\mu_{h,j}W(0,l_i^p,\Pi_h)}_{j \text{ repays}} + \underbrace{\mu_{h,i}(1-\mu_{h,j})W(0,l_i^p,\Pi_l)}_{j \text{ defaults}} + \underbrace{(1-\mu_{h,i})W(0,0,0)}_{i \text{ defaults}}. \tag{6}$$

To find the equilibrium, we start with the equilibrium conditions of the general goods market and solve backwards to find the equilibrium in the production market.

## 3.1 The General Goods Market (Subperiod B)

In subperiod B, households consume and produce the general good at home. The amount consumed is denoted by $x$, and $h$ denotes the produced amount. Households discount time with $\beta \in (0,1)$. Households enter subperiod B with heterogeneous portfolios of deposits ($d$), loans ($l$) and cash ($m$) and maximize the value function with respect to $x$, $h$ and $d_{+1}$. The maximization problem of a household entering subperiod B with the portfolio $(d,l,m)$ is

$$\begin{aligned} W(d,l,m) &= \max_{x,h,d_{+1}} \{U(x) - h + \beta V_{+1}(d_{+1})\} \\ \text{s.t.} \quad x - h &= \phi[(1+i_d)d + m + \tau M_{-1} - d_{+1} - (1+i)l], \end{aligned} \tag{7}$$

where $d_{+1}$ are deposits households place on the MFI for the subsequent period. The interest rate $i$ on loans depends on whether the household was a buyer ($i_d$) or a successful entrepreneur ($i_s$, $i_f$) in subperiod A. All values are stated in real terms. Households have to choose $x$, $h$ and $d_{+1}$, thereby satisfying the intertemporal budget constraint. The left-hand side of the budget constraint is consumption $x$ less the amount $h$ produced of the general good. The right-hand side consists of deposits of period $t$ charged with interest, the lump-sum transfer of the central bank less the deposits for period $t+1$, and loans charged with the respective interest. Substituting the budget constraint for $h$ gives

$$\begin{aligned} W(d,l,m) &= \max_{x,d_{+1}} \{U(x) - x + \beta V_{+1}(d_{+1}) \\ &\quad + \phi[(1+i_d)d + m + \tau M_{-1} - d_{+1} - (1+i)l]\}. \end{aligned}$$

The optimal quantities of good $x$ and deposits $d_{+1}$ for the subsequent period follow from the first order conditions:

$$U'(x) = 1, \tag{8}$$

$$\beta V'_{+1}(d_{+1}) = \phi. \tag{9}$$

The marginal value of deposits has to be equal to $\phi/\beta$ in equilibrium. The envelope condition for private saving is: $W_d = \phi(1 + i_d)$. The envelope condition of borrowing money for consumption is: $W_l = -\phi(1 + i_d)$. And lastly, the envelope condition of holding money is: $W_m = \phi$. All households will enter the next period with the same amount of deposits. This implies that at the beginning of the subsequent period, the money holdings are degenerate, and the liability side of the MFI's balance sheet is equal to the aggregate of all households' deposits $d_{+1}$. The general goods market serves to simplify calculations, since we do not have to keep track of the history of households' deposits.

## 3.2 The Production Goods Market (Subperiod A)

At the beginning of subperiod A, the preference shock determines whether households are buyers, producers or entrepreneurs. In the following, we will present the optimization problem for each group.

**Buyers**

Buyers choose how much to demand of good $q$, taking prices as given. For their expenses, they use their deposits and in addition have the possibility to take out consumption loans ($l_b$) from the MFI. The optimization problem of a representative buyer is:

$$\max_{q_b, l_b} \quad \{u(q_b) + W(0, l_b, d + l_b - pq_b)\}, \tag{10}$$

$$\text{s.t.} \quad pq_b \leq d + l_b, \tag{BC}$$

$$l_b \leq \bar{l}, \tag{LC}$$

where the budget constraint states that households can dispense up to the sum of deposits $d$ and the loan $l_b$. The loan constraint states that the buyer can borrow up to the limit $\bar{l}$. We will assume that buyers have no possibility to default and that the MFI can reclaim consumption loans without costs; thus, the (LC) is not binding. If this is the case, buyers optimally choose $q_b$ such that the following equation is satisfied:

$$u'(q_b) = \phi p(1 + i_d). \tag{11}$$

For the detailed derivation with the first-order conditions, see Appendix A.1.

**Producers**

Subsistence producers choose the amount $q_s$ that maximizes profit, thereby taking as given the price $p$. Producers incur utility costs $c(q_s)$ for producing the amount $q_s$. The optimization problem of a representative subsistence producer is:

$$\max_{q_s} \quad \{-c(q_s) + W(d, 0, pq_s)\}. \tag{12}$$

Because real balances enter the value function of submarket B in a linear fashion, the optimization problem of subsistence producers can be stated as

$$\phi \Pi_s = \max_{q_s} \{\phi p q_s - c(q_s)\}, \tag{13}$$

where real profit $\phi \Pi_s$ depends on the produced amount $q_s$ and the real price $\phi p$. Assuming a convex cost function gives the standard solution where producers set their marginal costs equal to the real price. Equation (14) displays the first order condition that maximizes the profit of a household at the subsistence level for a given real price.

$$c'(q_s) = \phi p. \tag{14}$$

**Entrepreneurs**

Entrepreneurs take out loans from the MFI and start small enterprises that produce with superior technology. To mitigate the agency problem between lender and borrower, the MFI offers group lending contracts with joint liability. In the following, we will present the optimization problem of a representative group with entrepreneurs $i$ and $j$ that takes as given the lending mechanism of the MFI. We assume that both entrepreneurs monitor each other and show afterwards how the MFI's group lending contract has to be designed to be incentive compatible. The optimization problem of entrepreneur $i$ when he monitors $j$ is:

$$\max_{l_i,w,q_e} E(\Pi) = \mu_{h,i}[pq_e - \mu_{h,j}(1+i_s)l_i - (1-\mu_{h,j})(1+i_f)l_i] - d - \delta_m l_j, \quad (15)$$

$$\text{s.t.} \quad d + l_i \geq w, \quad \text{(FC)}$$

$$w \geq K/\phi + \tilde{c}(q_e)/\phi + \Pi_s. \quad \text{(PC)}$$

With probability $\mu_{h,i}$, the project is successful and $i$ sells the amount $q_e$ at the production market. The actual interest rate that entrepreneur $i$ has to pay depends on whether $j$ repays his loan or not. With probability $\mu_{h,j}$, entrepreneur $j$ repays his loan and $i$ has to pay interest rate $i_s$. With probability $1-\mu_{h,j}$, entrepreneur $j$ defaults and $i$ has to pay interest rate $i_f$. The last term indicates the monitoring costs. Entrepreneur $i$ has to satisfy two constraints: First, the sum of deposits and the loan has to be greater or equal to the wage. Second, the wage has to be greater or equal to the disutility costs of producing $q_e$ plus the outside option of the employee. The outside option is production at the subsistence level, which achieves a profit of $\Pi_s$. See Section A.3 for the optimization problem of a subsistence producer. The participation constraint is satisfied if the wage is greater than the disutility of producing $q_e$ plus the foregone profit $\Pi_s$. Finally, the entrepreneur will only start the business if the expected profit is greater than his own outside option–principal and interest on deposits. The Lagrangian of the profit maximization problem with the two Lagrange multiplier $\lambda_l$ and $\lambda_w$ is

$$\begin{aligned} \mathrm{L}(q_e, l_i, w) &= \mu_{h,i}pq_e - \mu_{h,i}\mu_{h,j}(1+i_s)l_i - \mu_{h,i}(1-\mu_{h,j})(1+i_f)l_i - d - \delta_m l_j \\ &\quad -\lambda_l[w - d - l_i] - \lambda_w[\frac{K}{\phi} + \frac{\tilde{c}(q_e)}{\phi} + \Pi_s - w]. \end{aligned}$$

The first-order conditions are:

$$\begin{aligned} q_e: &\quad \phi\mu_{h,i}p &= \lambda_w \tilde{c}'(q_e), \\ l_i: &\quad \lambda_l &= \mu_{h,i}[\mu_{h,j}(1+i_s) + (1-\mu_{h,j})(1+i_f)], \\ w: &\quad \lambda_l &= \lambda_w, \\ \lambda_l: &\quad l_i &= w - d, \\ \lambda_w: &\quad w &= K/\phi + \tilde{c}(q_e)/\phi + \Pi_s. \end{aligned} \quad (16)$$

Combining the second and the third FOC gives $\lambda_w = \mu_{h,i}[\mu_{h,j}(1+i_s) - (1-\mu_{h,j})(1+i_f)]$. The fourth and the fifth FOC are the standard loan and wage constraints which have to hold with equality if the interest rate is greater than zero and the entrepreneur maximizes his profit. Substituting the Lagrange multiplier $\lambda_w$ in the first FOC and canceling $\mu_{h,i}$ on both sides gives the following condition which has to hold if entrepreneurs maximize profit:

$$[\mu_{h,j}(1+i_s) + (1-\mu_{h,j})(1+i_f)]\tilde{c}'(q_e) = (1+\bar{i})\tilde{c}'(q_e) = \phi p. \tag{17}$$

The term in the squared brackets is the expected interest rate an entrepreneur has to pay if production is successful. It depends on the two interest rates and also on the behavior of entrepreneur $j$.

# 4 Equilibrium

In this section, we assume that the microfinance institution maximizes social benefit and is not profit-oriented. But in contrast to a social planner, the institution cannot force households to behave. Moreover, outstanding credits have to be fully backed by deposits (no external sourcing). Furthermore, financial operations have to be sustainable, as we suppose that the MFI receives no subsidies. Therefore, the MFI offers group lending contracts with expected returns that are equal to the deposit rate. Even though the returns from specific groups are stochastic, aggregated returns of the MFI are fully predictable. The reason is that production failures are uncorrelated, and the law of large numbers applies.

**The Optimal Group Lending Contract**

The MFI has to design the group lending contract with joint liability such that entrepreneurs have incentives to monitor their peers.[13] Remember that an entrepreneur will behave if the peer monitors, since a detected entrepreneur would be punished by severe social sanctions. Hence, monitoring induces good

[13] In reality, collusion between entrepreneurs can be a serious threat for the success of group lending. However, the consideration of collusion is beyond the scope of our analysis and we therefore assume that entrepreneurs do not collude. See e.g., Laffont and Rey (2003) on collusion and group lending.

behavior, and production of the peer is successful with probability $\mu_h$. Assume that entrepreneur $i$ is monitored by the peer. It is optimal for entrepreneur $i$ to monitor if

$$\mu_{h,i}[pq_e - \mu_{h,j}(1+i_s)l_i - (1-\mu_{h,j})(1+i_f)l_i] - d - \delta_m l_j \geq \mu_{h,i}[pq_e - \mu_{l,j}(1+i_s)l_i - (1-\mu_{l,j})(1+i_f)l_i] - d.$$

In a symmetric equilibrium $l_i = l_j = l$. Then, it follows that the incentive constraint to monitor is satisfied if

$$\mu_h \Delta\mu(i_f - i_s) \geq \delta_m. \tag{18}$$

**Zero Profit Condition of the MFI**

The MFI has to pay interest rate $i_d$ on deposits. Thus, the expected return from the group lending contract has to be large enough. To break even, the two interest rates $i_s$ and $i_f$ have to satisfy the following condition:

$$2\mu_h^2(1+i_s) + 2\mu_h(1-\mu_h)(1+i_f) = 2(1+i_d), \tag{19}$$

where the first term of the left-hand side is gross repayment of the group if both households are successful (which occurs with probability $\mu_h^2$), and the second term is gross repayment if only one household is successful (which occurs with probability $2\mu_h(1-\mu_h)$). The right-hand side are the gross deposit costs the MFI has to pay. Entrepreneurs' loans and deposits have been normalized. To derive the interest rate $i_s$, we assume that Equation (18) is satisfied with equality and substitute for $i_f$. We obtain

$$\mu_h^2(1+i_s) + \mu_h(1-\mu_h)(1+i_s+\frac{\delta_m}{\mu_h\Delta\mu}) = 1+i_d. \tag{20}$$

**Business Funding**

Entrepreneurship exists if two constraints are satisfied: On the one hand, the MFI has to satisfy the incentive constraints of the entrepreneurs, and it has to respect the limited liability clause. The MFI will only give credit if profits are greater than the repayment obligation in the event that the peer defaults.

$$pq_e^* - (1 + i_f)[K/\phi + \tilde{c}(q_e^*)/\phi + \Pi_s - d] \geq 0. \tag{21}$$

On the other hand, entrepreneurs will only start the business if expected profits are greater than the principal and interest on deposits. Entrepreneurs only take out loans if the following condition is satisfied:

$$\begin{aligned} &\mu_h[pq_e^* - (1 + \bar{i})(K/\phi + \tilde{c}(q_e^*)/\phi + \Pi_s - d)] \\ &-\delta_m[K/\phi + \tilde{c}(q_e^*)/\phi + \Pi_s - d] \geq (1 - i_d)d. \end{aligned} \tag{22}$$

Whether the former or the latter constraint is more restrictive depends on the parameterization of the monitoring costs and the difference between $i_f$ and $\bar{i}$. Usually, Equation (22) is more restrictive. Furthermore, above the threshold exists a small range, where only part of the entrepreneurs are active. The rate of active entrepreneurs is determined through an indifference condition: Entrepreneurs enter production up to a rate where the expected profit is equal to the outside option.

**Equilibrium of the Financial and the Real Market**

In the equilibrium of the production market, supply has to equal demand. The demand side is the fraction of households hit by the preference shock $1 - n$. The supply side consists of the aggregate output of successful entrepreneurs and of subsistence producers. The market clearing condition of the production market, assuming no search frictions, is

$$n(1 - 2\theta)q_s + n\theta\mu_h q_e = (1 - n)q_b. \tag{23}$$

Combining the FOCs of entrepreneurs and subsistence producers with the FOC of buyers yield the relationship between the equilibrium quantities produced and consumed and the interest rates.

$$\begin{aligned} \frac{u'(q_b)}{c'(q_s)} &= (1 + i_d), && (24) \\ \frac{u'(q_b)}{\tilde{c}'(q_e)} &= (1 + \bar{i})(1 + i_d), && (25) \\ \text{where} \quad \bar{i} &\equiv \mu_h i_s + (1 - \mu_h)i_f. \end{aligned}$$

Rate $\bar{i}$ is the expected interest rate a successful entrepreneur has to pay, given that the peer behaves. The interest rates on deposits and on loans drive a wedge between the marginal utility of consumption and the marginal cost of production. The higher the interest rates are, the further away is the economy from the first best allocation where the ratio of the marginal utility of consumption and the marginal cost of production are equal.

At the beginning of the section, we ruled out external sourcing possibilities for the MFI. More specifically, we assume that all loans have to be fully covered by deposits and that the MFI is not dependent on subsidies. Hence, aggregated deposits have to be greater or equal to demanded loans. In the optimal allocation, the two measures are equal.

$$n(1-\theta)d = n\theta l + (1-n)l_b. \tag{26}$$

**Marginal Value of Deposits**

To obtain the marginal value of deposits, we first take the derivative of Equation (5).

$$\frac{\partial V(d)}{\partial d} = (1-n)\frac{u'(q_b)}{p} + n(1-\theta)(1+i_d)\phi + n\theta\frac{\partial E\{W\}}{\partial d}. \tag{27}$$

If the household is a buyer, he receives marginal utility of $u'(q_b)/p$. If he is a producer, he receives principal and interest in subperiod B and can consume the general good. If, instead, he is an entrepreneur, Lemma 1 below reveals that the marginal value of deposits is the same as for the producer.

**Lemma 1** *For $l > 0$, the marginal value of holding deposits for an entrepreneur is equal to $(1+i_d)\phi$.*

**Proof.** To verify Lemma 1, note that if entrepreneurs are not credit-constrained, then the equilibrium wage of the employee is independent of $d$. This implies that, when the entrepreneur increases his deposits, he is able to decrease the loan by the same amount. This implies that the marginal value of deposits is equal to the negative value of the expected marginal value of loans of an entrepreneur.

The expected marginal value of loans depends on the outcome of production and the respective interest rate that the entrepreneur has to pay. There are three outcomes: In the first outcome, the entrepreneur defaults and pays zero. In the second outcome, both entrepreneurs are successful, and the gross interest rate is $1 + i_s$. Finally, the peer entrepreneur defaults, and the gross interest rate is $1 + i_f$. Thus, the negative value of the ex-ante expected marginal value of loans is: $\phi\mu_h(1 + \bar{i})$. Using Equation (19) gives: $\phi\mu_h(1 + \bar{i}) = \phi(1 + i_d)$. ■

In the next step, we substitute $p$ by using the equilibrium condition of entrepreneurs (Equation (17)). Finally, we replace the left-hand side of Equation (27) by using the lagged intertemporal optimality condition (Equation (9)).[14] This gives the equilibrium relationship between the marginal utility of consumption and the marginal cost of industrial production conditional on the growth rate of money and the interest rate.

$$\frac{\gamma - \beta}{\beta} = (1 - n)\left[\frac{\mu_h u'(q_b)}{(1 + \bar{i})\tilde{c}'(q_e)} - 1\right] + n i_d. \tag{28}$$

Using the equilibrium condition of industrial production (Equation (25)) shows that one plus the deposit rate is equal to $\gamma/\beta$.

$$\frac{\gamma}{\beta} = (1 + r)\gamma = 1 + i_d. \tag{29}$$

**Real Value of Deposits**

The real value of deposits ($\phi d$) can be derived by using the clearing condition of the MFI (Equation (26)). Recall from the optimization problem of entrepreneurs that $l = K/\phi + \tilde{c}(q_e)/\phi + \Pi_s - d$, and from the budget constraint of buyers that $l_b = pq_b - d$. Using, once again, the optimality condition of the entrepreneur to replace $p$ and the market clearing condition of the production market gives the real value of deposits, where the superscript $^*$ denotes equilibrium values:

$$\phi d = (1 - n)(1 + \bar{i})\tilde{c}'(q_e^*)q_b^* + n\theta K + n\theta\tilde{c}(q_e^*) + n\theta[(1 + \bar{i})\tilde{c}'(q_e^*)q_s^* - c(q_s^*)]. \tag{30}$$

[14] We use the long-term relation of the growth rate of money and the real value of money: $\phi = \phi_{-1}\gamma$.

Equation (30) shows that the real value of deposits depends on the expected interest rate that producing households have to pay in order to receive a loan. The real price of money ($\phi$) can be determined by substituting $d$ by the money stock $M_{-1}$. All these conditions have to be satisfied in the symmetric, stationary equilibrium.

**Definition 1 (Equilibrium)** *The symmetric, stationary equilibrium* $\{q_b, q_e, q_s, x, p, \phi, i_d\}$ *satisfies the equilibrium equations (20), (21), (23), (24), (25), (26), (29) and (30). The interest rate* $i_d$*, the real price of money* $\phi$*, and the price* $p$ *result from the monetary policy of the central bank, which decides over the parameter values* $\{M, \gamma\}$*.*

# 5 Discussion

In this section, we analyze the economic equilibrium of our model. We suppose that the microfinance institution receives no subsidies and is in our case fully independent of commercial banks or the financial market. It provides basic financial services to the households. In Section 5.1, we offer a numerical example to present the equilibrium outcome of our model and give the intuition to our results. In Section 5.2, we compute the welfare costs for different money growth rates. Moreover, we compute the welfare costs of having no access to financial services for given monetary policies.

## 5.1 Numerical Example

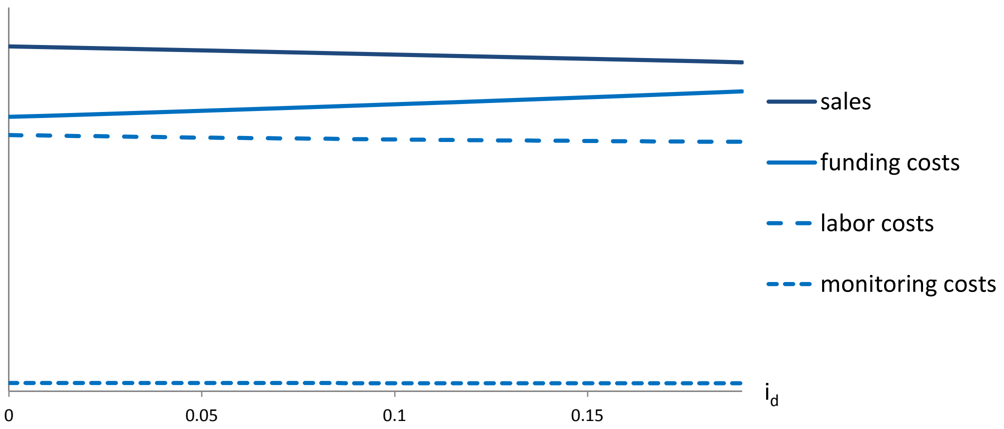
We suppose that a household is with probability 0.6 a buyer, with probability 0.3 a producer, and with probability 0.1 an entrepreneur with a business idea. For our numerical example, we use the same utility and disutility functions as in Lagos and Wright (2005). We assume that the disutility function of the employee is identical to the function of producers, but scaled down by the factor $a < 1$. Equation (31) displays the utility functions and the disutility functions of the households. The disutility function in market 2 is linear for all households.

$$u(q) = A\frac{(q+b)^{1-\eta} - b^{1-\eta}}{1-\eta}, \quad U(x) = D\log(x), \quad c(q) = q^{\rho}, \quad \tilde{c}(q) = aq^{\rho}, \quad (31)$$

where $D = 3$ and $\rho = 2.5$. For the cost function of entrepreneurs, we assume that $a = 0.25$ and the setup costs are $K = 0.2$. The success probability is $\mu_h = 0.9$ if the entrepreneur behaves, while it is $\mu_h = 0.3$ and he receives a private benefit of $B = 0.3$ if the entrepreneur shirks. Monitoring costs of entrepreneurs are $\delta_m l$, with $\delta_m = 0.05$. The remaining parameter values are $A = 1$, $b \approx 0$ and $\eta = 0.3$.[15]

Figure 2 displays sales and expenses of a representative entrepreneur as a function of the deposit rate $i_d$. The dark blue line indicates the sales of the enterprise. Sales are decreasing with the deposit rate. The other three lines display monitoring costs, labor costs and interest costs. The cost components are added up in the figure. Thus, the blue line marks the aggregated expenses of the enterprise.

Figure 2: Enterprise's sales and expenses




Increasing the inflation rate leads to a rise of the equilibrium deposit rate. Higher inflation rates, therefore, increase the funding costs of entrepreneurs. Funding costs are steadily increasing with the deposit rate. Monitoring costs are low and increase only marginally with the deposit rate. Labor costs account for the largest part of the expenses. Wages are slightly decreasing, since real output declines if the inflation rate increases. For deposit rates above 20 percent, the expected profits of entrepreneurs are smaller than their outside option–leaving deposits at the MFI and instead receiving interest in

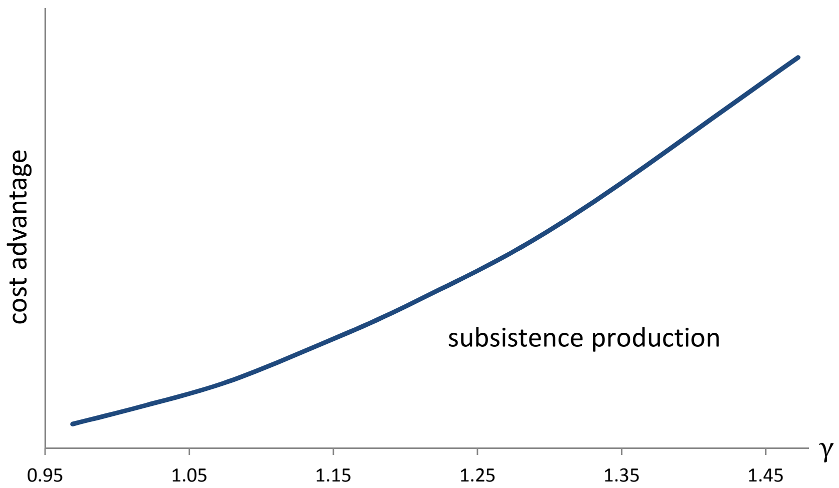
[15]Table 2 in the Appendix reports all underlying parameter values for the numerical example.

subperiod B. Thus, above this threshold rate, entrepreneurship collapses and entrepreneurs will be inactive.

The funding costs for entrepreneurs are relatively small if the rate of inflation is low. In such an environment, enterprises can resort to their superior technology. Enterprises' profits are greater than their outside option. However, if the inflation rate increases, subsistence production becomes more and more advantageous, for the reason that home production involves no external funding. Ultimately, very high inflation rates lead to a collapse of entrepreneurship.

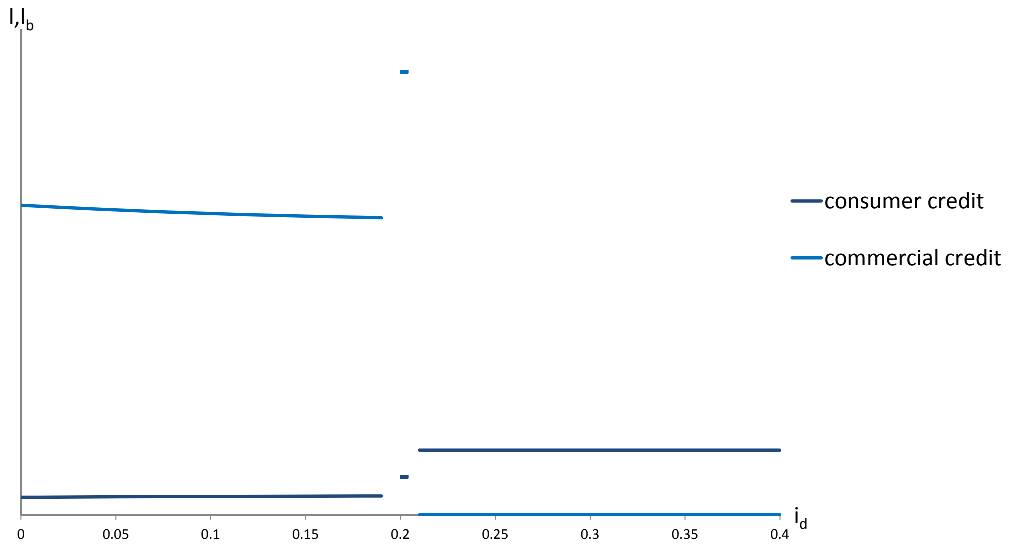
Figure 3 illustrates how the threshold value of entrepreneurs' production depends on the cost advantage and the inflation rate. In the area below the line, only subsistence production exists. For the case that the technology of enterprises is only slightly superior to the subsistence technology, the threshold value is very low. The more advanced the production technology is, the larger is the array where entrepreneurship is profitable.

Figure 3: The collapse of entrepreneurship




A further question of interest is how monetary policy affects the borrowing behavior of poor households. Figure 4 compares the credit volume of a consuming household to the credit volume of an entrepreneur as functions of the deposit rate. The commercial credit is nearly constant up to the threshold value, beyond which it collapses to zero. If the deposit rate is exactly 20 percent, only part of the entrepreneurs are active. At this rate entrepreneurs are indifferent between starting a business and their outside option. Contrary to commercial credits, the volume of the consumer credit increases regularly with the deposit rate. Above the threshold, the consumer credit jumps to a higher volume. If the deposit rate is 20 percent, the credit volume lies between the two values.

Figure 4: Consumer credit and commercial credit




We will show in Section 5.2 that the impact on social welfare is largest if households take out commercial and consumer loans simultaneously. The discovered link between monetary policy and the credit composition (commercial versus consumer credits) in our model is also important for another reason: The empirical finance literature finds a positive relationship between enterprise credits and economic growth, whereas the relationship between consumer credits and growth is insignificant (Beck et al., 2012).

## 5.2 Welfare Implications

Welfare of households depends on the amount of production and consumption. The equilibrium outcome is affected by the inflation rate, which in turn is induced by the central bank through the money growth rate. The aggregated steady state lifetime utility of households is

$$\mathcal{W} = \frac{1}{1-\beta}\{(1-n)u(q_b) - n\theta K - n\theta\tilde{c}(q_e) - n\theta\phi\delta_m l - n(1-2\theta)c(q_s) + U(x) - x\}. \tag{32}$$

The expected utility of one household at the beginning of the period consists of the expected net utility of the two subperiods. In subperiod A, it is the probability of being a consuming household times the utility of consumption minus the probability of being an entrepreneur or a producer times the disutility of production. In subperiod B, it is the utility of consumption minus the production of the general good, since all households consume the general good.[16]

To derive the welfare costs of inflation, we apply the standard approach used in search-theoretic monetary models (see e.g., Lagos and Wright (2005) or Craig and Rocheteau (2008)). We ask what fraction of consumption would households be willing to give up in order to change inflation from $\pi_0$ to $\pi_1$. We denote the fraction of consumption by $\Delta$. The aggregated steady state lifetime utility of a household that decreases its overall consumption by the fraction $\Delta$ is:

$$\begin{aligned}\mathcal{W}_\Delta = & \frac{1}{1-\beta}\{(1-n)u((1-\Delta)q_b) - n\theta(K + \tilde{c}(q_e) + \phi\delta_m l) \\ & -n(1-2\theta)c(q_s) + U((1-\Delta)x) - x\}.\end{aligned} \tag{33}$$

To obtain the welfare costs of increasing inflation from $\pi_0$ to $\pi_1$, we compute the $\Delta$, which sets $\mathcal{W}_\Delta(\pi_0) = \mathcal{W}(\pi_1)$.

Table 1 compares the real price, consumption and production quantities in equilibrium for inflation rates of -5%, 10%, 14%, 20% and 30%. Entrepreneurship collapses at an inflation rate of 14 percent–which is the inflation threshold in our numerical example. For the two cases with higher inflation rates, the output of enterprises is zero. In contrast, the output of subsistence producers decreases only slightly with higher inflation rates.

[16] The formula for aggregated steady state welfare is similar to the optimization problem of the social planner. However, in contrast to Section 2.1, the quantities produced and consumed are now chosen by the households.

Table 1: Welfare costs of inflation

| | $\gamma = \beta$ | $\gamma = 1.00$ | $\gamma = 1.14$ | $\gamma = 1.20$ | $\gamma = 1.30$ |
|---|---|---|---|---|---|
| $\phi$ p | 1.2803 | 1.2320 | 1.1548 | 1.1403 | 1.0668 |
| $q_b$ | 0.4389 | 0.4204 | 0.3371 | 0.2963 | 0.2834 |
| $q_s$ | 0.6400 | 0.6239 | 0.5975 | 0.5926 | 0.5668 |
| $q_e$ | 1.5035 | 1.4163 | 1.2429 | 0 | 0 |
| active | 1 | 1 | 0.44 | 0 | 0 |
| Welfare costs: | | | | | |
| MFI | – | 0% | 0.66% | 1.12% | 1.20% |
| no MFI | 1.01% | 1.02% | 1.17% | 1.25% | 1.41% |

Notes: *active* indicates the rate of active entrepreneurs; *no MFI* indicates that households have no access to financial services. In the benchmark case all households have access to the MFI, and the central bank implements the Friedman rule. Entrepreneurship collapses at an inflation rate of 14 percent.

In the lower part of the table, we display the welfare costs of inflation.[17] The first row depicts the costs of inflation, given that households have access to the microfinance institution, in comparison to the Friedman rule (our benchmark). For low rates, the welfare costs of inflation are small. However, welfare costs are much higher if inflation rates are high. The reason is that if inflation lies above the threshold value, entrepreneurship collapses. Hence, at moderate inflation rates welfare costs are negligible, but, above a certain threshold, welfare costs increase substantially.

Another important result of our model is that monetary policy affects the impact of microfinance. The last row of the table depicts the costs of inflation given that households have no access to the microfinance institution.[18] The comparison of welfare costs with and without a microfinance institution discloses that establishing microfinance institutions generally increases social welfare. However, the magnitude of the positive impact depends crucially on

[17] Lagos and Wright (2005) show that the costs of inflation can be significantly larger when trading frictions are taken into account or other pricing mechanisms are applied.

[18] Welfare costs without access to financial services are again measured relative to the benchmark case, where every household has access to financial services. To calculate social welfare without access to the MFI, we closely follow Berentsen et al. (2007): We first set $i_d = 0$ in Equation (28) to derive the equilibrium amounts of the production good and afterwards calculate social welfare.

the inflation rate. The welfare impact of microfinance is largest for moderate inflation rates, where external funding enables entrepreneurship and consumption smoothing. For inflation rates above the threshold, the welfare impact of microfinance is smaller, since households use loans solely for consumption smoothing.

# 6 Conclusion

In this paper, we analyzed the effects of microfinance and inflation on social welfare in developing countries. In our theoretical model, group lending mitigates the agency problem between entrepreneurs and the microfinance institution, which in turn allows entrepreneurship to emerge. At the same time, access to financial services affects households' decision to hold money balances, to save and to borrow. Under moderate inflation rates, entrepreneurship emerges, which in turn increases aggregate production and reduces real prices. Households receive interest payments on their deposits and benefit from the possibility to take out loans. Yet, subsistence producers, without profitable business ideas, are negatively affected by the general equilibrium effect on prices.

Our main result is that entrepreneurship collapses above an inflation threshold. A higher rate of inflation negatively affects entrepreneurship through two mechanisms. First, the standard real balance effect of inflation lowers output and expected real profits of entrepreneurs. Second, inflation increases the funding costs of entrepreneurs. The real balance effect affects subsistence producers and entrepreneurs alike, whereas the effect on funding costs applies only to entrepreneurs. There exists an inflation threshold beyond which entrepreneurship collapses, because entrepreneurs who rely on external funding are more affected by inflation than subsistence producers.

Our welfare analysis shows that the magnitude of the impact of microfinance on households' welfare crucially depends on the prevailing monetary policy regime. Microfinance has the largest impact on social welfare if inflation is moderate and households use loans to start small enterprises as well as for consumption smoothing. If inflation is high, the impact of microfinance on welfare decreases substantially, as entrepreneurship collapses and households use loans solely for consumption smoothing. Our findings imply that a reasonable monetary policy is especially important in developing countries, where households face high transaction costs and information problems when applying for small loans. Better knowledge of the relationship between monetary policy and economic development and growth is of great interest for central banks, the World Bank and development agencies, but primarily to improve the quality of life of the people in developing countries.

# Bibliography


Armendáriz, B. (1999). On the design of a credit agreement with peer monitoring. *Journal of Development Economics 60*(1), 79–104.

Armendáriz, B. and J. Morduch (2000). Microfinance beyond group lending. *The Economics of Transition 8*(2), 401–420.

Armendáriz, B. and J. Morduch (2010). *The Economics of Microfinance, Second Edition*, Volume 1 of *MIT Press Books*. The MIT Press.

Banerjee, A. V. and E. Duflo (2010). Giving credit where it is due. *Journal of Economic Perspectives 24*(3), 61–80.

Beck, T., B. Büyükkarabacak, F. Rioja, and N. Valev (2012). Who gets the credit? And does it matter? Household vs. firm lending across countries. *The B.E. Journal of Macroeconomics 12*(1).

Berentsen, A., G. Camera, and C. Waller (2007). Money, credit and banking. *Journal of Economic Theory 135*(1), 171–195.

Craig, B. and G. Rocheteau (2008). Inflation and welfare: A search approach. *Journal of Money, Credit and Banking 40*(1), 89–119.

Demirguc-Kunt, A. and L. Klapper (2012). Measuring financial inclusion: The Global Findex database. Policy Research Working Paper Series 6025, The World Bank.

Easterly, W. and S. Fischer (2001). Inflation and the poor. *Journal of Money, Credit and Banking 33*(2), 160–178.

Ghatak, M. and T. W. Guinnane (1999). The economics of lending with joint liability: Theory and practice. *Journal of Development Economics 60*(1), 195–228.

Holmstrom, B. and J. Tirole (1997). Financial intermediation, loanable funds, and the real sector. *The Quarterly Journal of Economics 112*(3), 663–691.

Kaboski, J. P. and R. M. Townsend (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica 79*(5), 1357–1406.

Kaboski, J. P. and R. M. Townsend (2012). The impact of credit on village economies. *American Economic Journal: Applied Economics 4*(2), 98–133.

Laffont, J.-J. and P. Rey (2003). Moral hazard, collusion and group lending. IDEI Working Papers 122, Institut d'Ëconomie Industrielle (IDEI), Toulouse.

Lagos, R. and R. Wright (2005). A unified framework for monetary theory and policy analysis. *Journal of Political Economy 113*(3), 463–484.

Levine, R. (2005). Finance and growth: Theory and evidence. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1 of *Handbook of Economic Growth*, Chapter 12, pp. 865–934. Elsevier.

Matin, I., D. Hulme, and S. Rutherford (2002). Finance for the poor: From microcredit to microfinancial services. *Journal of International Development 14*(2), 273–294.

Robinson, M. (2001). *The Microfinance Revolution: Sustainable Finance for the Poor*. Other World Bank Bks. World Bank.

Rutherford, S. (2006). GRAMEEN II: The first five years 2001–2005. Technical report, MicroSave. Dhaka, Bangladesh.

Rutherford, S., M. D. Maniruzzaman, and S. K. Sinha (2004). GRAMEEN II At the end of 2003: A 'grounded view' of how Grameen's new initiative is progressing in the villages. Technical report, MicroSave. Dhaka, Bangladesh.

Stiglitz, J. E. (1990). Peer monitoring and credit markets. *World Bank Economic Review 4*(3), 351–366.

World Bank (2012). New database shows three quarters of world's poor are "unbanked". Press Release No: 2012/400/Dec, Washington.

# A Appendix

## A.1 Optimization Problem of Buyers

The optimization problem of the buyer with access to the MFI is:

$$\max_{q_b, l_b} \quad \{u(q_b) + W(0, l_b, d + l_b - pq_b)\}, \tag{34}$$

$$\text{s.t.} \quad pq_b \leq d + l_b, \tag{BC}$$

$$l_b \leq \bar{l}. \tag{LC}$$

The Lagrangian of the maximization problem is:

$$\mathrm{L}(q_b, l_b) = u(q_b) + W(0, l_b, d + l_b - pq_b) - \lambda\phi[pq_b - d - l_b] - \lambda_l\phi[l_b - \bar{l}]. \tag{35}$$

The first-order conditions are:

$$\begin{array}{lccl} q_b : & u'(q_b) & = & pW_m + \lambda_l \phi p, \\ l_b : & (\lambda - \lambda_l)\phi & = & -W_l - W_m, \\ \lambda : & l_b & = & pq_b - d, \\ \lambda_l : & 0 & \geq & l_b - \bar{l}. \end{array} \tag{36}$$

Using the envelope conditions of holding money and borrowing money from Section 3.1 and substituting these into the second FOC gives $\lambda - \lambda_l = i_d$. The difference between the multiplier of the budget constraint and the multiplier of the loan constraint is equal to the marginal change in utility. If the lending constraint (LC) is binding, then $\lambda_l > 0$ and the buyer borrows $l_b = \bar{l}$. For the case that the lending constraint is not binding, $\lambda_l = 0$. In our model, we assume that the borrowing constraint will never be binding. In this case, the consumer chooses $l_b$ such that in equilibrium the multiplier $\lambda$ is equal to $i_d$. Using this result and again the envelope condition for the first FOC gives the following condition which has to hold if buyers maximize utility:

$$u'(q_b) = \phi p(1 + i_d). \tag{37}$$

## A.2 Optimization Problem of Subsistence Producers

Subsistence producers choose the amount $q_s$ that maximizes profit, thereby taking as given the price $p$. Producers incur utility costs $c(q_s)$ for producing the amount $q_s$. Subsistence producers' optimization problem is:

$$\max_{q_s} \quad \{-c(q_s) + W(d, 0, pq_s)\}. \tag{38}$$

The first-order condition is:

$$q_s : \quad c'(q_b) \quad = \quad pW_m = \phi p. \tag{39}$$

## A.3 Optimization Problem of Unbanked Households

The unbanked households can neither hold deposits nor take out loans. Buyers choose how much to demand of good $q$. Their optimization problem is formulated as follows:

$$\max_{q_b} \quad \{u(q_b) + W(0, 0, m - pq_b)\}, \tag{40}$$

$$\text{s.t.} \quad pq_b \leq m, \tag{BC}$$

where the constraint states that households can dispense up to their money holdings $m$. The Lagrangian of the maximization problem of the buyer is

$$\mathrm{L}(q_b) = u(q_b) + W(0, 0, m - pq_b) - \lambda_m \phi [pq_b - m]. \tag{41}$$

The first-order conditions are:

$$\begin{aligned} q_b : \quad u'(q_b) &= W_m p + \lambda_m \phi p, \\ \lambda_m : \quad m &= pq_b. \end{aligned} \tag{42}$$

The envelope condition for holding money states that $W_m = \phi$. In the optimum, buyers choose $q_b$ such that the following equation is satisfied:

$$u'(q_b) = \phi p (1 + \lambda_m). \tag{43}$$

If the budget restriction of the buyer is binding, then he will consume: $q_b = m/p$. If the buyer's money balance is sufficiently large, he will consume the amount of the production good that sets marginal utility equal to the real price. In this case, $\lambda_m$ is equal to zero.

Subsistence producers choose the amount $q_s$ that maximizes profit, thereby taking as given the price $p$. Households incur utility costs $c(q_s)$ for producing the amount $q_s$. Their optimization problem is formulated as follows:

$$\max_{q_s} \quad \{-c(q_s) + W(0, 0, pq_s)\}. \tag{44}$$

The Lagrangian of the maximization problem of the subsistence producer is

$$\mathrm{L}(q_s) = -c(q_s) + W(0, 0, m + pq_s). \tag{45}$$

The first-order condition is:

$$c'(q_s) = W_m p = \phi p. \tag{46}$$

## A.4 Parameter Values

Table 2: Parameter values of the numerical example

| Parameter | Values | Description |
|---|---|---|
| A | 1 | parameter of the utility function |
| $a$ | 0.25 | entrepreneurs' cost factor |
| B | 0.3 | entrepreneurs' private benefit |
| b | 0 | parameter of the utility function |
| $\beta$ | 0.95 | discount factor |
| D | 3 | parameter of the utility function |
| $\delta_m$ | 0.05 | monitoring costs |
| $\eta$ | 0.3 | parameter of the utility function |
| K | 0.2 | setup costs of entrepreneurs |
| $\mu_h$ | 0.9 | success rate, high |
| $\mu_l$ | 0.3 | success rate, low |
| n | 0.4 | rate of producers and entrepreneurs |
| $\theta$ | 0.25 | ratio of entrepreneurs to producers |
| $\rho$ | 2.5 | parameter of the production function |

Notes: Standard parameter values are similar to Lagos and Wright (2005).

# Chapter II

## The Doping Threshold in Sport Contests

**Abstract**

We analyze the doping behavior of heterogeneous athletes in an environment of private information. In a n-player strategic game, modeled as an all-pay auction, each athlete has private information about his actual physical ability and choses the amount of performance-enhancing drugs. The use of doping substances is costly but not further regulated. The main finding of the analysis is the existence of a doping threshold. In our leading case only strong athletes dope. The level of the doping threshold is increasing in the doping costs and decreasing in the prize level. Furthermore, increasing the number of athletes affects the doping decision in two ways. More competition increases the incentives to dope for strong athletes. At the same time, we find a discouragement effect for weak athletes.

| | |
|---|---|
| Keywords: | Auctions, Contests, Doping, Heterogeneity, Private Information. |
| JEL Classification: | C72, D44, D82. |

# 1 Introduction

Since the 1990s, the number of athletes tested positive for doping has substantially increased. Positive doping cases have been covered by the media particularly in professional cycling. The Festina affair in 1998 and the Fuentes scandal in 2006, followed by extensive legal investigations, show that many favorites and even entire cycling teams doped systematically.[1] The recently published material from the investigations of the U.S. Anti-Doping Agency against Lance Armstrong and the U.S. Postal Service Team reveal the actual dimension of doping in professional cycling (USADA, 2012). Doping is, however, by no means a new phenomenon. Written sources show that already in ancient Greece athletes used stimulants and dubious mixtures to enhance their strength and endurance (Verroken, 2005).

A closer look at doping practices reveals that athletes either dope during the training period, directly before a competition, or do both. Performance-enhancing drugs are popular because of their immediate and strong impact on individual performance. Drugs instantaneously improve performance, whereas training is time consuming and affects the performance only in the long run. Moreover, many drugs are only detectable for a short period of time. These features make doping especially attractive for athletes who wish to further enhance their performance shortly before a contest. A better understanding of athletes' incentives to take drugs is an important prerequisite to increase the efficiency of anti-doping policies.

Anti-doping agencies and the International Olympic Committee (IOC) have the objective to establish a doping-free environment for sports contests. The regulator, however, cannot directly control whether athletes take performance-enhancing drugs. Asymmetric information makes detection difficult and expensive. The current policies of anti-doping agencies are out-of-competition as well as in-competition doping controls and the sanctioning of convicted athletes. The ongoing doping cases in sports such as athletics, professional cycling or weight lifting show that despite severe sanctions and public humiliation in the case of detection, doping remains present in professional sports. Some

[1] Dilger et al. (2007) give a short review of the history of doping and present recent doping scandals in professional cycling and in athletics.

experts even believe that doping has increased in particular sports due to the ongoing commercialization and the development of more effective drugs.

Asymmetric information can occur at several levels. The existing doping literature has primarily focused on asymmetric information between athletes and the regulator, leaving aside considerations about asymmetric information between athletes. Muehlheusser (2006) emphasizes the relevance of informational asymmetries in sports and recommends taking them into account when designing contests. Each athlete has private information about his actual physical ability, but can only guess the abilities of his rivals. The behavior of athletes in camouflaging and even misrepresenting their actual ability indicates that this informational advantage is important.[2] An athlete's decision to use performance-enhancing drugs will hence not only depend on regulations, but also on private information concerning his ability. So far, this issue has been neglected by the doping literature. The aim of this paper is to investigate the rationale of doping in a heterogeneous n-player game under private information.

Many researchers have studied the doping problem in the context of actual or potential anti-doping regulations. Berentsen (2002) is one of the first to analyze anti-doping regulations in a strategic two-player game.[3] Cheating and doping have recently been introduced into the theory of contests and tournaments. Their common basis is the Lazear-Rosen tournament model, extended to include a regulator who audits the athletes. Kräkel (2007), for example, analyzes the doping behavior of heterogeneous athletes who optimize the use of doping and legal inputs.[4] Another strand of the doping literature has focused on fair play norms and on peer group approval based on past doping decisions (Eber, 2008, 2011; Strulik, 2012). In both approaches an equilibrium without doping is possible; however, a reliable coordination mechanism is needed to

[2] For instance, athletes play down their actual form or, conversely, conceal injuries and illnesses at press conferences.

[3] For similar contributions, see e.g., Eber and Thépot (1999), Maennig (2002), Haugen (2004). Berentsen and Lengwiler (2004) analyze doping and fraudulent accounting in an evolutionary game.

[4] Analogous is Stowe and Gilpatric (2010) who focus instead on the doping decision and different regulation regimes. Curry and Mongrain (2009) investigate the effects of the prize structure. Finally, Gilpatric (2011) analyzes how enforcement affects the effort levels.

make this equilibrium stable.[5]

Our paper is closely related to the literature that analyzes cheating in tournament models. In contrast to the random component in Lazear-Rosen tournament models, we assume that athletes' abilities are heterogeneously distributed and private information. As the interest lies in identifying how heterogeneity affects doping behavior under private information, we assume that taking performance-enhancing drugs is costly and ignore further anti-doping regulations for the moment. The information structure of the contest is the following: In stage one, *nature* independently draws athletes' abilities from a distribution. The number of athletes and the distribution is common knowledge. The actual ability of the athlete, however, is private information. In stage two, athletes may improve their performance by taking performance-enhancing drugs. They base their doping decision on their actual ability and their beliefs about the abilities of their competitors. Finally, in stage three, the athlete with the greatest performance–the combination of ability and the chosen amount of doping–wins the prize money.

Our private information setting with heterogeneous athletes yields new insights which complement the results of existing doping literature. We analyze how the prize amount, the doping costs, the number of athletes and the distribution of abilities affect athletes' doping behavior. We show that under private information not all athletes take performance-enhancing drugs. For the majority of underlying parameter values, there exists a doping threshold. In our leading case, strong athletes dope, and athletes beneath the doping threshold have no incentive to dope. The doping behavior of an athlete depends crucially on his actual ability and the degree of competition. If the degree of competition increases, strong athletes take larger amounts of drugs. At the same time there exists a discouragement effect for athletes with low abilities. The anticipation of encountering stronger athletes in the contest discourages weak athletes from doping. For nonstandard parameterizations, three other equilibrium outcomes occur.

The paper is structured as follows: In Section 2, we describe the model

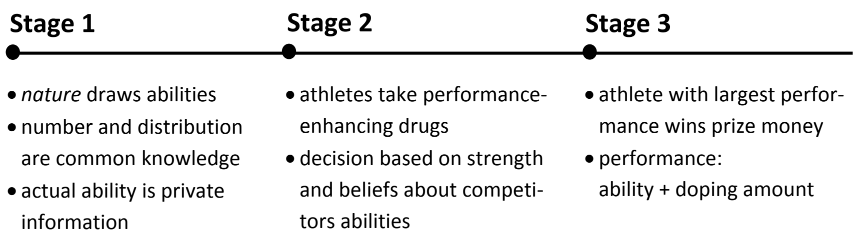
[5] Bird and Wagner (1997) were the first who proposed decentralized mechanisms based on social norms to solve the doping problem. An example of such a mechanism is whistleblowing (see e.g., Berentsen et al., 2008).

and present the main results. In Section 3, we formally derive the equilibrium outcome with the doping threshold. In Section 4, we discuss the results of the doping model. First, we show how the doping threshold depends on the number of athletes and on the ratio of prize money to doping costs. Second, we analyze the influence of the distribution of abilities. Finally, we display the equilibrium results under special parameterizations. Section 5 concludes.

## 2 The Model and Main Results

In a contest, $n$ athletes compete against each other. The winner receives prize money $v$. All athletes are risk-neutral and maximize their expected payoff. Figure 1 displays the information structure of the contest. In stage one, *nature* independently draws athletes' abilities $a_i$, for $i = 1, ..., n$. The abilities are drawn from a given cumulative distribution function $F$ with support $[0, 1]$. More specifically, we assume a power function distribution $F(a) = a^{\alpha}$. This flexible functional form allows us to analyze the doping incentives for different shapes of the distribution. At the same time it ensures an explicit solution. The number of athletes and the distribution are common knowledge. The actual ability of the athlete, however, is private information.

Figure 1: The information structure of the contest




In stage two, athletes may improve their performance by taking performance-enhancing drugs. The athlete's performance $p_i$ is a linear combination of his ability $a_i$ and the amount of doping $d_i$ he chooses to take: $p_i = a_i + d_i$. We assume that athletes are free to choose an arbitrary doping amount and are not limited to the two discrete options of *doping* or *not doping*.[6] Athletes base

[6] The model can readily be adapted to apply to a binary decision between *doping* and *not doping*. Note that this modification would not change the qualitative results.

their doping decision on their actual ability and their beliefs about the abilities of their competitors. Doping substances are not free of charge and the athlete has to pay for his doping substances before the contest begins. Since we are interested in how heterogeneity affects doping behavior under private information, we assume that taking performance-enhancing drugs is costly and ignore further anti-doping regulations for the moment.[7]

Finally, in stage three, the athlete with the greatest performance–the combination of ability and the chosen amount of doping–wins the prize money. For simplicity, the athlete with the greatest performance wins with certainty.[8] In a world without doping, the athlete with the greatest ability would win the contest. However, as athletes can choose arbitrary amounts of doping, it is, in principal, possible for a weaker athlete to beat a more talented athlete. Each athlete thus faces a trade-off between gaining a higher likelihood of winning through doping and the increased costs.

The doping amount is the difference between the athlete's performance $p_i$ and his ability $a_i$. If an athlete's performance is equal to his ability, then he does not dope and his doping costs are zero. Every athlete's doping cost function is thus a function of the difference $p_i - a_i$. The doping costs are $c(p - a_i)$, where the parameter $c$ indicates the magnitude of marginal doping costs. The linearity of the function ensures a closed-form solution. The cost function is the same for all athletes. We denote the ratio of prize money to marginal doping costs by $w$ ($w \equiv v/c$). From now on, we will assume that the parameter values of $\alpha$, $n$ and $w$ satisfy $\alpha(n-1) \geq 1$ and $w\alpha(n-1) > 1$ and use the term *leading case* for these parameterizations. In Section 4.3, we will relax these assumptions and address special cases.[9]

The athlete's probability of winning depends on his chosen performance $p$ and on the performances of the other athletes. The athlete only wins the contest if his performance $p(a_i)$ is greater than the performances of all other athletes. On the other hand, the costs accrue even if the athlete does not win

[7] This is comparable to a situation with toothless anti-doping regulations. For example, if athletes can easily manipulate their test results.

[8] See e.g. Kovenock et al. (1996) for all-pay auctions with complete information. For all-pay auctions with private information, see among others Amann and Leininger (1996) and Feess et al. (2008).

[9] See Appendix A.2 for an overview of all possible cases and their underlying conditions.

the contest. Equation (1) displays the payoffs of an athlete with ability $a_i$ who chooses the performance $p(a_i)$.

$$\Pi(a_i) = \begin{cases} v - c(p(a_i) - a_i) & \text{if} \quad p(a_i) > \max_{j \neq i} p(a_j), \\ -c(p(a_i) - a_i) & \text{if} \quad p(a_i) < \max_{j \neq i} p(a_j), \end{cases} \tag{1}$$

$$\text{nonnegativity constraint:} \qquad p(a_i) \geq a_i. \tag{2}$$

Negative doping amounts are ruled out by assumption ($d_i \geq 0$). Therefore, the nonnegativity constraint $p(a_i) \geq a_i$ has to hold for every possible $a_i$. Furthermore, in case of a tie between $m$ athletes ($m \leq n$), the prize money is split up equally between the $m$ athletes.

We are interested in athletes' optimal doping behavior. More precisely, we present the symmetric equilibrium performance function of the athletes. In equilibrium there exists a doping threshold. Athletes with ability below the threshold do not dope, whereas athletes with ability above the threshold do dope. Theorem 1 presents the equilibrium outcome of the doping model.

**Theorem 1** *There exists a symmetric Nash equilibrium in pure-strategies with the doping threshold $a^\star$, for parameterizations that satisfy the conditions of the leading case. In the symmetric equilibrium, an athlete with ability $a$ chooses the performance*

$$p(a) = \begin{cases} a & \text{if } a < a^\star, \\ a^\star + w\left[a^{\alpha(n-1)} - {a^\star}^{\alpha(n-1)}\right] & \text{if } a \geq a^\star. \end{cases} \tag{3}$$

*The unique doping threshold $a^\star$ is*

$$a^\star = [\alpha w(n-1)]^{\frac{1}{1-\alpha(n-1)}}. \tag{4}$$

**Proof.** See Appendix A.1.[10] ■

[10] The proof relies on the intermediate results of Section 3. Therefore, we recommend that readers cover Section 3 before turning to the proof.

# 3 The Doping Equilibrium

The actual contest is similar to an all-pay auction where every athlete has to pay for his personal doping amounts. We are interested in an equilibrium performance function which is increasing in $a$. If this is the case, then the probability of winning $G(a)$ is the c.d.f. of the highest order statistic $A_{(n-1:n-1)}$ of the remaining athletes.[11] Given the assumed power function distribution, the probability of having the greatest ability is $G(a) = \Pr\{A_{(n-1:n-1)} \leq a\} = F(a)^{n-1} = a^{\alpha(n-1)}$.

In order to find the symmetric Nash equilibrium in the doping contest, we apply the usual approach used in auction theory. First, we assume that in a symmetric Nash equilibrium an athlete with ability $a$ chooses the performance $p(a)$ and then formulate the expected return for this athlete (using the highest order statistic). Every athlete can deviate from his equilibrium strategy by choosing another performance. However, it does not make sense to choose a performance lower than $p(0)$ or higher than $p(1)$. In the first instance, one would never win, and, in the second instance, one would always win, but have to pay too much. For this reason, deviations from the equilibrium performance function can be modeled as follows: An athlete with ability $a$ who pretends to have a different ability $x$ chooses the associated performance $p(x)$ in the contest through adjusting the doping amount.

The expected utility function of an athlete with ability $a$, who pretends to have ability $x$, is the product of the prize $v$ multiplied by the winning probability of an athlete with ability $x$ minus the cost of doping necessary in order to achieve the performance $p(x)$. His expected utility is thus: $u(a, x) = vG(x) - c(p(x) - a)$. The athlete will choose the $x$ that maximizes his expected utility. He can only imitate performances that are equal to or greater than his ability, since the amount of doping cannot be negative. In equilibrium, the nonnegativity constraint $p(a) \geq a$ has to hold for every possible $a$.

The main contribution of this paper is the equilibrium performance function where the nonnegativity constraint is binding. However, in a fist step, the doping equilibrium is derived for cases where the constraint is not binding.[12]

[11] For further information on order statistics, see David and Nagaraja (2003).

[12] The nonnoegativity constraint is not binding if $\alpha < (n-1)^{-1}$ and $w > 1$.

This is done in order to introduce the basic solution technique and to emphasize the importance of the nonnegativity constraint. If the nonnegativity constraint is not binding, then, the equilibrium performance function can be derived in the following way.

To obtain the optimum, we differentiate the utility function with respect to $x$ and set it equal to zero. We obtain the FOC: $vG'(x) - cp'(x) = 0$. The optimal behavior of an athlete with ability $a_i$ is to imitate the strategy of an athlete with ability $x$ so that the FOC is satisfied.

In addition to the FOC, the incentive compatibility (IC) constraint has to be satisfied for every possible $a$ in the Nash equilibrium. There must be no gain in deviating from the equilibrium strategy. The IC constraint is satisfied if $u(a, a) \geqslant u(a, x)$ for all $a, x$. We assume that if an athlete is indifferent between $u(a, a)$ and $u(a, x)$, he will choose the equilibrium strategy $a$. In a symmetric equilibrium, the optimal $x$ corresponds to the athlete's own ability $a$. Therefore, we insert $x = a$ and $p(x) = p(a)$ into the FOC and obtain $vG'(a) - cp'(a) = 0$. The FOC states that the expected marginal return has to be equal to the marginal costs of increasing the winning probability. Dividing by $c$ and solving for $p'(a)$ gives $p'(a) = vG'(a)/c$. We see that only the ratio of prize money to marginal doping costs matters. Therefore, we use $w$ and obtain the basic equation to derive the equilibrium performance function:

$$p'(a) = wG'(a). \tag{5}$$

Our assumption of a linear doping cost function implies that an athlete with ability $a_i$ is indifferent between his equilibrium performance $p(a_i)$ and all other $p(a)$'s for which the nonnegativity constraint of doping is not binding.[13]

[13] This particular circumstance has to be kept in mind when we investigate the equilibrium and for the proof of Theorem 1. If, instead, we used a doping cost function with slightly decreasing marginal costs or a quadratic doping cost function, then the implied single crossing property would guarantee a strictly separating equilibrium. For example, assume that the doping cost function is: $c(a) = c\exp(-\rho a)$, where $\rho$ is very small. As $\rho$ goes to zero the marginal costs go to $c$ ($\lim_{\rho\to 0} c(a) = c$). Therefore, it seems plausible that the equilibrium outcome in the limit is similar to the outcome of a linear doping cost function.

Integrating from 0 to $a$ gives the performance function $p(a)$ for the cases where the nonnegativity constraint is not binding.

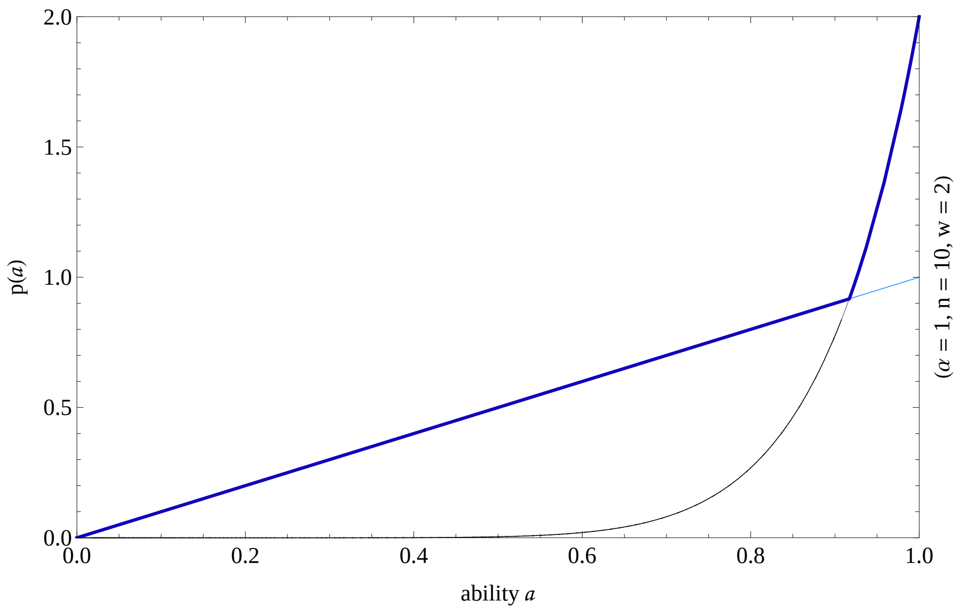
$$p(a) = wa^{\alpha(n-1)} \qquad \text{for} \quad 0 \leq a \leq 1. \tag{6}$$

The equilibrium performance function is similar to an equilibrium bid function of a standard all-pay auction.

Having specified the performance function for parameterizations where the nonnegativity constraint is not binding, we now turn to parameterizations of the leading case. Under these assumptions, the nonnegativity constraint is only binding for weak abilities. If we use performance function (6) for cases where the constraint is binding, then this would result in a performance function that lies partly below the underlying ability.

Figure 2 depicts such a situation ($\alpha = 1, n = 10, w = 2$). For athletes with an ability lower than the intersection (0.92), the performance function lies below the underlying ability. Hence, the nonnegativity constraint is violated. One could argue that augmenting the performance function of the lower section of the curve to the 45-degree line (to their corresponding ability) would solve the problem. However, we will show below, that this is not a Nash equilibrium.

Figure 2: Optimization that neglects the nonnegativity constraint



In the following, we will derive the equilibrium results of Theorem 1. To ensure that an athlete's performance is equal to or greater than his ability, we introduce the reserve ability approach.[14] This approach to find the equilibrium subject to the nonnegativity constraint is non-standard in auction literature. Therefore, we derive the equilibrium with the doping threshold step by step and, if necessary, provide further explanations.

The **reserve ability approach** is a three-step procedure to find the equilibrium outcome. First, assign an arbitrary reserve ability and assume that athletes with abilities below the reserve ability choose their ability as performance. Second, derive the equilibrium doping behavior of athletes with abilities above the reserve ability. And finally, determine the proper reserve ability such that athletes with abilities below it behave optimally. The approach can only be applied if the constraint is binding for a closed interval that includes the lowest ability and the c.d.f. is continuous.

The introduction of the reserve ability approach ensures that the performance of weak athletes is equal to their ability, such that $p(a)$ is no longer smaller than $a$. In contrast to an auction with a reserve price, weaker athletes can still win, since their performance is the sum of the chosen doping quantity and their ability. The reserve ability is non-effective under parameterizations where the constraint is not binding and performance outstrips ability over the whole support.

In order to obtain the performance function $p(a)$ of an athlete with ability $a$, we integrate Equation (5) from an arbitrary reserve ability $a^r$ with respect to $a$.

$$\int_{a^r}^{a} p'(z)dz = w \int_{a^r}^{a} G'(z)dz = w[F(a)^{n-1} - F(a^r)^{n-1}].$$

We solve the integral on the left-hand side of the equation above and take the obtained term $p(a^r)$ over to the right-hand side. Since the performance function at position $a^r$ must be equal to $a^r$, we can replace $p(a^r)$ by $a^r$. Finally, we substitute the power distribution function. We thus obtain the performance function of athletes, given the underlying distribution of abilities and

[14] The term *reserve ability* comes from the concept of the *reserve price*. The reserve price is used in auction theory in order to close out bid valuations that are too low. The bids have to at least meet the reserve price, which prevents bidders with valuations lower than $r$ from placing a bid. For a good overview on auctions and the reserve price see Krishna (2002).

the reserve ability $a^r$.

$$p(a, a^r) = a^r + w\left[F(a)^{n-1} - F(a^r)^{n-1}\right] = a^r + w[a^{\alpha(n-1)} - a^{r^{\alpha(n-1)}}] \text{ for } a^r \leq a \leq 1. \tag{7}$$

The proper reserve ability satisfies two conditions. First, the reserve ability $a^r$ has to be chosen such that the performance function $p(a, a^r)$ of athletes with ability above the reserve ability does not sink below the 45-degree line ($p(a, a^r) \geqslant a$ for $a \in [a^r, 1]$). Second, $a^r$ must not be chosen too large such that athletes with abilities below the reserve ability have no incentives to deviate. For our continuous power distribution function, these two conditions are satisfied if $p(a, a^r)$ has the same slope as the 45-degree line at position $a = a^r$. Hence, the slope of the performance function is equal to 1 at the proper reserve ability, which we denote the doping threshold. The reason for this smoothness condition can be seen in the athletes' optimization problem. In equilibrium, every athlete chooses his performance such that weaker athletes have no interest in imitating that performance. Therefore, the equilibrium performance function does not have a kink at the doping threshold.[15]

To determine the proper reserve ability $a^r$, we set the derivative of the performance function (7) equal to one. By solving the equation for $a$, we obtain Equation (4) which is the doping threshold. The doping threshold $a^\star$ defines the ability level where the athlete is indifferent between doping and not doping. Athletes with an ability beneath the threshold do not dope, and athletes with an ability above this value take performance-enhancing drugs. The doping threshold depends on the distribution of abilities, the number of competing athletes, and the ratio of prize money to marginal doping costs. In Proposition 2, we show that under the parameterization of the leading case the doping threshold exists.

**Proposition 2** *If* $\alpha(n-1) > 1$ *and* $\alpha(n-1)w > 1$*, then there exists a unique doping threshold* $a^\star \in (0, 1)$*, and strong athletes dope, and weak athletes abstain from doping.*

**Proof.** The derivative of Equation (7) is continuous and strictly increasing if $\alpha(n-1) > 1$. Given that $a^r = 0$, the derivative $p_a(0, 0)$ is zero. If the

[15] For the mathematical proof see Appendix A.1.

derivative $p_a(1, 0) = \alpha(n - 1)w > 1$, then it follows that a unique solution of the doping threshold $a^\star$ must exist over the support $\in [0, 1]$. ■

The next step is to derive the performance function and the doping amount of the athletes. The athlete's performance function is obtained by inserting Equation (4) in Equation (7). The result is Equation (3), which is the equilibrium behavior of the athletes.

Besides the performance function, the doping behavior of the athletes is of interest. As the performance function $p(a)$ is the sum of ability $a$ and the doping quantity $d$, the doping function can be simply derived from the performance function. The doping quantity is an athlete's performance minus his ability. Having determined the doping threshold, the performance function and the doping quantity, we can now describe the equilibrium behavior of athletes in the doping model. Section 4 will discuss the outcomes of the doping model in more detail. Using comparative statics and figures, we will show how different values of $\alpha$, $n$ and $w$ influence the doping threshold, the performance function and the doping function.

# 4 Discussion

In our model, athletes usually choose an amount of doping such that the marginal costs of doping are equal to the marginal expected increase in prize revenue. Depending on the distribution of abilities, however, it is possible that the marginal return to be gained by doping is smaller than the marginal costs of doping. In the leading case, weak athletes abstain from doping, because it would cost more to imitate the performance of a slightly stronger athlete than the extra return from having a higher probability of winning. Thus, we identify a discouragement effect for weak athletes, similar to the discouragement effect in Lazear-Rosen tournament models with head-starts or handicaps (see e.g., Weigelt et al. (1989) and Schotter and Weigelt (1992)). Crucial for an athlete's doping decision is his actual ability and the shape of the winning probability function $G$. The function becomes more convex, the more competitors there are. This is the reason that a threshold value exists for the majority of parameterizations, below which it is optimal to abstain from doping.

## 4.1 The Number of Athletes and the Costs of Doping

In discussing the results, we first investigate the effects of $n$ and $w$. For this, we will assume that $\alpha = 1$. The abilities are uniformly distributed between 0 and 1. First, we will consider the doping threshold. With a uniform distribution of abilities, Equation (4) becomes

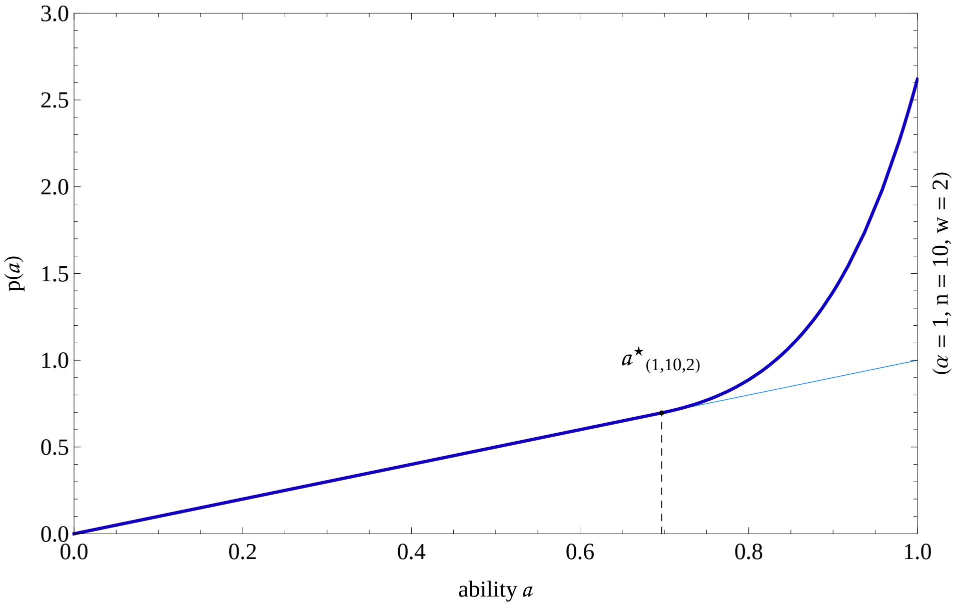
$$a^{\star}_{(\alpha=1,n,w)} = [w(n-1)]^{\frac{1}{2-n}} . \tag{8}$$

The number of competitors is decisive for the size of the doping threshold. If only two athletes compete, then the function of the doping threshold does not exist. In a two-player contest, the performance function consequently has the following appearance: $p(a) = wa$ for $w > 1$. The doping quantity, thus, increases linearly in line with ability. If $w \leq 1$, the two athletes do not dope, because doping is too expensive. Hence, either both athletes dope or neither dopes.

In contests with more than two athletes, a doping threshold does exist. An athlete is indifferent between doping and not doping if his ability is equal to the threshold. The more athletes participate, the higher the value of the doping threshold. A higher number of participants makes it increasingly unattractive for weak athletes to dope. Furthermore, the ratio of prize money to marginal doping costs $w$ determines the level of the doping threshold. If $w$ is large, then the participants can win a large amount of prize money and the doping costs are relatively low. If $w$ becomes smaller, that is, the relative doping costs increase, then the doping threshold rises.

Figure 3 shows the performance function for a contest between 10 athletes and a ratio of prize money to marginal doping costs of 2. The performance function (thick line) represents an athlete's ability up to the threshold. Beyond this point, the athlete's performance is greater than his ability. The doping threshold lies at approximately 0.7. An athlete's doping amount is the difference between his performance and his ability. Up to the doping threshold, the doping amount is equal to zero. Doping does not pay in this array, and athletes' true abilities determine the outcome. For abilities above the doping threshold, the doping function has a positive value. In other words, athletes above this threshold will dope. The more talented such an athlete is, the more he will dope, so that the athlete with the greatest ability will dope the most.

Figure 3: Performance function $p(a)$




Under standard parameterizations, an athlete will have difficulties winning without resorting to doping. In order to illustrate this, we use the example of an athlete with ability $a_{\text{fair}} = 1$. This athlete always wins in a world where no athlete resorts to doping. However, if he decides not to dope, his chances of winning falls dramatically. In order to calculate his chance of winning in this contest, we need ability $a^{\circ}$, for which the performance function p($a^{\circ}$) is equal to 1.[16] Our best athlete, who is also honest, is only able to win against athletes whose $a < a^{\circ}$. His likelihood of winning corresponds to the highest order statistic of $a^{\circ}$. $\text{G}(a_{\text{fair}} = 1) = a^{\circ\, n-1} = 0.19$. In a doping environment, his chance of winning thus falls from 100 percent to just 20 percent. This low winning probability is due to the fact that he does not simply have to compete against one competitor whose ability should be smaller than his ($a_i < a^{\circ}$), but has to win against all nine competitors.

Commercialization has caused prize money in certain sports to surge and has allowed successful athletes to skim additional cash from private companies.

[16] We set the performance function equal to 1 and then solve for $a$. For the contest with $n = 10$ and $w = 2$ we get: $a^{\circ} = 0.832$.

Nowadays, it is quite common, that companies employ the images of successful athletes in corporate sponsoring events to position a brand or to ameliorate their images. Thus, the ratio $w$ seems to have increased rather than decreased over the last two decades. In our model, a larger $w$ leads to a smaller doping threshold and raises the amount of doping substances used. The implications of our model are supported by the observation that there are more doping cases in popular sports than in sports where the prize amount is lower, or where doping offers only a small competitive advantage.

## 4.2 The Distribution of Abilities

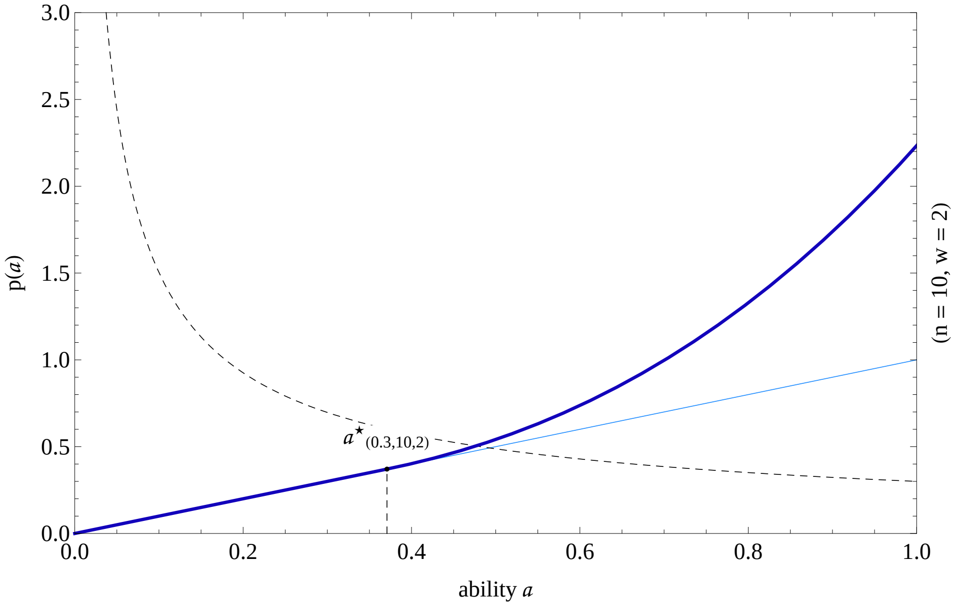
The nature of the distribution affects the doping threshold and the equilibrium amount of doping substances used. The power distribution with an arbitrary $\alpha$ illustrates different distributions of abilities. For $\alpha < 1$, the density of the ability distribution is the highest for small $a$'s, while on the other hand for $\alpha > 1$, there are relatively more strong than weak athletes. We accentuate the importance of the distribution of abilities by contrasting the outcome of a contest of numerous strong athletes with the outcome of a contest of numerous weak athletes.

Figure 4 compares two different distributions of abilities. In both cases, 10 athletes compete against each other, and the ratio of prize money to marginal doping costs is equal to 2. The dashed line is the density of athletes' abilities. In Graph (a), the density of the power function distribution has a parameter value of $\alpha = 0.3$, and in Graph (b) it has one of $\alpha = 3$. The thin line is the performance of the athletes when they do not dope. The athlete's behavior–when doping takes place–is illustrated by the performance function.

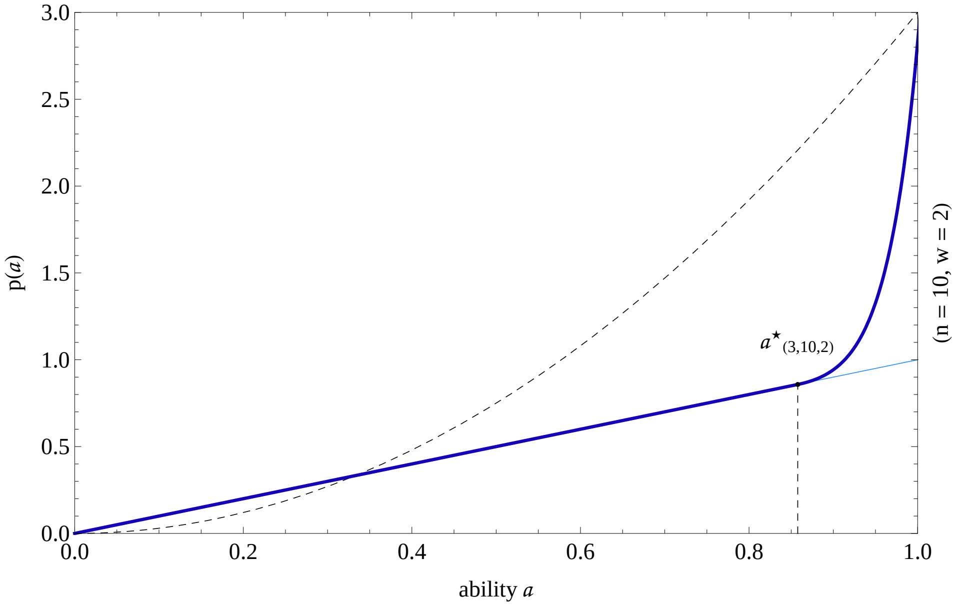
Graph (a) illustrates the outcome of a contest with numerous weak athletes. Here, the doping threshold is very low. Since there are only a few strong athletes, even relatively weak athletes make use of doping. The explanation for this result is that the probability of coming up against a stronger athlete in the contest is relatively small. With a low probability of strong athletes, the performance function increases gradually.

Figure 4: Comparison of different distributions

(a) Many weak contestants: $\alpha = 0.3$




(b) Many strong contestants: $\alpha = 3$




Graph (b) illustrates the outcome of a contest with numerous strong athletes. Here, doping behavior is quite different in comparison to the first case. The doping threshold is higher. This is because a weak athlete can expect to come up against a stronger competitor, given the higher probability of strong athletes. The performance function increases much more sharply than it did in Graph (a). The reason is that the degree of competition is more intense among strong athletes. This leads the strongest athlete to take more doping substances than he would have in the first case. In Graph (a), the performance of the strongest athlete is 2.2, while it is above 2.6 in Graph (b). Hence, although the doping threshold is further to the right, the actual performance of an athlete with $a = 1$ is greater than in a contest with numerous weak athletes.

This section shows that in addition to the number of athletes and the ratio of prize money to marginal doping costs, the distribution of abilities also plays a significant role. An athlete's strength relative to his opponents and, as in Dilger and Tolsdorf (2010), the competitive pressure which weighs on an athlete from competitors with similar abilities are crucial for the doping decision. However, note that a high doping threshold does not always imply that strong athletes dope less than in circumstances with a lower doping threshold. In summary, competitive pressure and the distribution of abilities play decisive roles in an athlete's choice of the optimal doping quantity.

## 4.3 Special Cases

All parameter values for $\alpha$, $n$ and $w$ that meet the assumptions of the leading case lead to an equilibrium outcome where weak athletes abstain from doping. Relaxing the conditions of the distribution of abilities and allowing for extreme values of the ratio of prize money to doping costs makes three other outcomes possible. The leading case arises in the majority of underlying parameter values. The other three cases are special cases that emerge only under exceptional circumstances: if doping costs are very high or low, or if the distribution of abilities is extremely skewed to the left.

For the prevailing equilibrium outcome, the derivative of Equation (7) with respect to $a$ plays a central role. The derivative is the expected marginal gain divided by the marginal costs of doping. The second derivative shows that the first derivative is increasing if $\alpha(n-1) > 1$ and decreasing if $\alpha(n-1) < 1$ for every $a \in (0, 1]$. This critical value is the key to distinguish between the different doping outcomes. The underlying parameter values of $\alpha$, $n$ and $w$

determine which outcome arises. Three additional outcomes are possible. In the first case, no athlete dopes. In the second case, only weak athletes dope. And in the third case, everybody dopes.[17]

The outcome that nobody dopes results if the ratio of prize money to marginal doping costs is low in comparison to the product of the shape parameter and the number of rivals. Nobody dopes because the marginal doping costs are greater than the marginal increase in the expected return. Hence, everybody would lower his expected utility by doping.

**Proposition 3** *If $\alpha(n-1) \geq 1$ and $\alpha(n-1)w \leq 1$, then no doping threshold $a^\star \in [0,1]$ exists and nobody dopes.*

**Proof.** The derivative of Equation (7) with respect to $a$ at $a^r = a$ is continuous and strictly increasing in $a$ if $\alpha(n-1) > 1$. The derivative at $a = 0$ is zero. If the derivative $p_1'(1,1) = \alpha(n-1)w < 1$, then it follows that no doping threshold $a^\star$ exists over the support $[0,1]$. ■

The two other outcomes arise when the distribution of abilities is highly skewed to the left. In case (ii) only weak athletes dope, and in case (iii) everybody dopes. We investigate equilibrium outcomes for distributions of abilities that satisfy the inequality $\alpha(n-1) < 1$. Then, the derivative of Equation (7) is decreasing. Furthermore, $p_1(0, a^r)$ is not defined under this condition. It is the case that $\lim_{a\to 0} p_1(a, a^r) = \infty$. Therefore, the nonnegativity constraint is not binding for small $a$'s. Depending on the underlying parameter values, it is possible that the constraint is not binding over the whole support. In such a case, the equilibrium performance function is Equation (6). Generalizing the performance function in such a way that it displays the equilibrium outcomes of case (ii) and (iii), gives

$$p(a) = \max\{wa^{\alpha(n-1)}, a\} \qquad \text{for} \quad 0 \leq a \leq 1. \tag{9}$$

The doping quantity is the difference between an athlete's performance and his ability. In the equilibrium outcome, only weak athletes take performance-enhancing drugs. The formula of the doping threshold in case (ii) differs from

[17] Appendix A.2 presents the conditions, the parameter values have to meet for each of the three special cases.

the leading case. The doping threshold in case (ii) is given as follows:

$$a^{\star}_{\alpha,n,w} = w^{\frac{1}{1-\alpha(n-1)}}. \tag{10}$$

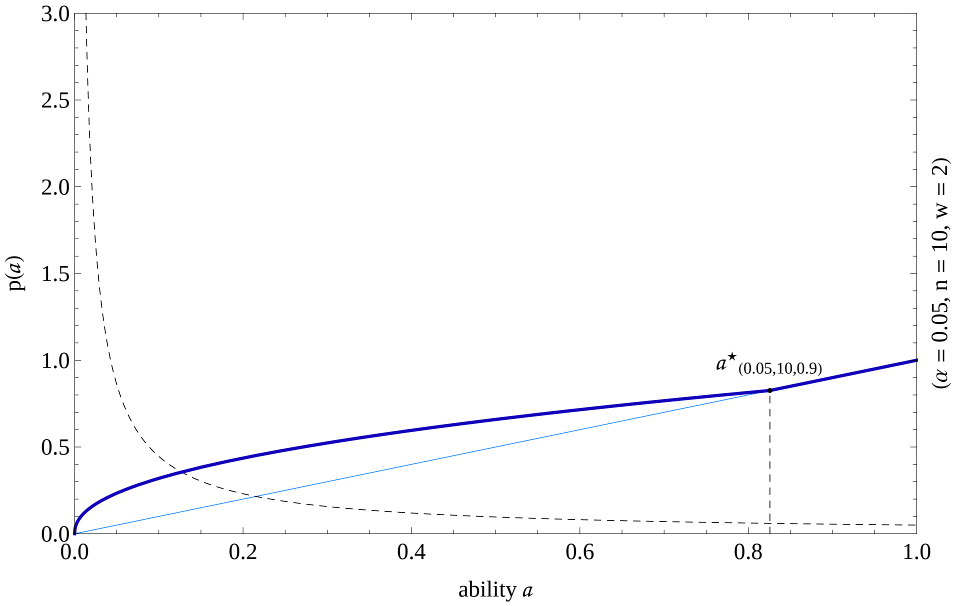
Proposition 4 presents the outcome of case (ii) and (iii) and the underlying conditions.

**Proposition 4** *If $\alpha(n-1) < 1$, then two doping outcomes are possible, depending on the underlying value of $w$. In the first outcome (ii), only weak athletes dope. If $w \leq 1$, then there exists a unique doping threshold $a^{\star} \in [0,1]$ and only weak athletes have an incentive to dope.*
*In the second outcome (iii), everybody dopes. If $w > 1$, then no doping threshold exists over the support $[0,1]$, and every athlete has an incentive to dope.*

**Proof.** If $\alpha(n-1) < 1$, then the derivative of Equation (7) with respect to $a$ for $a^r = 0$ is strictly decreasing in $a$. Furthermore, the derivative $p_1(0, a^r)$ is not defined. It can be shown that $\lim_{a \to 0} p_1(a, a^r) = \infty$. First, consider special case (ii). The performance function $p(a, 0)$ at $a = 1$, not considering the doping nonnegativity constraint, would be $p(1,0) = w1^{\alpha(n-1)} = w \leq 1$. Therefore, there exists a unique intersection point with the 45-degree line. Hence, there exists a doping threshold $a^{\star}$ over the support $\in [0,1]$.
Using the same reasoning for special case (iii), no doping threshold exists over the support $[0,1]$ if $w > 1$. ■

Figure 5 displays the performance function $p(a)$ in the special case (ii). The dashed line is the density distribution of abilities. The distribution is very skewed to the left, which implies that mostly weak athletes compete in the contest. Ten athletes participate in the contest, and the ratio of prize money to marginal doping costs is assumed to be 2. The equilibrium outcome is that weak athletes dope. The doping amount starts to decrease after a certain ability level and is zero for abilities above the doping threshold. The reason for this outcome is that the degree of competition is greatest for weak abilities. Strong athletes do not dope, because the probability of encountering a stronger athlete and the ratio of prize money to marginal doping costs are so small that it is not optimal to increase their performance over their ability level under private information.

Figure 5: Weak athletes dope, special case (ii)




At the doping threshold, the performance function has a kink. Having argued in the leading case that there can be no kink at the doping threshold, in the special case (ii) this is different. Given that all other athletes play the symmetric equilibrium of the performance function in Equation (9), it can be shown that an athlete with an ability $a_i$ below the doping threshold is indifferent to imitating the performance of an athlete with an ability $a_j < a^\star$. If he would instead imitate the performance of an athlete with an ability $a_j > a^\star$, his utility decreases. For athletes with an ability above the doping threshold the nonnegativity constraint is binding. This implies that the marginal return of increasing the performance is lower than the marginal costs of doping.

# 5 Conclusion

We study the doping behavior in an environment of heterogeneous agents and private information. Our setting yields new insights which complement existing results of the doping literature. For the majority of underlying parameter values a doping threshold exists. In our leading case, weak athletes will abstain from doping even without doping controls. Athletes with abilities above the doping threshold resort to doping substances. The doping behavior of athletes and the level of the doping threshold are sensitive to the underlying parameterization. Three other equilibrium outcomes occur when we investigate nonstandard parameterizations. In this paper, we restrict our attention to doping in sport contests. However, our private information setting may also be of interest in areas such as promotion tournaments or public procurement.

Our findings can be summarized as follows. First, an athlete's doping decision depends on the ratio of prize money to marginal doping costs and not on absolute values. In the doping model, a higher ratio decreases the doping threshold and more athletes dope. Thus, the ongoing commercialization and new discoveries of the pharmaceutical industry have increased the incentives to resort to doping. Second, increasing the number of athletes affects the doping decision in two ways. The increased competition forces strong athletes to take larger amounts of drugs. At the same time a discouragement effect exists for weak athletes. The anticipation of facing a higher probability of encountering stronger athletes discourages weak athletes from doping. Finally, our comparison of outcomes between a contest with many weak athletes and a contest with many strong athletes shows that competitive pressure and the distribution of abilities play decisive roles in athletes' doping behavior.

The results of our model would be even more convincing if we could test our findings empirically. Unfortunately, hardly any empirical studies about doping exist, since doping is not directly observable. However, there are indications that support our findings. Empirical evidence shows that there are more positive doping cases in commercial sports with high prize amounts and where athletes receive substantial payments from sponsorships. On the other hand, doping is only rarely detected in technical sports such as tennis, where the use of performance-enhancing drugs helps only marginally.

A promising extension of our model would be the inclusion of a regulator who checks the pool of athletes. But, such a model that combines our private information setting with the asymmetric information problem between athletes and the regulator would be demanding. This is because we would have to incorporate disqualifications into our model, which makes closed-form solutions impossible.

# Bibliography

Amann, E. and W. Leininger (1996). Asymmetric all-pay auctions with incomplete information: The two-player case. *Games and Economic Behavior 14*(1), 1–18.

Berentsen, A. (2002). The economics of doping. *European Journal of Political Economy 18*(1), 109–127.

Berentsen, A., E. Bruegger, and S. Loertscher (2008). On cheating, doping and whistleblowing. *European Journal of Political Economy 24*(2), 415–436.

Berentsen, A. and Y. Lengwiler (2004). Fraudulent accounting and other doping games. *Journal of Institutional and Theoretical Economics (JITE) 160*(3), 402–415.

Bird, E. J. and G. G. Wagner (1997). Sport as a common property resource: A solution to the dilemmas of doping. *Journal of Conflict Resolution 41*(6), 749–766.

Curry, P. A. and S. Mongrain (2009). Deterrence in rank-order tournaments. *Review of Law & Economics 5*(1), 723–740.

David, H. and H. Nagaraja (2003). *Order Statistics* (Third ed.). New Jersey: Wiley.

Dilger, A., B. Frick, and F. Tolsdorf (2007). Are athletes doped? Some theoretical arguments and empirical evidence. *Contemporary Economic Policy 25*(4), 604–615.

Dilger, A. and F. Tolsdorf (2010). Doping und Wettbewerbsintensität. *Schmollers Jahrbuch (Journal of Applied Social Science Studies) 130*(1), 95–116.

Eber, N. (2008). The performance-enhancing drug game reconsidered: A fair play approach. *Journal of Sports Economics 93*(3), 318–327.

Eber, N. (2011). Fair play in contests. *Journal of Economics 103*(3), 253–270.

Eber, N. and J. Thépot (1999). Doping in sport and competition design. Discussion Papers (REL - Recherches Economiques de Louvain) 1999044, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Feess, E., G. Muehlheusser, and M. Walzl (2008). Unfair contests. *Journal of Economics 93*(3), 267–291.

Gilpatric, S. M. (2011). Cheating in contests. *Economic Inquiry 49*(4), 1042–1053.

Haugen, K. K. (2004). The performance-enhancing drug game. *Journal of Sports Economics 5*(1), 67–86.

Kovenock, D., M. R. Baye, and C. G. de Vries (1996). The all-pay auction with complete information. *Economic Theory 8*(2), 291–305.

Kräkel, M. (2007). Doping and cheating in contest-like situations. *European Journal of Political Economy 23*(4), 988–1006.

Krishna, V. (2002). *Auction Theory*. San Diego: Academic Press.

Maennig, W. (2002). On the economics of doping and corruption in international sports. *Journal of Sports Economics 3*(1), 61–89.

Muehlheusser, G. (2006). Implications from the theory of contests for modelling and designing sports competitions. In V. Andreff and S. Szymanski (Eds.), *Handbook on the economics of sport*, pp. 342–348. Edward Elgar.

Schotter, A. and K. Weigelt (1992). Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. *The Quarterly Journal of Economics 107*(2), 511–539.

Stowe, C. J. and S. M. Gilpatric (2010). Cheating and enforcement in asymmetric rank-order tournaments. *Southern Economic Journal 77*(1), 1–14.

Strulik, H. (2012). Riding high: Success in sports and the rise of doping cultures. *Scandinavian Journal of Economics 114*(2), 539–574.

USADA (2012, [online]). U.S. postal service pro cycling team investigation. URL http://cyclinginvestigation.usada.org. Accessed 28 November 2012.

Verroken, M. (2005). Drug use and abuse in sport. In D. R. Mottram (Ed.), *Drugs in sport* (Fourth ed.)., Chapter 2, pp. 29–63. Abingdon: Routledge Taylor & Francis Group.

WADA (2009, [online]). World anti-doping code. URL http://www.wada-ama.org. Accessed 11 January 2013.

Weigelt, K., J. Dukerich, and A. Schotter (1989). Reactions to discrimination in an incentive pay compensation scheme: A game-theoretic approach. *Organizational Behavior and Human Decision Processes 44*(1), 26–44.

# A Appendix

## A.1 Proof of Theorem 1

The equilibrium performance function (Equation (3)) with the doping threshold (Equation (4)) is a Nash equilibrium if no athlete with ability $a_i$ for all $a_i \in [0, 1]$ is better off by unilaterally deviating from the equilibrium strategy. Note that the performance function is strictly increasing in $a$. It follows that, if an athlete imitates the strategy of an athlete with ability $a_j$, his winning probability is $G(a_j)$. By using the term imitating, we mean that the athlete choses the performance $p(a_j)$ of an athlete with ability $a_j$. Note that an athlete with ability $a_i$ can only imitate athletes with a performance greater than his ability ($p(a_j) \geq a_i$). Thus, we can limit the verification on the range $p(a_j) \geq a_i$, for all $a_i \in [0, 1]$.

Before starting with the proof, we ask the question: How large would the doping costs have to be for an athlete with ability $a_i$ to win with probability $G(a_j)$, such that he would be indifferent to his equilibrium strategy? An athlete is indifferent if the difference of the winning probability times prize money is equal to the difference of doping costs. We denote the performance level where the athlete would be indifferent by $\hat{p}_{a_i}(G(a_j))$. The performance level satisfies the following equation:

$$vG(a_j) - c(\hat{p}_{a_i}(G(a_j)) - a_i) = vG(a_i) - c(p(a_i) - a_i). \qquad (11)$$

Generalizing Equation (11) for an arbitrary $a \geq a_i$ gives an indifference function.

**Definition 2** *The indifference function $\hat{p}_{a_i}(a)$ indicates the performance level at which an athlete with ability $a_i$ is indifferent to his equilibrium strategy $p(a_i)$ if he were to win with the probability of an athlete with ability $a$. Function $\hat{p}_{a_i}(a)$ is defined in the range $a \in [a_i, 1]$. More formally the function is*

$$\hat{p}_{a_i}(a) = p(a_i) + w[G(a) - G(a_i)] = p(a_i) + w[a^{\alpha(n-1)} - a_i^{\alpha(n-1)}] \text{ for } a \geq a_i. \qquad (12)$$

Comparing the functional form of Equation (12) with Equation (7) shows that the only difference is the threshold value. For abilities $a_i > a^\star$, the indifference function is even identical to the equilibrium performance function within the range of $a \in [a^\star, 1]$. This is a direct result of the assumed linear doping cost function. The indifference function illustrates whether an athlete would have an interest in deviating from the equilibrium. If the performance function $p(a)$ runs below (above) $\hat{p}_{a_i}(a)$, then the athlete is better off (worse off) by deviating. The proof of Theorem 1 consists of two steps. First, we prove that the performance function is a Nash equilibrium ((i), (ii)). Second, we prove that the doping threshold $a^\star$ is unique.

**Proof.** For athletes with ability $a_i$ above the doping threshold, the indifference function is identical to the performance function. (i) Hence, for all $a_i \in [a^\star, 1]$ no athlete can be better off by imitating another performance $(p(a_j) \geq a_i)$.

For athletes with abilities $a_i < a^\star$ below the doping threshold, the indifference function is not identical to the performance function. Note that the derivative with respect to $a$ of $\hat{p}_{a_i}(a)$ is equal to the derivative of Equation (7). For parameterizations of the leading case, the derivative $\hat{p}_{a_i}{}'(a)$ at $a = a_i$ is strictly increasing in $a_i$. Remember that $\hat{p}_{a_i}(a)$ for $a_i = a^\star$ is identical to the performance function. Therefore, the derivative of $\hat{p}_{a_i}(a)$ at $a = a^\star$ is equal to 1. (ii) This implies that $\hat{p}_{a_i}(a)$ runs below the performance function $p(a)$ in the range of $t \in [a_i, 1]$ for all $a_i \in [0, a^\star)$. Hence, an athlete with ability $a_i$ is worse off by imitating the strategy of an ability $a > a_i$.

Finally, we prove that the doping threshold $a^\star$ is unique. Suppose that the optimal threshold $\widetilde{a^\star}$ is smaller than $a^\star$. Then, the performance function would violate the nonnegativity constraint, since the derivative of the performance function is smaller than 1 for all $\widetilde{a^\star} \in [0, a^\star)$. Now suppose that the optimal threshold $\widetilde{a^\star}$ is greater than $a^\star$. Then, $\hat{p}_{a_i}(a)$ of an athlete with ability $a_i$ which is slightly smaller than $\widetilde{a^\star}$ would run above the performance function, since the derivative of $\hat{p}_{a_i}(a_i)$ for $a_i \in (a^\star, 1]$ is greater than 1. Hence, the athlete with ability $\widetilde{a^\star} - \epsilon$ would be better off by deviating from his equilibrium strategy. It follows that the doping threshold $a^\star$ is optimally chosen, and therefore the performance function is a Nash equilibrium. ■

## A.2 The Four Doping Outcomes

The following description presents the different doping outcomes. Moreover, the conditions on the parameterization values of $\alpha$, $n$ and $w$ that lead to the four outcomes are displayed.

Table 1: The doping outcomes

| **(A) Leading Case** | **(B) Special Cases** |
|---|---|
| | **(i) Nobody dopes** |
| | $\alpha \geq (n-1)^{-1} \cap w\alpha(n-1) \leq 1.$ |
| **Strong athletes dope** | **(ii) Weak athletes dope** |
| $\alpha \geq (n-1)^{-1} \cap w\alpha(n-1) > 1.$ | $\alpha < (n-1)^{-1} \cap w \leq 1.$ |
| | **(iii) Everybody dopes** |
| | $\alpha < (n-1)^{-1} \cap w > 1.$ |

# Chapter III

## Energy Policy Challenges and the Input Mix in Swiss Manufacturing*

**Abstract**

Switzerland's energy policy is faced by unseen challenges owing to the scheduled nuclear phaseout and the continuation of the Kyoto protocol. As a small open economy, effects of unilateral energy policies on manufacturing industries may be especially relevant for Switzerland. However, estimates of sectoral substitution elasticities between input factors used to address important questions in this context do not exist for Swiss manufacturing. We close this gap by using a newly assembled data set covering the period from 1997 to 2008 and by estimating substitution elasticities employing a translog cost function. Moreover, we examine the implications of energy price increases putting emphasis on the input mix in Swiss manufacturing industries.

| | |
|---|---|
| Keywords: | Energy Policy, Substitution Elasticities, Swiss Manufacturing |
| JEL Classification: | C33, Q41 |

* This chapter has been written in collaboration with Lukas Mohler from the University of Basel. We thankfully acknowledge the financial support of the Swiss Federal Office of Energy (SFOE). The views expressed in this chapter are those of the authors and do not necessarily represent those of the SFOE.

# 1 Introduction

In 2011, shortly after the incident in Fukushima, the Swiss government and parliament have decided to aim for a nuclear phaseout. At the end of 2012, the Swiss government has furthermore prolonged the reduction targets for greenhouse gas emissions until the year 2020. Only few countries have conclusively decided to abandon nuclear energy production and many countries around the world have not ratified the Kyoto protocol. Switzerland is therefore faced with energy policy challenges unmatched by other countries. The dimension of these challenges is even intensified by the fact that Switzerland, as a small open economy, is highly integrated into the world economy, and that manufacturing exports are one of the cornerstones of Swiss prosperity. Hence, the effects of policy measures largely depends on whether similar measures are implemented in other countries.

The actual implementation of policy instruments aimed at achieving these goals is still rather unclear which makes it hard to assess possible consequences for the Swiss economy in detail. However, the implied future policy adjustments are sure to yield important effects on the demand and the supply side of the energy provision in Switzerland, changing absolute and relative prices of different energy sources and other input factors. Specifically, price increases of oil, fuel or electricity, as they are already announced by the government, will reduce the demand for energy. They will also change the demand for other input factors like labor, capital or material, in turn causing adjustments of the relative use of these inputs.

Elasticities of substitution between input factors measure these adjustments in the input mix of firms and are, therefore, important inputs for policy analysis and forecasting. Such elasticities have been estimated abundantly and the empirical literature has shown that the magnitude of these elasticities depends on a variety of factors. Importantly, these findings indicate that elasticities vary significantly between countries, industries and time periods. It is hence essential that policy evaluations are based on elasticities that are representative for the situation at hand. However, such elasticities of substitution between input factors covering recent years have not been estimated for Swiss manufacturing at a sectoral level, the reason being a lack of data availability.

Switzerland is now in a situation where it is of paramount importance to obtain a better understanding of how manufacturing firms adjust their behavior upon relative changes of factor prices. In this article, we address this issue and estimate sectoral elasticities of substitution between energy, capital, labor and material inputs for Swiss manufacturing covering the period from 1997 to 2008. We evaluate how the substitutability between input factors actually works in firms of different industries and focus on how energy price increases affect the input mix and production costs of manufacturing industries.

The results of our analysis can be summarized as follows. First, employing a translog cost function, we present estimates of own- and cross-price elasticities (CPE) as well as Morishima elasticities of substitution (MES) using a newly assembled data set on Swiss manufacturing industries. We observe that labor and capital are estimated as CPE substitutes in most industries, while the evidence is mixed regarding the substitution pattern regarding energy and capital and particularly regarding energy and labor. All factor inputs are estimated as being MES substitutes. Standard errors of some estimates are, however, relatively large.

Second, and to obtain more precise estimates, we provide a pooled regression using Swiss data as well as additional data on nine high-income OECD countries with economic attributes similar to Switzerland. Standard errors decrease substantially and we observe elasticities of smaller magnitude. The substitution patterns remain similar to the case using Swiss data only, as capital and labor are still estimated as substitutes, while evidence on the substitutability of labor and energy as well as capital and energy is mixed and sector heterogeneity remains substantial. Again, all inputs tend to be MES substitutes with very few exceptions.

Third, we compare our results with results of recent contributions from the elasticity literature, concentrating on the substitutability versus complementarity debate. Estimated substitution patterns are generally similar to other studies. However, sectoral heterogeneity is substantial and hence, individual sector results may largely differ in comparison to studies analyzing other countries or time periods.

Fourth, we show how an increase of the price of energy affects the input mix of manufacturing industries by interpreting the estimated CPE and MES.

Again, sector heterogeneity is large, but in an average manufacturing firm, energy price increases lead to a reduction in labor expenses and an increase of material and especially of capital expenses. The manufacturing-wide reduction of labor expenses upon price increases of energy and its compensation by capital and material inputs is one important take-away of our contribution. We also show that in a typical firm, the *physical use ratios* of energy relative to the use of the other factors decrease. However, due to the inelastic nature of energy own-price elasticities in the more reliable OECD regressions, the *cost ratios* of energy relative to the other factors increase on average.

The remainder of this article is structured as follows. In Section 2, we discuss the theoretical model employed and illustrate how economic elasticities are calculated. In Section 3, we provide an overview of the data set and present the estimation strategy. Section 4 discusses the regression results and illustrates the substitution patterns of manufacturing sectors. In Section 5, we analyze how energy price increases affect the input mix and costs of manufacturing industries and discuss policy implications. Section 6 concludes.

## 2 Modeling Approach

We assume an industry-specific production technology with input factors capital (K), labor (L), energy (E), and material (M), its cost function taking the form of a transcendental logarithmic (translog) function. The translog cost function was proposed in Christensen et al. (1973) and subsequently adopted in a seminal paper by Berndt and Christensen (1973) to estimate the substitution patterns in U.S. manufacturing. Since it does not impose any prior constraints on the elasticities and on the optimal path of input factor adjustments, it is one of the preferred functional forms used in the literature.[1]

It embodies the following important implications for parameterization. First, the sum of the factor shares equals one in every sector, and second, symmetry of the cross-price derivatives must be satisfied ($\beta_{ij} = \beta_{ji}$ for $i, j = 1, .., I$). The

[1] It is standard to estimate elasticities from the translog model using cost functions instead of production functions. One reason for this approach is that no arbitrary restrictions of the production patterns have to be assumed (Jorgenson, 1986). Another reason is that the data requirements are less rigorous than for estimating production functions, since cost functions only require data on factor shares and factor unit prices.

first property induces singularity, because the sum of the error terms of the four factor shares has to equal zero for every sector. We obtain a non-singular system by normalizing the price of material which allows us to omit the factor share equation of this input.[2] The second property is satisfied by imposing constraints directly on the regression equation (see Section 3). Furthermore, we note that the necessary concavity properties are not generally met using translog cost functions. We test for this property in Section 4.

The factor share equations can be obtained by differentiating the logarithm of the translog cost function with respect to the logarithm of the prices and are given by

$$s_{in,t} = \beta_{in} + \sum_{j} \beta_{ijn} \ln(p_{jn,t}) + \beta_{iny} \ln y_{n,t} + \beta_{in,t} t, \tag{1}$$

where $i$ and $j$ are indices for input factors, $n$ is the industry index and $t$ is a time index. Factor share $s_{in,t}$ of input $i$ in industry $n$ thus equals the sum of a constant $\beta_{in}$ and the own-price and cross-price derivatives $\beta_{ijn}$ (also referred to as share elasticities) multiplied by the logarithm of the input prices. Moreover, we add the logarithm of output $y_{n,t}$ and a time trend $t$. The coefficient $\beta_{iny}$ allows for various forms of returns to scale, and the time trend parameter $\beta_{in,t}$ additionally permits us to account for non-neutral technological change.

Three types of economic elasticities derived from the estimates of $\beta_{ijn}$ are commonly used in the literature: the cross-price elasticity, the Morishima elasticity of substitution and the Allen partial elasticity of substitution (AES) which is a normalized version of the CPE. We refrain from calculating AES, as it has been argued by many scholars that they do not yield any useful interpretation in the case of more than two input factors.[3]

The $\text{CPE}_{x_i,p_j}$, measures the relative change in quantity $x_i$ of a production factor $i$ due to the relative price change in $p_j$ of a factor $j$. In the remainder of the paper, we denote such elasticities by $\text{CPE}_{ij}$, where the first index indicates the factor whose quantity is affected, and the second one the factor experiencing the initial price change. These elasticities can be derived from

[2]The elasticities of the input factor material can be derived by using the imposed restrictions.
[3]See for example Thompson and Taylor (1995), Blackorby and Russell (1989) or Frondel (2011).

the estimated share elasticities ($\beta_{ij}$) and the observed factor shares as follows:[4]

$$\text{CPE}_{ij} \equiv \frac{\partial \ln x_i}{\partial \ln p_j} = \frac{\hat{\beta}_{ij} + s_i s_j}{s_i}, \qquad \text{CPE}_{ii} \equiv \frac{\partial \ln x_i}{\partial \ln p_i} = \frac{\hat{\beta}_{ii} + s_i s_i - s_i}{s_i}, \tag{2}$$

where the second expression is used to calculate the own-price elasticities.

The $\text{MES}_{x_i/x_j,p_j}$ measures the relative change in the quantity ratio $x_i/x_j$ of production factors $i, j$ due to a relative price change in $p_j$ of factor $j$. We denote these elasticities by $\text{MES}_{ij}$. To obtain the MES, the own-price elasticity is subtracted from the CPE as follows,

$$\text{MES}_{ij} \equiv \frac{\partial \ln(x_i/x_j)}{\partial \ln p_j} = \text{CPE}_{ij} - \text{CPE}_{jj}. \tag{3}$$

Following Koetse et al. (2008), the standard errors of the elasticities can be derived by using the Delta Method (see Appendix A.1).

Using Equation (3) and remembering that the own-price elasticity is always negative, it follows that $\text{MES}_{ij} > \text{CPE}_{ij}$. The MES primarily indicates how factor intensities in production change upon price increases. Moreover, the MES can also be used to assess the change in the relative cost shares between two factors $i$ and $j$ upon a price increase in factor $j$:

$$\begin{aligned} \frac{\partial \ln(p_i x_i / p_j x_j)}{\partial \ln p_j} &= \frac{\partial \ln(p_i x_i)}{\partial \ln p_j} - \frac{\partial \ln(p_j x_j)}{\partial \ln p_j} = \frac{\partial \ln x_i}{\partial \ln p_j} - 1 - \frac{\partial \ln x_j}{\partial \ln p_j} \\ &= \text{CPE}_{ij} - \text{CPE}_{jj} - 1 = \text{MES}_{ij} - 1, \end{aligned} \tag{4}$$

where $p_i x_i / \sum_k p_k x_k = s_i$. Hence, the share ratio $s_i/s_j$ increases in $p_j$ if $\text{MES}_{ij} > 1$.

# 3 Data and Estimation Strategy

We use industry-specific data containing information on cost shares, prices and deflators of the input factors capital, labor, energy and material, as well as on

[4] Anderson and Thursby (1986) argue in favor of using the means of factor shares, since this provides desirable properties regarding the distribution of the elasticity estimators. We omit the industry index $n$ for brevity's sake. As $s_i$ and $s_j$ we use factor share averages computed over the full time span of our analysis, see Table 2.

Table 1: Definition of manufacturing industries

| Aggregates | (industries[a]) | Description | (Short name) |
|---|---|---|---|
| 1 | (15,16) | Food products and beverages | (Food) |
| 2 | (17,18,19) | Textiles, textile products, leather and footwear | (Textiles) |
| 3 | (20) | Wood and products of wood and cork | (Wood) |
| 4 | (21,22) | Pulp, paper products, printing and publishing | (Pulp, paper) |
| 5 | (24) | Chemicals and chemical products | (Chemicals) |
| 6 | (25) | Rubber and plastic products | (Rubber, plastics) |
| 7 | (26) | Other non-metallic mineral products | (Other non-metals) |
| 8 | (27,28) | Basic metals and fabricated metal products | (Metals) |
| 9 | (29) | Machinery and equipment | (Machinery) |
| 10 | (30,...,33) | Electrical and optical equipment | (Electrical equip.) |
| 11 | (34,35) | Transport equipment | (Transport equip.) |

Notes: [a] Industries according to NOGA 2002 industrial classification of Switzerland, 2-digit.

revenue and production, covering Swiss manufacturing sectors for the period 1997–2008. To increase the sample size, we also estimate a specification where we pool across countries using additional data on nine OECD countries from 1995 to 2005, i.e., from Belgium, Denmark, France, Germany, Great Britain, Japan, the Netherlands, Spain and Sweden. The Swiss Federal Statistical Office (SFSO) and the Swiss Federal Office of Energy (SFOE) are the main sources of the data on Swiss manufacturing industries. Data for the other countries stem from the database EU KLEMS.[5] All data are generally available at the ISIC 3.1 2-digit level. While this classification defines 23 manufacturing industries, we aggregate these ISIC industries into 11 industries due to data limitations.[6] Table 1 provides an overview of these industries.

Table 2 displays the average cost shares of capital, labor, energy and material in Swiss manufacturing industries. In addition, averages over the full country sample, labeled OECD, is displayed. While average factor shares in Switzerland are of similar magnitude, we note that labor and material shares

[5]An extensive description of the KLEMS database can be found in Timmer et al. (2007) and O'Mahony and Timmer (2009).

[6]We exclude ISIC sectors 23 (manufacture of coke, refined petroleum products), 36 (furniture and other manufacturing) and 37 (recycling) due to the lack of data availability in Switzerland.

Table 2: Average input cost shares

| | Industry | Capital CHE | Capital OECD | Labor CHE | Labor OECD | Energy CHE | Energy OECD | Material CHE | Material OECD |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Food | 18.6% | 25.4% | 17.7% | 16.5% | 1.4% | 1.8% | 62.3% | 56.3% |
| 2 | Textiles | 18.9% | 23.3% | 30.0% | 27.1% | 1.8% | 2.0% | 49.3% | 47.5% |
| 3 | Wood | 20.5% | 24.0% | 33.6% | 26.0% | 1.3% | 2.8% | 44.5% | 47.2% |
| 4 | Pulp, paper | 24.7% | 31.4% | 34.4% | 27.1% | 3.5% | 2.6% | 37.5% | 38.9% |
| 5 | Chemicals | 35.7% | 37.2% | 16.7% | 17.1% | 1.5% | 5.2% | 46.1% | 40.5% |
| 6 | Rubber, plastics | 21.5% | 26.8% | 28.4% | 25.9% | 1.4% | 3.1% | 48.7% | 44.3% |
| 7 | Other non-metals | 25.2% | 32.0% | 29.7% | 27.1% | 4.5% | 6.1% | 40.6% | 34.7% |
| 8 | Metals | 22.4% | 23.4% | 34.3% | 26.5% | 2.3% | 3.6% | 41.0% | 46.5% |
| 9 | Machinery | 19.7% | 24.8% | 28.7% | 27.4% | 0.6% | 1.2% | 51.0% | 46.5% |
| 10 | Electrical equipm. | 25.8% | 28.0% | 26.2% | 24.5% | 0.8% | 1.1% | 47.2% | 46.5% |
| 11 | Transport equipm. | 16.7% | 20.0% | 28.6% | 19.5% | 1.4% | 1.1% | 53.2% | 59.5% |

Notes: Swiss average shares are contrasted with the overall OECD average shares including data from Switzerland, Belgium, Denmark, France, Germany, Great Britain, Japan, the Netherlands, Spain and Sweden.

are mostly above the OECD average, while capital and energy intensities are slightly below. This pattern is quite intuitive given the high labor costs of skill-intensive production and the relatively low capital costs in Switzerland. Section A.2 in the Appendix describes the data in more detail.[7]

To estimate the share elasticities of the translog cost functions as in Equation (1), we use a Seemingly Unrelated Regression (SUR) approach to take the cross-correlations between the error terms (covariance of the disturbances) in the system of factor share equations into account. If the error terms are correlated, SUR provides more efficient estimates than separate OLS estimation.[8]

We directly impose the symmetry restrictions on the estimation procedure by setting up constraints in the form of $\beta_{ij} = \beta_{ji}$. Furthermore, we estimate different specifications using present and lagged values of input prices to ac-

[7]Furthermore, detailed description of the Swiss data set can be found in Mohler and Müller (2011) and Mohler and Müller (2012). These SFOE reports are available on *http://www.ewg-bfe.ch*.

[8]Due to the limited number of time-series observations, we are not able to estimate the coefficients simultaneously for all industries of Swiss manufacturing. However, we are still able to estimate the three share equations of each industry in one step. In the pooled regression, we estimate all share equations of all industries simultaneously.

count for inertia in the adjustment process of the input factors as well as for endogeneity issues. We find that the concavity property (also see next section) is best met when using a one-year lag, and we therefore use this specification in the present paper. In the pooled specification, we additionally add country fixed effects to achieve results that are comparable to the Swiss estimates.[9]

# 4 Results

We first present the sectoral estimates for Swiss manufacturing before discussing the pooled OECD regressions. CPE and MES for labor, capital and energy are discussed in turn.[10] Key statistics of the estimations are shown in Appendix A.4.

## 4.1 Estimates from Swiss Data

Table 3 displays own-price elasticities for manufacturing industries in Switzerland. Well-behaved production functions have to satisfy the concavity property, i.e., own-price elasticities are required to be negative. The table illustrates that while there is one positive point estimate of the labor own-price elasticities, it is not significantly different from zero and hence the concavity property is met.

To make an example regarding the interpretation of these elasticities, the own-price elasticity of energy (-1.20) in industry 4, pulp and paper, implies that an energy price increase of one percent induces the quantity of energy utilized

[9] It has been argued that pooling across countries yield long-run elasticities–since a substantial part of cross-section differences may stand for long-term disparities between countries–whereas time-series data yield short-run elasticities, see for example Griffin and Gregory (1976), Berndt and Wood (1979), Griffin (1981) or Berndt and Wood (1981). However, this argument is very controversial, i.e., it is questioned, whether cross-section differences really are a manifestation of long-run effects, see Apostolakis (1990) for a detailed overview. Hence, to compare the pooled results to the Swiss estimates, we include country fixed effects in the pooled regression to prevent the factor share coefficients from taking on cross-section effects and concentrate on the time-series variation.

[10] We refrain from presenting elasticities concerning the factor material. The inclusion of materials into the modeling approach is nonetheless very important to omit biased results for the other inputs as is shown by Berndt and Wood (1979).

to produce a constant amount of output to decrease by 1.20 percent. Since the own-price elasticity is smaller than -1, not only the quantity but also the cost share of this factor will decrease upon a price increase. We also observe that the energy elasticities are in most cases not found to be significantly different from zero. This is also illustrated by Figure 1 in the Appendix.

Table 3: Own-price elasticities, Swiss manufacturing industries

| | Industry | $CPE_{LL}$ | $CPE_{KK}$ | $CPE_{EE}$ |
|---|---|---|---|---|
| 1 | Food | 0.02 | -0.78* | -1.15 |
| | | (0.45) | (0.25) | (1.21) |
| 2 | Textiles[a] | -0.25 | -0.51 | -2.23* |
| | | (0.65) | (0.71) | (0.37) |
| 3 | Wood | -1.58 | -1.14* | -3.18 |
| | | (0.82) | (0.54) | (3.49) |
| 4 | Pulp, paper[a] | -0.84* | -1.47 | -1.20* |
| | | (0.09) | (1.04) | (0.55) |
| 5 | Chemicals | -1.83* | -0.65* | -1.42 |
| | | (0.31) | (0.26) | (1.41) |
| 6 | Rubber, plastics[a] | -1.08* | -0.38 | -1.37 |
| | | (0.26) | (0.23) | (1.88) |
| 7 | Other non-metals | -1.87* | -0.75* | -1.29 |
| | | (0.34) | (0.33) | (0.87) |
| 8 | Metals | -1.30* | -1.31* | -0.87 |
| | | (0.13) | (0.16) | (1.65) |
| 9 | Machinery[a] | -1.06* | -1.29* | -0.27 |
| | | (0.17) | (0.55) | (2.85) |
| 10 | Electrical equipm. | -0.92* | -1.55* | -0.86 |
| | | (0.16) | (0.30) | (1.47) |
| 11 | Transport equipm.[a] | -1.62* | -0.91 | -0.29 |
| | | (0.34) | (0.98) | (3.44) |
| Concavity Test: | | | | |
| Negative[b] | | **10 (8)** | **11 (7)** | **11 (2)** |
| Positive[b] | | **1 (0)** | **0 (0)** | **0 (0)** |

Notes: Swiss sectoral data. Asymptotic standard errors derived with the delta method. * Indicates significance at the 5% level. [a] Specification contains employment as a control variable. [b] Number of significant estimates in parentheses.

Table 4 displays all six CPE between the input factors capital, labor and energy. Considering industry 1, food products, the CPE that evaluates the impact of an energy price change on labor use, $CPE_{LE}$ (-0.20), indicates that if energy prices increase by one percent, the quantity of labor utilized to produce the constant amount of output will decrease by 0.2 percent. As expected, elas-

Table 4: Cross-price elasticities, Swiss manufacturing industries

| Industry | $CPE_{LE}$ | $CPE_{EL}$ | $CPE_{KE}$ | $CPE_{EK}$ | $CPE_{KL}$ | $CPE_{LK}$ |
|---|---|---|---|---|---|---|
| 1 Food | -0.20 | -2.51 | 0.10 | 1.30 | 0.54* | 0.57* |
| | (0.15) | (1.87) | (0.10) | (1.36) | (0.17) | (0.18) |
| 2 Textiles[a] | -0.01 | -0.15 | 0.20* | 2.09* | 0.15 | 0.10 |
| | (0.06) | (1.02) | (0.07) | (0.72) | (0.59) | (0.37) |
| 3 Wood | 0.14 | 3.62 | 0.17 | 2.57 | 0.23 | 0.14 |
| | (0.20) | (5.12) | (0.24) | (3.73) | (0.72) | (0.44) |
| 4 Pulp, paper[a] | 0.01 | 0.06 | 0.16 | 1.17 | 1.48* | 1.06* |
| | (0.04) | (0.39) | (0.18) | (1.28) | (0.26) | (0.18) |
| 5 Chemicals | -0.08 | -0.94 | 0.05 | 1.09 | 0.30* | 0.65* |
| | (0.13) | (1.41) | (0.06) | (1.38) | (0.12) | (0.25) |
| 6 Rubber, plastics[a] | 0.11 | 2.32 | -0.01 | -0.21 | 0.49* | 0.37* |
| | (0.10) | (2.02) | (0.13) | (1.97) | (0.20) | (0.15) |
| 7 Other non-metals | 0.21 | 1.38 | 0.11 | 0.60 | 1.63* | 1.39* |
| | (0.16) | (1.04) | (0.18) | (1.04) | (0.25) | (0.21) |
| 8 Metals | 0.09 | 1.34 | 0.03 | 0.25 | 1.11* | 0.73* |
| | (0.07) | (1.10) | (0.10) | (0.98) | (0.14) | (0.09) |
| 9 Machinery[a] | 0.05 | 2.38 | -0.00 | -0.09 | 1.13* | 0.78* |
| | (0.07) | (3.47) | (0.11) | (3.78) | (0.25) | (0.17) |
| 10 Electrical equipm. | -0.12 | -3.97 | 0.04 | 1.17 | 0.74* | 0.73* |
| | (0.07) | (2.35) | (0.04) | (1.32) | (0.08) | (0.08) |
| 11 Transport equipm.[a] | -0.08 | -1.66 | -0.22 | -2.56 | -0.61 | -0.36 |
| | (0.16) | (3.18) | (0.31) | (3.64) | (0.61) | (0.36) |
| Substitutes or Complements: | | | | | | |
| Substitutes[b] | **6 (0)** | **6 (0)** | **8 (1)** | **8 (1)** | **10 (8)** | **10 (8)** |
| Complements[b] | **5 (0)** | **5 (0)** | **3 (0)** | **3 (0)** | **1 (0)** | **1 (0)** |

Notes: Swiss sectoral data. Asymptotic standard errors derived with the delta method. * Indicates significance at the 5% level. [a] Specification contains employment as a control variable. [b] Number of significant estimates in parentheses.

ticities that evaluate the effect of energy price changes (columns one and three) are generally small in magnitude due to the relatively small energy share in production, while relative price changes of more important input factors yield a larger impact on the use of other factors.

Interpreting all point estimates, while elasticities between capital and labor as well as between energy and capital give some evidence for the substitutability of these factors, the picture is less clear regarding energy and labor inputs. Table 4 also reveals that the standard errors of the estimates are relatively large (also see Figure 2 in the Appendix), and most elasticities that are significantly different from zero are found in the last two columns, between capital and labor.

Analogously, Table 5 displays the MES. As these state how the ratio of two input quantities changes due to a price change of the factor in the denominator, a reading example shows that the energy-labor MES in industry 1, $\text{MES}_{LE}$ (0.95), implies that an energy price change of one percent leads to an increase of the labor-energy quantity ratio of 0.95 percent. It is furthermore noteworthy that most MES are positive, implying substitutability between all input factors. Again, the standard errors are relatively large, except for the elasticities in the last two columns of the table. Figure 3 in the Appendix visualizes the corresponding confidence intervals.

Table 5: Morishima elasticities, Swiss manufacturing industries

| Industry | $MES_{LE}$ | $MES_{EL}$ | $MES_{KE}$ | $MES_{EK}$ | $MES_{KL}$ | $MES_{LK}$ |
|---|---|---|---|---|---|---|
| 1 Food | 0.95 | -2.53 | 1.24 | 2.08 | 0.52 | 1.35* |
| | (1.15) | (1.83) | (1.30) | (1.51) | (0.55) | (0.38) |
| 2 Textiles[a] | 2.22* | 0.10 | 2.43* | 2.60* | 0.41 | 0.60 |
| | (0.40) | (1.13) | (0.42) | (1.27) | (0.47) | (0.52) |
| 3 Wood | 3.32 | 5.20 | 3.35 | 3.71 | 1.81 | 1.28 |
| | (3.63) | (5.28) | (3.69) | (3.93) | (1.11) | (0.71) |
| 4 Pulp, paper[a] | 1.21* | 0.91* | 1.36* | 2.64* | 2.33* | 2.54* |
| | (0.56) | (0.40) | (0.51) | (1.18) | (0.30) | (1.11) |
| 5 Chemicals | 1.34 | 0.89 | 1.47 | 1.74 | 2.13* | 1.30* |
| | (1.46) | (1.45) | (1.46) | (1.41) | (0.35) | (0.44) |
| 6 Rubber, plastics[a] | 1.48 | 3.40 | 1.35 | 0.17 | 1.57* | 0.75* |
| | (1.86) | (1.94) | (1.84) | (1.85) | (0.32) | (0.25) |
| 7 Other non-metals | 1.50 | 3.24* | 1.40 | 1.34 | 3.50* | 2.13* |
| | (0.83) | (1.00) | (1.02) | (1.27) | (0.46) | (0.41) |
| 8 Metals | 0.96 | 2.64* | 0.90 | 1.55 | 2.41* | 2.03* |
| | (1.70) | (1.11) | (1.73) | (1.06) | (0.18) | (0.18) |
| 9 Machinery[a] | 0.32 | 3.44 | 0.27 | 1.20 | 2.19* | 2.07* |
| | (2.88) | (3.53) | (2.93) | (3.86) | (0.20) | (0.42) |
| 10 Electrical equipm. | 0.74 | -3.05 | 0.90 | 2.73* | 1.66* | 2.28* |
| | (1.47) | (2.38) | (1.49) | (1.35) | (0.22) | (0.35) |
| 11 Transport equipm.[a] | 0.21 | -0.05 | 0.07 | -1.65 | 1.00 | 0.55 |
| | (3.53) | (3.31) | (3.59) | (3.91) | (0.72) | (1.11) |
| Substitutes or Complements: | | | | | | |
| Substitutes[b] | **11 (2)** | **8 (3)** | **11 (2)** | **10 (3)** | **11 (7)** | **11 (8)** |
| Complements[b] | **0 (0)** | **3 (0)** | **0 (0)** | **1 (0)** | **0 (0)** | **0 (0)** |

Notes: Swiss sectoral data. Asymptotic standard errors derived with the delta method. * Indicates significance at the 5% level. [a] Specification contains employment as a control variable. [b] Number of significant estimates in parentheses.

## 4.2 Estimates from Pooled OECD Data

The large standard errors found above make a clear assessment of the the degree of factor substitutability in Switzerland difficult. To obtain more precise estimates, we complement the Swiss data with data from further high-income OECD countries and pool these data in the regression. This specification again yields sectoral-specific substitution elasticities but it does imply equality of the substitution patterns across countries. We use EU KLEMS data for Belgium, Denmark, France, Germany, Great Britain, Japan, the Netherlands, Spain and Sweden, i.e., high-income OECD countries with economic attributes similar to Switzerland.

Table 6: Own-price elasticities, pooled across OECD countries

| | Industry | $CPE_{LL}$ | $CPE_{KK}$ | $CPE_{EE}$ |
|---|---|---|---|---|
| 1 | Food | -0.50* | -0.57* | -0.01 |
| | | (0.05) | (0.05) | (0.12) |
| 2 | Textiles[a] | -0.40* | -1.35* | 0.16 |
| | | (0.04) | (0.19) | (0.14) |
| 3 | Wood[a] | -0.49* | -0.41* | -0.11 |
| | | (0.06) | (0.15) | (0.18) |
| 4 | Pulp, paper | -0.37* | -0.64* | -0.63* |
| | | (0.04) | (0.09) | (0.07) |
| 5 | Chemicals | -0.65* | -0.41* | -0.87* |
| | | (0.08) | (0.06) | (0.14) |
| 6 | Rubber, plastics[a] | -0.40* | -0.60* | -0.25 |
| | | (0.05) | (0.11) | (0.14) |
| 7 | Other non-metals | -0.50* | -0.90* | -0.72* |
| | | (0.06) | (0.09) | (0.08) |
| 8 | Metals | -0.28* | -0.49* | -0.48* |
| | | (0.06) | (0.06) | (0.10) |
| 9 | Machinery | -0.36* | -0.51* | -0.73* |
| | | (0.06) | (0.10) | (0.14) |
| 10 | Electrical equipm. | -0.63* | -0.80* | -1.94* |
| | | (0.07) | (0.13) | (0.15) |
| 11 | Transport equipm. | -0.62* | -0.98* | -0.65* |
| | | (0.11) | (0.19) | (0.14) |
| Concavity Test: | | | | |
| Negative[b] | | **11 (11)** | **11 (11)** | **10 (7)** |
| Positive[b] | | **0 (0)** | **0 (0)** | **1 (0)** |

Notes: Pooled estimates with country fixed effects using data from Switzerland and 9 OECD countries. Asymptotic standard errors derived with the delta method. * Indicates significance at the 5% level. [a] Specification contains employment. [b] Number of significant estimates in parentheses.

Table 6 displays the resulting own-price elasticities. Note that standard errors are substantially smaller and most of the estimates are significantly smaller than zero. Generally, estimates of the pooled specification are similar in magnitude as the Swiss estimates, except in some cases, where the Swiss elasticities exhibit mostly high standard errors and insignificant results. All but three own-price elasticities in the pooled specification lie in the range between 0 and -1. All own-price elasticities are smaller than zero, indicating that concavity conditions are well met. Figure 4 in the Appendix depicts point estimates as well as confidence intervals graphically.

Table 7: Cross-price elasticities, pooled across OECD countries

| Industry | $CPE_{LE}$ | $CPE_{EL}$ | $CPE_{KE}$ | $CPE_{EK}$ | $CPE_{KL}$ | $CPE_{LK}$ |
|---|---|---|---|---|---|---|
| 1 Food | 0.03* | 0.39* | -0.04* | -0.99* | 0.11* | 0.18* |
| | (0.01) | (0.20) | (0.01) | (0.22) | (0.03) | (0.05) |
| 2 Textiles[a] | 0.03* | 0.54* | -0.11* | -1.10* | 0.44* | 0.22* |
| | (0.01) | (0.19) | (0.03) | (0.26) | (0.09) | (0.04) |
| 3 Wood[a] | 0.04* | 0.74* | -0.11* | -1.74* | 0.41* | 0.30* |
| | (0.01) | (0.31) | (0.03) | (0.42) | (0.08) | (0.06) |
| 4 Pulp, paper | 0.00 | 0.03 | 0.03* | 0.28* | 0.10* | 0.10* |
| | (0.01) | (0.08) | (0.01) | (0.14) | (0.05) | (0.05) |
| 5 Chemicals | -0.12* | -0.76* | 0.06* | 0.93* | 0.15* | 0.38* |
| | (0.03) | (0.18) | (0.02) | (0.26) | (0.03) | (0.08) |
| 6 Rubber, plastics[a] | 0.06* | 0.71* | -0.13* | -1.45* | 0.11* | 0.10* |
| | (0.02) | (0.20) | (0.03) | (0.32) | (0.06) | (0.05) |
| 7 Other non-metals | 0.14* | 0.66* | 0.04 | 0.21 | 0.38* | 0.46* |
| | (0.02) | (0.10) | (0.02) | (0.14) | (0.04) | (0.05) |
| 8 Metals | -0.04* | -0.35* | 0.03 | 0.26 | 0.02 | 0.02 |
| | (0.02) | (0.13) | (0.02) | (0.15) | (0.05) | (0.05) |
| 9 Machinery | -0.05* | -1.79* | 0.04* | 1.24* | 0.29* | 0.24* |
| | (0.01) | (0.24) | (0.01) | (0.28) | (0.06) | (0.05) |
| 10 Electrical equipm. | 0.01 | 0.12 | 0.15* | 3.65* | 0.15* | 0.17* |
| | (0.01) | (0.26) | (0.02) | (0.42) | (0.07) | (0.08) |
| 11 Transport equipm. | 0.05* | 0.83* | -0.05* | -0.95* | -0.25* | -0.29* |
| | (0.01) | (0.23) | (0.02) | (0.31) | (0.10) | (0.12) |
| Substitutes or Complements: | | | | | | |
| Substitutes[b] | **8 (6)** | **8 (6)** | **6 (4)** | **6 (4)** | **10 (9)** | **10 (9)** |
| Complements[b] | **3 (3)** | **3 (3)** | **5 (5)** | **5 (5)** | **1 (1)** | **1 (1)** |

Notes: Pooled estimates with country fixed effects using data from Switzerland and 9 OECD countries. Asymptotic standard errors derived with the delta method. * Indicates significance at the 5% level. [a] Specification contains employment. [b] Number of significant estimates in parentheses.

Table 7 displays the corresponding CPE. Compared to the Swiss estimates, the elasticities are often of smaller magnitude. As in the Swiss case, evidence on the substitutability is strong in the case of capital and labor. However, it is mixed in the case of labor and energy as well as capital and energy. More than 80% of all estimates are found to be significantly different from zero. Figure 5 in the Appendix depicts point estimates as well as confidence intervals graphically.

Table 8 displays the corresponding MES. It is found that all factors are estimated to be mainly MES substitutes, as is the case for Swiss manufacturing. Again, more than 80% of all estimates are found to be significantly different from zero. Figure 6 in the Appendix depicts point estimates as well as confidence intervals graphically.

To summarize, we find that all inputs tend to be MES substitutes using Swiss and pooled OECD data. Additionally, while labor and capital are predominantly CPE substitutes, labor and energy as well as capital and energy are found to be CPE complements in a number of cases under both specifications. Besides the expected difference in the accuracy of the estimates thanks to the larger sample size, the more moderate magnitude of the elasticities of the pooled specification is a further noteworthy observation. It may be due to the lesser importance of outliers as a result of the larger sample size or due to an "averaging-out" effect in our country sample.

Table 8: Morishima elasticities, pooled across OECD countries

| Industry | $MES_{LE}$ | $MES_{EL}$ | $MES_{KE}$ | $MES_{EK}$ | $MES_{KL}$ | $MES_{LK}$ |
|---|---|---|---|---|---|---|
| 1 Food | 0.04 | 0.89* | -0.03 | -0.42 | 0.61* | 0.75* |
| | (0.12) | (0.21) | (0.12) | (0.23) | (0.06) | (0.08) |
| 2 Textiles[a] | -0.13 | 0.94* | -0.27 | 0.25 | 0.84* | 1.56* |
| | (0.14) | (0.20) | (0.14) | (0.32) | (0.10) | (0.19) |
| 3 Wood[a] | 0.14 | 1.23* | -0.00 | -1.33* | 0.90* | 0.71* |
| | (0.18) | (0.32) | (0.18) | (0.45) | (0.11) | (0.16) |
| 4 Pulp, paper | 0.63* | 0.40* | 0.66* | 0.92* | 0.47* | 0.74* |
| | (0.07) | (0.09) | (0.07) | (0.17) | (0.07) | (0.10) |
| 5 Chemicals | 0.75* | -0.11 | 0.92* | 1.33* | 0.80* | 0.79* |
| | (0.15) | (0.20) | (0.14) | (0.27) | (0.08) | (0.10) |
| 6 Rubber, plastics[a] | 0.31* | 1.11* | 0.12 | -0.85* | 0.52* | 0.71* |
| | (0.14) | (0.20) | (0.14) | (0.34) | (0.08) | (0.13) |
| 7 Other non-metals | 0.87* | 1.16* | 0.76* | 1.11* | 0.88* | 1.36* |
| | (0.08) | (0.11) | (0.08) | (0.16) | (0.07) | (0.10) |
| 8 Metals | 0.44* | -0.07 | 0.51* | 0.75* | 0.30* | 0.51* |
| | (0.10) | (0.15) | (0.10) | (0.16) | (0.08) | (0.08) |
| 9 Machinery | 0.68* | -1.43* | 0.78* | 1.75* | 0.65* | 0.75* |
| | (0.14) | (0.25) | (0.14) | (0.30) | (0.08) | (0.11) |
| 10 Electrical equipm. | 1.94* | 0.75* | 2.08* | 4.45* | 0.78* | 0.98* |
| | (0.15) | (0.27) | (0.15) | (0.44) | (0.10) | (0.16) |
| 11 Transport equipm. | 0.70* | 1.45* | 0.61* | 0.02 | 0.37* | 0.68* |
| | (0.14) | (0.26) | (0.14) | (0.36) | (0.15) | (0.23) |
| Substitutes or Complements: | | | | | | |
| Substitutes[b] | **10 (8)** | **8 (8)** | **8 (7)** | **8 (6)** | **11 (11)** | **11 (11)** |
| Complements[b] | **1 (0)** | **3 (1)** | **3 (0)** | **3 (2)** | **0 (0)** | **0 (0)** |

Notes: Pooled estimates with country fixed effects using data from Switzerland and 9 OECD countries. Asymptotic standard errors derived with the delta method. * Indicates significance at the 5% level. [a] Specification contains employment. [b] Number of significant estimates in parentheses.

# 5 Discussion and Policy Relevance

We first compare our estimation results to findings from the existing literature in Section 5.1, emphasizing the complementarity versus substitutability debate. In Section 5.2, we assess how energy price increases affect the input mix of manufacturing industries using our estimation results. Finally, in Section 5.3 we provide policy implications of our results.

## 5.1 A Comparison with the Existing Literature

Table 9 lists a selection of recent contributions that utilize the translog specification to estimate elasticities of substitution.[11] Some of these studies estimate elasticities of aggregate manufacturing while others compute industry-specific elasticities. Three of the contributions use detailed firm-level data. Columns two to eight of the table summarize further characteristics of the estimation approach and the data employed. The last column reveals the resulting range

Table 9: Comparison of assumptions and results of translog studies

| Study | Struct. | Country | Industries | Period | Estimat. | Technology | Elasticities | Estimates |
|---|---|---|---|---|---|---|---|---|
| M-M[a] | KLEM | CHE | 11 | 1997-08 | SURE | factor-specific | CPE, MES | -5.2 to 5.9 |
| M-M[a] | KLEM | 9 C. | 11 | 1995-05 | SURE | factor-specific | CPE, MES | -1.9 to 4.4 |
| A-B[b] | KLE | DEN | micro pan. | 1993-97 | SURE | factor-specific | CPE | -1.5 to 2.6 |
| C[c] | KLE | GRE | aggregate | 1970-90 | IZEF* | Hicks-neutral | CPE | -1.3 to 1.0 |
| D-G[d] | KLEM | CAN | 2 (4-digit) | 1961-03 | SURE | factor-specific | CPE, MES | -0.4 to 0.8 |
| K-T[e] | KLE | USA | 3 | 2007 | SURE | Hicks-neutral | CPE, MES | -12.3 to 9.6 |
| M-V[f] | KLE | 3 C. | aggregate | 1980-96 | SURE | Hicks-neutral | AES, CPE | -0.4 to 0.4 |
| N-S[g] | KLEM | USA | micro data | 1991 | IZ3SLS** | factor-specific | CPE, MES | -3.8 to 3.2 |
| T-T[h] | KLE(M) | FIN | micro data | 2000-09 | SURE | factor-specific | AES, CPE | -1.4 to 1.6 |
| W-O[i] | KLE | GER | aggregate | 1976-94 | I3SE | factor-specific | CPE, MES | -1.3 to 1.6 |
| Y[j] | KLE | SWE | 9 (2-digit) | 1965-89 | SURE | factor-specific | AES, CPE | -0.5 to 1.0 |

Notes: Comparison of studies that estimate elasticities from a translog specification. * Nonlinear iterative Zellner estimation. ** Iterative Zellner three stage least squares. [a] We report economic elasticities for 11 sectoral aggregates for Switzerland. [b] We pool data over 9 OECD countries. [c] Arnberg and Bjorner (2007) distinguish between electricity and other energy; Cross-section estimates as well as fixed-effects estimates; Additionally they apply a linear logit specification. [d] Christopoulos (2000) distinguishes between crude oil, electricity and diesel (energy sources nested); estimates long-run elasticities from a dynamic structure; elasticities containing diesel are larger than the others. [e] Dissou and Ghazal (2010) Additionally to the translog cost function they apply a Symmetric Generalized McFadden (SGM) cost function. [f] Krishnapillai and Thompson (2012) use electricity instead of energy; aggregate 4-digit industries to three categories. [g] Medina and Vega-Cervera (2001) estimate elasticities of manufacturing for Italy, Portugal and Spain. [h] Nguyen and Streitwieser (1999) use a cross-section data set; estimate elasticities for small plants and large plants. [i] Tamminen and Tuomaala (2012) present own-price elasticities and MES for 71 sectors; factor material is nested; differentiate between electricity and other energy; include the production factor outside services. [j] Welsch and Ochsen (2005) show only elasticities containing the factor energy; differentiate between low-skilled and high-skilled labor; consider technological change and trade orientation. [k] Yi (2000) distinguishes between electricity and fuels; estimates a dynamic translog (TL) version and a general Leontief (GL) version.

[11] For a comparison of elasticity results of earlier studies we refer to Apostolakis (1990) and Koetse et al. (2008).

of the elasticity estimates' magnitude. We observe that the size of our estimates are largely in line with other contributions that provide industry-specific elasticities while estimates of aggregate manufacturing elasticities are usually of lower magnitude.

Regarding the substitutability versus complementarity debate, we first observe that the production factors capital and labor are predominantly CPE substitutes and MES substitutes using the Swiss as well as the OECD sample. This is in line with the existing literature which usually predicts substitutability between these two factors for manufacturing industries as well as for aggregate manufacturing.[12]

Second, energy and labor are also usually estimated to be CPE substitutes and MES substitutes at the aggregate level in the existing literature. However, empirical findings are less clear on an industry-specific level, indicating complementarity in a number of cases. Using Swiss data, we find that evidence on CPE substitutability is mixed while labor and energy are clearly MES substitutes. Using the OECD sample, there is slightly more evidence for the substitutability of these factors: In about two thirds of the industries, labor and energy are CPE substitutes while in all but one industries, these factors are MES substitutes.

Third, existing empirical evidence is especially mixed on the question of whether capital and energy are substitutes or complements. A comparison of empirical papers by Apostolakis (1990) shows that capital and energy tend to be complements in studies using time-series data, while in cross-section and pooled time-series studies capital and energy are estimated as being substitutes. Further differences in the modeling approach (KLEM vs. KLE), the level of sector aggregation, or the region and time period covered also influence the estimation results as is emphasized by Koetse et al. (2008). Using our OECD estimates, we find that capital and energy are CPE substitutes in about half and MES substitutes in about two thirds of the industries. Using Swiss data only and relying on the point estimates, the two factors are CPE substitutes in eight industries and complements in the remaining three and MES substitutes in all but one industry. The last result is in line with in-

[12]Capital and labor may also be complements in a few particular industries, as shown, for instance, in Dissou and Ghazal (2010).

sights from Thompson and Taylor (1995) who emphasize that using the MES, empirical evidence clearly suggests substitutability between those two factors.

## 5.2 Energy Price Increases and the Input Mix

The economic substitution elasticities estimated above describe the cost minimizing behavior of firms upon price changes holding output constant and allowing for the adjustment of all factor inputs simultaneously. In such a setting, complementarity of certain input factors is only possible in a production function environment with three or more production factors. In other words, if there exists complementarity between two factors and the price of one of them increases, then the remaining factors have to offset the reduction in the use of the first two factors to hold output constant.

There are two take-aways from these clarifications: First, all relevant production factors must be included into the estimation, since the exclusion of factors that are not homothetically separable in prices results in biased elasticity estimates (Jorgenson, 1986). Second, when interpreting the adjustment behavior of firms using these elasticities, it is mandatory to consider all factor adjustments simultaneously, since they are largely influenced by each other. In our view, this is often neglected in the existing literature, that mostly concentrates on the interplay between two selected factors (usually energy and capital). We therefore provide an interpretation of our estimates considering all factor adjustments in the following. In the light of recent developments especially in Switzerland but also in other OECD economies, we focus in our analysis on the implications of energy price increases for the input mix of manufacturing industries.

Table 10 displays the industry-specific effects of a relative energy price increase on the factor expenses and on total production costs using the pooled OECD data.[13] We first note that within the translog framework, total production costs adjustments are directly proportional to the factor share of the factor experiencing the price change: The last column indicates, for example, that a 1% increase of the energy price yields a total cost increase of 0.018% given the sector's initial (average) cost of 1.8% (see to Table 2).

Table 10: Effects of an energy price increase on factor expenses

| Industry | | $\Delta p_E x_E$ | $\Delta p_L x_L$ | $\Delta p_K x_K$ | $\Delta p_M x_M$ | $\Delta TC$ |
|---|---|---|---|---|---|---|
| 1 | Food | 0.990 | 0.028 | -0.042 | 0.015 | *0.018* |
| 2 | Textiles | 1.162 | 0.027 | -0.112 | 0.015 | *0.020* |
| 3 | Wood | 0.891 | 0.035 | -0.113 | 0.036 | *0.028* |
| 4 | Pulp, paper | 0.365 | 0.003 | 0.029 | 0.025 | *0.026* |
| 5 | Chemicals | 0.125 | -0.116 | 0.056 | 0.036 | *0.052* |
| 6 | Rubber, plastics | 0.744 | 0.057 | -0.130 | 0.045 | *0.031* |
| 7 | Other non-metals | 0.270 | 0.144 | 0.038 | -0.025 | *0.061* |
| 8 | Metals | 0.514 | -0.044 | 0.033 | 0.040 | *0.036* |
| 9 | Machinery | 0.258 | -0.054 | 0.045 | 0.022 | *0.012* |
| 10 | Electrical equipm. | -0.958 | 0.005 | 0.145 | -0.045 | *0.011* |
| 11 | Transport equipm. | 0.340 | 0.049 | -0.048 | 0.013 | *0.011* |
| Weighted avg. | | 0.184 | -0.013 | 0.030 | 0.011 | *0.025* |
| No. positive/negative | | 10/1 | 8/3 | 6/5 | 9/2 | 11/0 |

Notes: Estimation results from the pooled OECD sample; Values denote percentage changes (%); $s_E$ is the share of energy expenses; $\Delta p_i x_i$ is the relative change of input $i$'s expenses; $\Delta TC$ is the relative change of total production costs.

More interesting are hence the effects on the expenses of the various input factors displayed in the preceding four columns. On average, energy expenses increase 0.18% upon an energy price rise of 1.00% due to the inelastic nature of the own-price elasticities in all but one industry using the OECD data. Mainly driven by the large negative effect in the chemical industry, labor expenses decrease by 0.01%, implying that labor and energy are CPE complements on average. In contrast, capital and material expenses increase by 0.03% and 0.01% upon an energy price rise, indicating substitutability versus energy on average. Since factor prices of labor, material and capital are assumed to be constant, changes in the expenses of these factors are equivalent to changes in the factor use. Hence, in some industries like the chemical industry, the change in the use of input factors can be quite substantial with a 0.12% decrease of labor use, a 0.06% increase of capital use and a 0.04% increase of material use upon a 1.00% increase in the energy price despite the limited energy intensity of production. As the last row of the table shows, substantial differences between sectors are observable.

[13] We want to emphasize that this is a partial equilibrium analysis, which does neither take into account general equilibrium effects on prices of other inputs nor demand side effects.

In Table 13 in the Appendix, the same results are shown using Swiss data only. Energy expenses decrease on average due to the elastic point estimates of the energy own-price elasticities. However, this result must be put into perspective since standard errors are large and it is not possible to show that elasticities are indeed significantly smaller than -1. Reaction of the other factor expenses are on average similar to results of the OECD sample; also observing reductions in labor expenses and increases in capital and material use.

Table 11 displays the effects of an energy price increase on the ratios of the physical use (left panel) as well as the costs (rights panel) of energy relative to the other factors, hence using the MES as determinants. The first three columns indicate that the ratio of physical energy use relative to the use of the three other factors of production decreases on average, where the decrease is strongest for the energy-capital ratio. The last three columns of the table imply that despite the relative decrease in the employed quantity of energy, energy cost ratios rise on average relative to all other input factors, an outcome due to the inelastic nature of the own-price elasticities of the factor energy.

Table 11: Effects of an energy price increase on factor use and cost ratios

| | | Use ratio of Energy to: | | | Cost ratio of Energy to: | | |
|---|---|---|---|---|---|---|---|
| | Industry | Labor | Capital | Material | Labor | Capital | Material |
| 1 | Food | -0.04 | 0.03 | -0.03 | 0.96 | 1.03 | 0.97 |
| 2 | Textiles | 0.13 | 0.27 | 0.15 | 1.13 | 1.27 | 1.15 |
| 3 | Wood | -0.14 | 0.00 | -0.14 | 0.86 | 1.00 | 0.86 |
| 4 | Pulp, paper | -0.63 | -0.66 | -0.65 | 0.37 | 0.34 | 0.35 |
| 5 | Chemicals | -0.75 | -0.92 | -0.90 | 0.25 | 0.08 | 0.10 |
| 6 | Rubber, plastics | -0.31 | -0.12 | -0.30 | 0.69 | 0.88 | 0.70 |
| 7 | Other non-metals | -0.87 | -0.76 | -0.70 | 0.13 | 0.24 | 0.30 |
| 8 | Metals | -0.44 | -0.51 | -0.52 | 0.56 | 0.49 | 0.48 |
| 9 | Machinery | -0.68 | -0.78 | -0.76 | 0.32 | 0.22 | 0.24 |
| 10 | Electrical equipm. | -1.94 | -2.08 | -1.89 | -0.94 | -1.08 | -0.89 |
| 11 | Transport equipm. | -0.70 | -0.61 | -0.67 | 0.30 | 0.39 | 0.33 |
| Weighted avg. | | -0.80 | -0.84 | -0.82 | 0.20 | 0.16 | 0.18 |
| No. positive/negative | | 1/10 | 3/10 | 1/10 | 10/1 | 10/1 | 10/1 |

Notes: Estimation results from the pooled OECD sample; Values denote percentage changes (%); values of column two to four are $-MES_{iE}$ and values of column five to seven are $1 - MES_{iE}$, for $i = L, K, M$.

The increase is strongest regarding labor costs, a follow-up of the CPE complementarity between energy and labor. Tables 10 and 11, hence, nicely demonstrate the relationship between MES and CPE while still emphasizing the different interpretation of these measures.

In Table 14 in the Appendix, we show analogous results using Swiss data. Again, the general patterns are similar to the OECD sample: Ratios of physical energy use decrease relative to the other input factors, the strongest effect being on the energy-capital use ratio. In contrast to the OECD sample, however, cost ratios of energy decrease, a consequence of the elastic but insignificant own-price elasticities of energy using the Swiss data set. In analogy to the OECD sample, the reduction of the energy-labor cost ratio is weakest owing to the CPE complementarity of these factors.

## 5.3 Policy Implications

The implementation of policies to mitigate climate change as well as to achieve the nuclear phaseout in Switzerland will most likely change absolute and relative prices of different energy sources as energy prices are expected to rise substantially.[14] These changes will alter the factor content of industrial production and increase production costs in Switzerland. If these policies are implemented unilaterally, the policy induced price changes only affect the Swiss economy. Manufacturing industries are, however, often exposed to international competition, which leads to low pass-through possibilities of changes in production costs. Furthermore, unilateral implementation of mitigation or abatement policies involve counterproductive effects such as carbon leakage.[15] In summary, unilateral energy policies are prone to have unwanted reverse effects on the Swiss but also the world economy.

[14] For example, the IEA (2012) notes that several aspects of Switzerland's energy policy suggest an increased likelihood of higher energy prices: GHG emission reduction targets, policies aimed at the nuclear phaseout, investments in electricity grids and capacity, and convergence with price levels in surrounding countries. Furthermore, scenarios of the Swiss government expect substantial price increases of more than 50% for most energy carriers, see Swiss Federal Office of Energy (2011).

[15] Carbon leakage has been investigated by e.g., Felder and Rutherford (1993) and Juergens et al. (2013).

For these reasons, it is well recognized that climate and energy policies are best implemented in a multilateral framework to mitigate detrimental effects on the environment and, at the same time, prevent negative repercussions on the domestic (energy-intensive) production. Table 12 shows the average energy shares of Swiss industries along with an openness indicator and the revenue share in total Swiss manufacturing as a static measure for the industries' importance. Using the indicators in the first two columns, the textile, the chemical and the metal industry seem to be most exposed by a unilateral policy, the latter two industries accounting for two thirds of Swiss manufacturing revenues.

Table 12: Exposure of Swiss manufacturing industries

| Industry | | Energy share | Openness | Revenue share |
|---|---|---|---|---|
| 1 | Food | 1.4% | 45% | 10.1% |
| **2** | **Textiles** | **1.8%** | **345%** | **1.3%** |
| 3 | Wood | 1.3% | 30% | 3.1% |
| 4 | Pulp, paper | 3.5% | 70% | 5.4% |
| **5** | **Chemicals** | **1.5%** | **157%** | **22.9%** |
| 6 | Rubber, plastics | 1.4% | 118% | 3.0% |
| 7 | Other non-metals | 4.5% | 72% | 2.2% |
| **8** | **Metals** | **2.3%** | **123%** | **10.0%** |
| 9 | Machinery | 0.6% | 133% | 14.2% |
| 10 | Electrical equipm. | 0.8% | 109% | 25.1% |
| 11 | Transport equipm. | 1.4% | 323% | 2.7% |
| Weighted avg. (sum) | | 1.4% | 122% | (100%) |

Notes: The table displays average energy shares from 1997 to 2008. Openness is defined as the value of imports and exports relative to the total production value. The revenue shares are relative to aggregate manufacturing.

In this context, our analysis shows that the change of the input mix can be quite substantial in selected industries upon an energy price increase, even in the short run and given that production remains at a constant level: On average, energy and labor inputs will decrease relative to capital and material inputs upon energy price increases. The results seem intuitive as energy savings may result from investments into new technologies but also from out-sourcing the production of energy-intensive intermediates. Results also indicate the possibility that the labor force is partially displaced even at constant production levels. Given that energy prices are expected to increase by 50% or more in the long-run and the possibility of unwanted negative repercussions due to unilateral implementation, effects could be expected to be more severe although the short-run elasticities estimated in our contribution cannot answer such questions conclusively.

Nonetheless, our results in combination with the high exposure to international competition of certain industries give evidence to the notion that unilateral climate and energy policies are therefore accompanied by a great deal of uncertainty and substantial risks. It is not by chance that the IEA recommends Switzerland to align its energy policies with its major trade partners (IEA, 2012): "[Switzerland should] pursue closer integration with European energy markets and closest possible alignment of its energy policies with those of the European Union". If extensive climate and energy policies are adapted without international coordination, our result that reveals substantial changes in the input mix even in the short run hint at the possibility that reverse effects that are clearly beyond the scope of our analysis may be triggered. For example, alterations of factor prices of relatively immobile factors like labor, a certain degree of deindustrialization or even negative spill-overs on the environment. It remains to note that alternatives to multilateral commitment exist, e.g. supplementary measures such as border tax adjustments or tax exempts for energy intensive industries.[16] Both possibilities have inherent drawbacks. While the border tax adjustments are very difficult to implement from a technical point of view, tax exempts undermine the original purpose of energy policy measures.

# 6 Concluding Remarks

We estimate sectoral substitution elasticities for manufacturing industries in Switzerland and additionally present such elasticities from a sample of high-income OECD countries with economic attributes similar to Switzerland. In

[16]See for instance, Kuik and Hofkes (2010) and van Asselt and Brewer (2010) on border tax adjustments.

our analysis, we illustrate how energy price increases affect the input mix of manufacturing industries in the short run. We consider the relative change of factor expenses and factor ratios as well as the increase of total production costs upon an increase in relative energy prices. Finally, we discuss potential challenges that have to be considered when implementing future policy measures that increase relative energy prices.

Our empirical results suggest that a relative price increase of energy substantially affects the input mix of manufacturing firms in the short run. We show that the decrease in the use of energy is in many industries compensated by increasing expenses on labor, capital and material inputs. However, weighting industries according to their size results in a manufacturing-wide decrease of labor expenses upon an energy price increase, implying energy-labor complementarity. Our results give important insights about the flexibility of production in the short run and complement results of computational general equilibrium models that rather place emphasis on the middle- and long-run impact of policy measures.

It remains to emphasize that in the long run, even rigid factors can be adjusted as firms can adapt new technologies or innovate production. Better knowledge of these adjustment processes are of great interest for policy makers in the field of energy and environmental economics. Effects of technological innovations and the adaptation of new technologies are thus important topics of future research.

# Bibliography


Anderson, R. G. and J. G. Thursby (1986). Confidence intervals for elasticity estimators in translog models. *The Review of Economics and Statistics 68*(4), 647–656.

Apostolakis, B. E. (1990). Energy–capital substitutability/complementarity: The dichotomy. *Energy Economics 12*(1), 48–58.

Arnberg, S. and T. B. Bjorner (2007). Substitution between energy, capital and labor within industrial companies: A micro panel data analysis. *Resource and Energy Economics 29 (2)*, 122–136.

Berndt, E. R. and L. R. Christensen (1973). The translog function and the substitution of equipment, structures, and labor in U.S. manufacturing 1929-68. *Journal of Econometrics 1*(1), 81–113.

Berndt, E. R. and D. O. Wood (1979). Engineering and econometric interpretations of energy-capital complementarity. *American Economic Review 69*(3), 342–354.

Berndt, E. R. and D. O. Wood (1981). Engineering and econometric interpretations of energy-capital complementarity: Reply and further results. *American Economic Review 71*(5), 1105–1110.

Blackorby, C. and R. R. Russell (1989). Will the real elasticity of substitution please stand up? (A comparison of the Allen/Uzawa and Morishima elasticities). *American Economic Review 79*(4), 882–888.

Christensen, L. R., D. W. Jorgenson, and L. J. Lau (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics 55*(1), 28–45.

Christopoulos, D. K. (2000). The demand for energy in Greek manufacturing. *Energy Economics 22*(5), 569–586.

Dissou, Y. and R. Ghazal (2010). Energy substitutability in Canadian manufacturing econometric estimation with bootstrap confidence intervals. *The Energy Journal 31*(1), 121–148.

Felder, S. and T. F. Rutherford (1993). Unilateral CO2 reductions and carbon leakage: The consequences of international trade in oil and basic materials. *Journal of Environmental Economics and Management 25*(2), 162–176.

Frondel, M. (2011). Modeling energy and non-energy substitution: A survey of elasticities. Ruhr Economic Papers 0256, Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Ruhr-Universität Bochum, Universität Dortmund, Universität Duisburg-Essen.

Greene, W. H. (2000). *Econometric Analysis*. Prentice Hall.

Griffin, J. M. (1981). Engineering and econometric interpretations of energy-capital complementarity: Comment. *American Economic Review 71*(5), 1100–1104.

Griffin, J. M. and P. R. Gregory (1976). An intercountry translog model of energy substitution responses. *American Economic Review 66*(5), 845–857.

IEA (2012). *Energy Policies of IEA Countries: Switzerland 2012*. International Energy Agency (IEA). OECD Publishing.

Jorgenson, D. W. (1986). Econometric methods for modeling producer behavior. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 3 of *Handbook of Econometrics*, Chapter 31, pp. 1841–1915. Elsevier.

Juergens, I., J. Barreiro-Hurlé, and A. Vasa (2013). Identifying carbon leakage sectors in the EU ETS and implications of results. *Climate Policy 13*(1), 89–109.

Koetse, M. J., H. L. de Groot, and R. J. Florax (2008). Capital-energy substitution and shifts in factor demand: A meta-analysis. *Energy Economics 30*(5), 2236–2251.

Krishnapillai, S. and H. Thompson (2012). Cross section translog production and elasticity of substitution in U.S. manufacturing industry. *International Journal of Energy Economics and Policy 2*(2), 50–54.

Kuik, O. and M. Hofkes (2010). Border adjustment for European emissions trading: Competitiveness and carbon leakage. *Energy Policy 38*(4), 1741–1748.

Medina, J. and J. A. Vega-Cervera (2001). Energy and the non-energy inputs substitution: Evidence for Italy, Portugal and Spain. *Applied Energy 68*(2), 203–214.

Mohler, L. and D. Müller (2011). Energy-related data for Swiss manufacturing sectors. EWG Publication No. 290342, SFOE, Bern.

Mohler, L. and D. Müller (2012). Substitution elasticities in swiss manufacturing. EWG Publication No. 290707, SFOE, Bern.

Nguyen, S. V. and M. L. Streitwieser (1999). Factor substitution in U.S. manufacturing: Does plant size matter? *Small Business Economics 12*(1), 41–57.

O'Mahony, M. and M. P. Timmer (2009). Output, input and productivity measures at the industry level: The EU KLEMS database. *Economic Journal 119*(538), 374–403.

Swiss Federal Office of Energy (2011). Grundlagen für die Energiestrategie des Bundesrates. Bern.

Tamminen, S. and E. Tuomaala (2012). Variation in price and substitution elasticities between sectors? A microdata analysis. Working Papers 34, Government Institute for Economic Research Finland (VATT).

Thompson, P. and T. G. Taylor (1995). The capital-energy substitutability debate: A new look. *The Review of Economics and Statistics 77*(3), 565–569.

Timmer, M. P., M. O'Mahony, and B. van Ark (2007). EU KLEMS growth and productivity accounts: An overview. *International Productivity Monitor 14*, 71–85.

van Asselt, H. and T. Brewer (2010). Addressing competitiveness and leakage concerns in climate policy: An analysis of border adjustment measures in the US and the EU. *Energy Policy 38*(1), 42–51.

Welsch, H. and C. Ochsen (2005). The determinants of aggregate energy use in West Germany: Factor substitution, technological change, and trade. *Energy Economics 27*(1), 93–111.

Yi, F. (2000). Dynamic energy-demand models: A comparison. *Energy Economics 22*(2), 285–297.

# A Appendix

## A.1 Variances of the Elasticities

We derive the standard errors of the elasticities by using the Delta method (see e.g., Greene (2000)). The variance of the CPE is:

$$\text{var}(\text{CPE}_{ij}) = \left(\frac{1}{s_i}\right)^2 \text{var}(\hat{\beta}_{ij}). \tag{5}$$

The expression determining the variance of own-price elasticities can be derived analogously and is omitted here. The variance of the MES is given by

$$\begin{aligned} \text{var}(\text{MES}_{ij}) &= \text{var}(\text{CPE}_{ij}) + \text{var}(\text{CPE}_{jj}) - 2\text{cov}(\text{CPE}_{ij}, \text{CPE}_{jj}) \\ &= \left(\frac{1}{s_i}\right)^2 \text{var}(\hat{\beta}_{ij}) + \left(\frac{1}{s_j}\right)^2 \text{var}(\hat{\beta}_{jj}) - 2\left(\frac{1}{s_i s_j}\right) \text{cov}(\hat{\beta}_{ij}, \hat{\beta}_{jj}). \end{aligned} \tag{6}$$

## A.2 Data

Data for the OECD countries (Belgium, Denmark, France, Germany, Great Britain, Japan, the Netherlands, Spain and Sweden) is taken from the EU KLEMS database.[17] This database provides sectoral data on prices and quantities of different input and output variables for about 30 countries. Data coverage varies substantially across countries, and we selected the countries with comparable coverage to the Swiss data set (see below). We use data from 1995 to 2005 with the exception of France (1998 to 2005) and Great Britain (1996 to 2005). A description of this database is presented in Timmer et al. (2007) and O'Mahony and Timmer (2009).

Most data used to estimate the Swiss elasticities has been assembled during a precursory project, see Mohler and Müller (2011). We use data from 1997 to 2008 for the analysis of Swiss manufacturing industries. We use sales by industry as surveyed by the Swiss Federal Statistical Office (SFSO) in the Produktions- und Wertschöpfungsstatistik as our output variable. Material and labor expenditures stem from the same survey. Energy expenditures are calculated by using the survey EVID from the Swiss Federal Office of Energy

[17] The EU KLEMS database is readily available for download at *http://www.euklems.net.*

(SFOE), which comprises the physical quantities of different energy sources used by manufacturing industries. Unfortunately, this data set is not available at the 2-digit NOGA level. Hence, we aggregate the relevant NOGA 2-digit industries into 11 manufacturing industries as laid out in Table 1. We then calculate the energy expenditures of each industry by using energy prices published by the IEA and the SFOE and the physical quantities from the EVID database. Capital expenditures are calculated as a residual by taking the difference of sales, material, labor and energy expenditures. As for the price indices used in the analysis, we employ output and material price indices available from the OECD. We use a wage index published by the SFSO. Sectoral energy price indices are calculated by using the above-mentioned data on physical energy use and the energy prices using a Laspeyres price index approach. Capital price indices are not available for different manufacturing industries in Switzerland. We use an investment/capital goods import price index published by the Swiss Federal Customs Administration (EZV) to proxy capital price changes, as we are mainly interested in the price evolution of physical capital.

## A.3 Energy Price Effects, Swiss Sample

Table 13: Effects of an energy price increase on factor expenses, Switzerland

| Industry | | $\Delta p_E x_E$ | $\Delta p_L x_L$ | $\Delta p_K x_K$ | $\Delta p_M x_M$ | $\Delta TC$ |
|---|---|---|---|---|---|---|
| 1 | Food | -0.157 | -0.200 | 0.098 | 0.053 | *0.014* |
| 2 | Textiles | -1.253 | -0.009 | 0.198 | 0.011 | *0.018* |
| 3 | Wood | -2.210 | 0.144 | 0.168 | -0.091 | *0.013* |
| 4 | Pulp, paper | -0.212 | 0.006 | 0.164 | -0.003 | *0.034* |
| 5 | Chemicals | -0.436 | -0.084 | 0.046 | 0.041 | *0.015* |
| 6 | Rubber, plastics | -0.380 | 0.113 | -0.013 | -0.021 | *0.014* |
| 7 | Other non-metals | -0.306 | 0.207 | 0.105 | -0.075 | *0.044* |
| 8 | Metals | 0.120 | 0.090 | 0.025 | -0.040 | *0.023* |
| 9 | Machinery | 0.724 | 0.047 | -0.003 | -0.022 | *0.006* |
| 10 | Electrical equipm. | 0.129 | -0.117 | 0.035 | 0.060 | *0.008* |
| 11 | Transport equipm. | 0.707 | -0.083 | -0.220 | 0.122 | *0.014* |
| Weighted avg. | | -0.064 | -0.043 | 0.044 | 0.021 | *0.014* |
| No. positive/negative | | 4/7 | 6/5 | 8/3 | 5/6 | 11/0 |

Notes: Estimation results from the Swiss sample; Values denote percentage changes (%); $s_E$ is the share of energy expenses; $\Delta p_i x_i$ is the relative change of input $i$'s expenses; $\Delta\ TC$ is the relative change of total production costs.

Table 14: Effects of an energy price increase on factor use and cost ratios, Switzerland

| Industry | | Use ratio of Energy to: Labor | Capital | Material | Cost ratio of Energy to: Labor | Capital | Material |
|---|---|---|---|---|---|---|---|
| 1 | Food | -0.95 | -1.24 | -1.20 | 0.05 | -0.24 | -0.20 |
| 2 | Textiles | -2.22 | -2.43 | -2.24 | -1.22 | -1.43 | -1.24 |
| 3 | Wood | -3.32 | -3.35 | -3.09 | -2.32 | -2.35 | -2.09 |
| 4 | Pulp, paper | -1.21 | -1.36 | -1.20 | -0.21 | -0.36 | -0.20 |
| 5 | Chemicals | -1.34 | -1.47 | -1.46 | -0.34 | -0.47 | -0.46 |
| 6 | Rubber, plastics | -1.48 | -1.35 | -1.35 | -0.48 | -0.35 | -0.35 |
| 7 | Other non-metals | -1.50 | -1.40 | -1.22 | -0.50 | -0.40 | -0.22 |
| 8 | Metals | -0.96 | -0.90 | -0.83 | 0.04 | 0.10 | 0.17 |
| 9 | Machinery | -0.32 | -0.27 | -0.25 | 0.68 | 0.73 | 0.75 |
| 10 | Electrical equipm. | -0.74 | -0.90 | -0.92 | 0.26 | 0.10 | 0.08 |
| 11 | Transport equipm. | -0.21 | -0.07 | -0.41 | 0.79 | 0.93 | 0.59 |
| Weighted avg. | | -1.01 | -1.10 | -1.07 | -0.01 | -0.10 | -0.07 |
| No. positive/negative | | 0/11 | 0/11 | 0/11 | 5/6 | 4/7 | 4/7 |

Notes: Estimation results from the Swiss sample; Values denote percentage changes (%); values of column two to four are $-MES_{iE}$ and values of column five to seven are $1 - MES_{iE}$, for $i = L, K, M$.

## A.4 Key Estimation Statistics

Table 15: Key statistics, Swiss sample

| Industry | Equation | Observations | RMSE | R-squared | p-value |
|---|---|---|---|---|---|
| 1 | Capital | 11 | 0.004 | 0.46 | 0.005 |
| | Labor | 11 | 0.002 | 0.85 | 0.000 |
| | Energy | 11 | 0.001 | 0.29 | 0.631 |
| 2 | Capital | 11 | 0.009 | 0.18 | 0.391 |
| | Labor | 11 | 0.006 | 0.73 | 0.000 |
| | Energy | 11 | 0.001 | 0.71 | 0.077 |
| 3 | Capital | 11 | 0.008 | 0.45 | 0.190 |
| | Labor | 11 | 0.014 | 0.63 | 0.000 |
| | Energy | 11 | 0.000 | 0.79 | 0.863 |
| 4 | Capital | 11 | 0.005 | 0.65 | 0.000 |
| | Labor | 11 | 0.002 | 0.98 | 0.000 |
| | Energy | 11 | 0.000 | 0.89 | 0.000 |
| 5 | Capital | 11 | 0.008 | 0.92 | 0.000 |
| | Labor | 11 | 0.003 | 0.93 | 0.000 |
| | Energy | 11 | 0.000 | 0.96 | 0.000 |
| 6 | Capital | 11 | 0.004 | 0.35 | 0.000 |
| | Labor | 11 | 0.003 | 0.96 | 0.000 |
| | Energy | 11 | 0.001 | 0.40 | 0.070 |
| 7 | Capital | 11 | 0.007 | 0.75 | 0.000 |
| | Labor | 11 | 0.005 | 0.94 | 0.000 |
| | Energy | 11 | 0.002 | 0.67 | 0.022 |
| 8 | Capital | 11 | 0.004 | 0.79 | 0.000 |
| | Labor | 11 | 0.004 | 0.95 | 0.000 |
| | Energy | 11 | 0.001 | 0.93 | 0.000 |
| 9 | Capital | 11 | 0.006 | 0.59 | 0.000 |
| | Labor | 11 | 0.002 | 0.99 | 0.000 |
| | Energy | 11 | 0.000 | 0.43 | 0.477 |
| 10 | Capital | 11 | 0.007 | 0.88 | 0.000 |
| | Labor | 11 | 0.001 | 1.00 | 0.000 |
| | Energy | 11 | 0.000 | 0.96 | 0.000 |
| 11 | Capital | 11 | 0.010 | 0.41 | 0.142 |
| | Labor | 11 | 0.008 | 0.89 | 0.000 |
| | Energy | 11 | 0.001 | 0.61 | 0.501 |

Notes: For each sector we estimated a SUR consisting of three share equations. We list the number of observations, the root mean square error, the r-squared and the p-value for each equation.

Table 16: Key statistics, pooled OECD sample

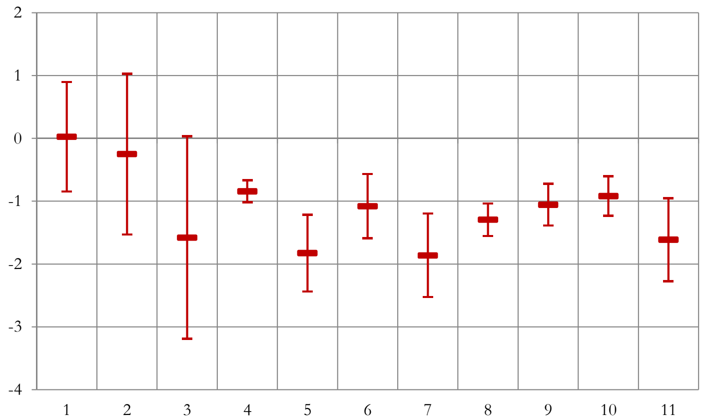
| Industries | Equation | Observations | RMSE | R-squared | p-value |
|---|---|---|---|---|---|
| 1 | Capital | 87 | 0.010 | 0.94 | 0.000 |
| | Labor | 87 | 0.004 | 0.98 | 0.000 |
| | Energy | 87 | 0.002 | 0.82 | 0.000 |
| 2 | Capital | 87 | 0.014 | 0.93 | 0.000 |
| | Labor | 87 | 0.010 | 0.96 | 0.000 |
| | Energy | 87 | 0.002 | 0.82 | 0.000 |
| 3 | Capital | 87 | 0.018 | 0.88 | 0.000 |
| | Labor | 87 | 0.013 | 0.94 | 0.000 |
| | Energy | 87 | 0.004 | 0.96 | 0.000 |
| 4 | Capital | 87 | 0.017 | 0.90 | 0.000 |
| | Labor | 87 | 0.011 | 0.96 | 0.000 |
| | Energy | 87 | 0.002 | 0.93 | 0.000 |
| 5 | Capital | 87 | 0.015 | 0.95 | 0.000 |
| | Labor | 87 | 0.008 | 0.96 | 0.000 |
| | Energy | 87 | 0.007 | 0.94 | 0.000 |
| 6 | Capital | 87 | 0.016 | 0.86 | 0.000 |
| | Labor | 87 | 0.010 | 0.92 | 0.000 |
| | Energy | 87 | 0.004 | 0.96 | 0.000 |
| 7 | Capital | 87 | 0.013 | 0.95 | 0.000 |
| | Labor | 87 | 0.008 | 0.95 | 0.000 |
| | Energy | 87 | 0.006 | 0.78 | 0.000 |
| 8 | Capital | 87 | 0.011 | 0.94 | 0.000 |
| | Labor | 87 | 0.010 | 0.97 | 0.000 |
| | Energy | 87 | 0.006 | 0.78 | 0.000 |
| 9 | Capital | 87 | 0.014 | 0.92 | 0.000 |
| | Labor | 87 | 0.010 | 0.93 | 0.000 |
| | Energy | 87 | 0.002 | 0.89 | 0.000 |
| 10 | Capital | 87 | 0.022 | 0.91 | 0.000 |
| | Labor | 87 | 0.010 | 0.93 | 0.000 |
| | Energy | 87 | 0.002 | 0.72 | 0.000 |
| 11 | Capital | 87 | 0.013 | 0.93 | 0.000 |
| | Labor | 87 | 0.010 | 0.97 | 0.000 |
| | Energy | 87 | 0.001 | 0.86 | 0.000 |

Notes: Data pooled across countries including country fixed effects. For each sector we estimated a SUR consisting of three share equations. We list the number of observations, the root mean square error, the r-squared and the p-value for each equation.

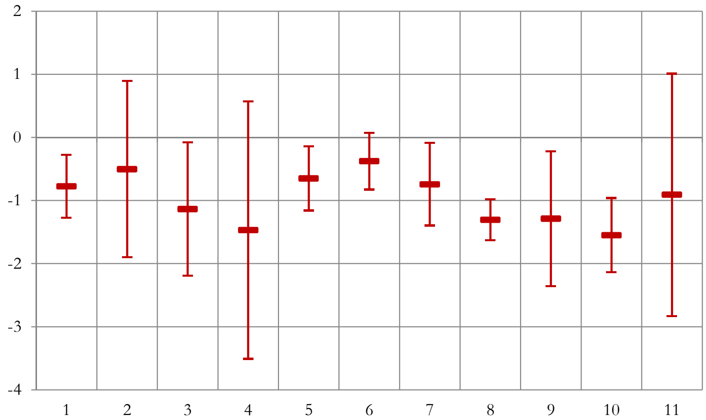
## A.5 Estimation Results: Switzerland

Figure 1: Own-price elasticities, Swiss manufacturing industries

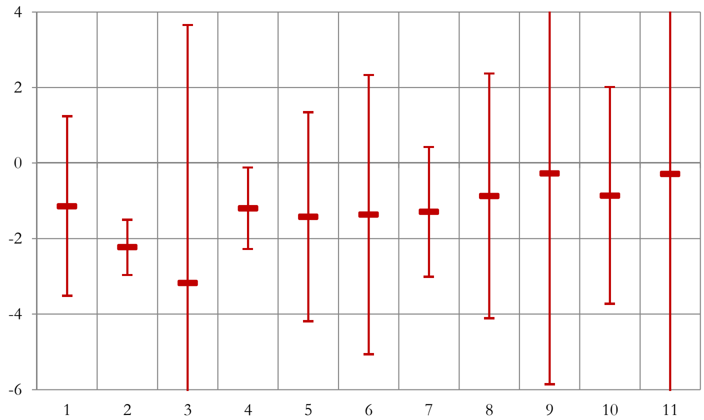
(a) $\text{CPE}_{LL}$




(b) $\text{CPE}_{KK}$



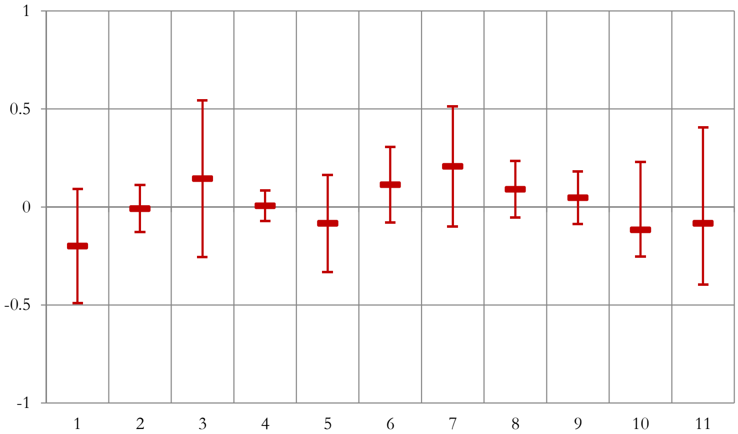

(c) $\text{CPE}_{EE}$



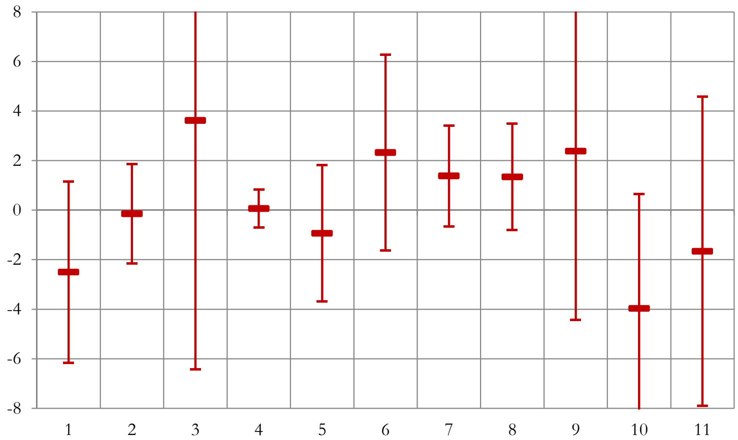

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries in Switzerland. We use the translog cost function as functional form.

Figure 2: Cross-price elasticities, Swiss manufacturing industries

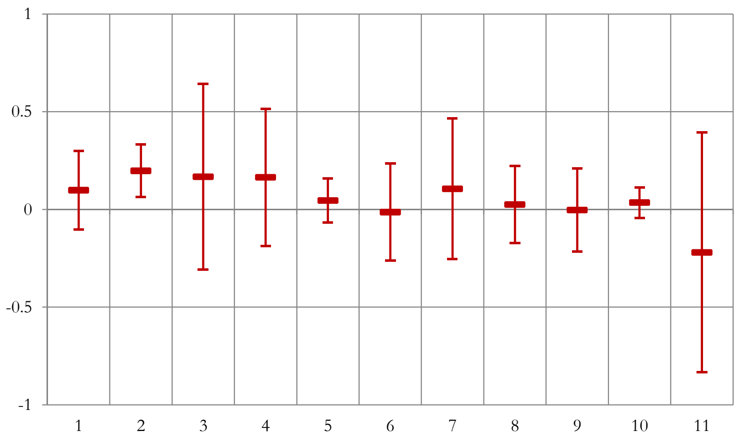
(a) $\text{CPE}_{LE}$




(b) $\text{CPE}_{EL}$



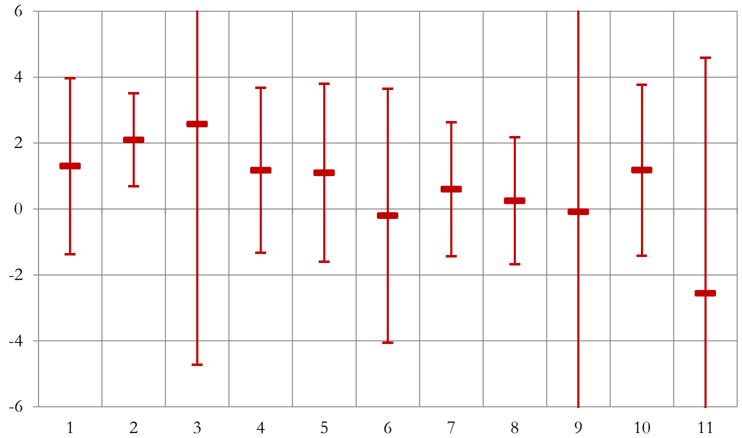

(c) $\text{CPE}_{KE}$



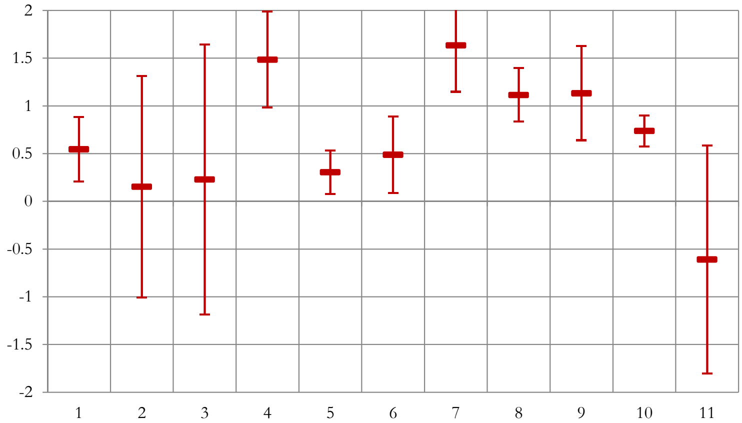

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries in Switzerland. We use the translog cost function as functional form.

Figure 2: Cross-price elasticities, Swiss manufacturing industries (cont.)

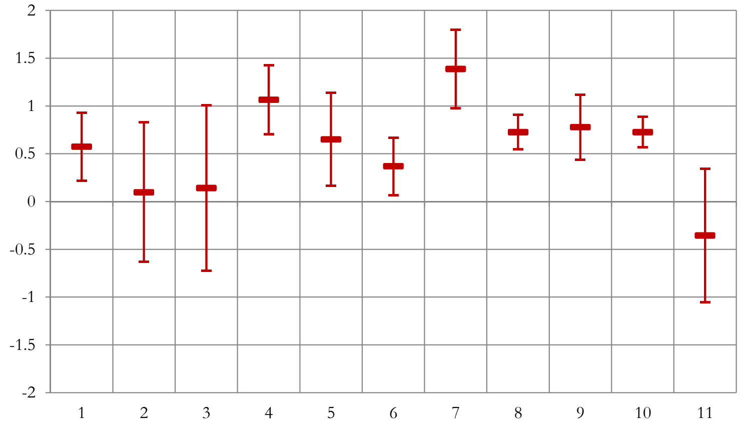
(d) $\text{CPE}_{EK}$




(e) $\text{CPE}_{KL}$



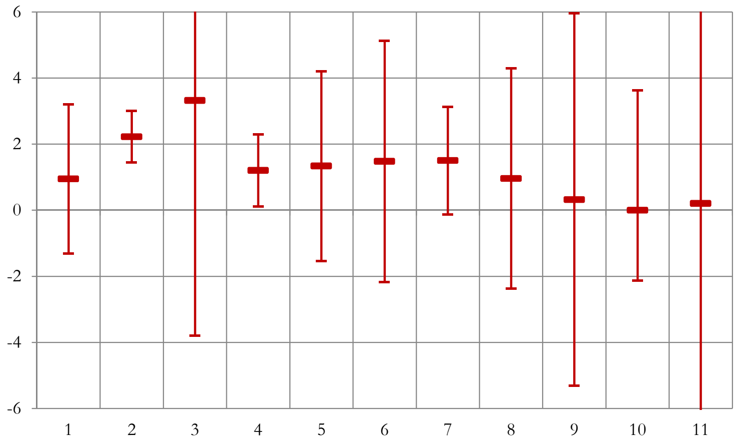

(f) $\text{CPE}_{LK}$



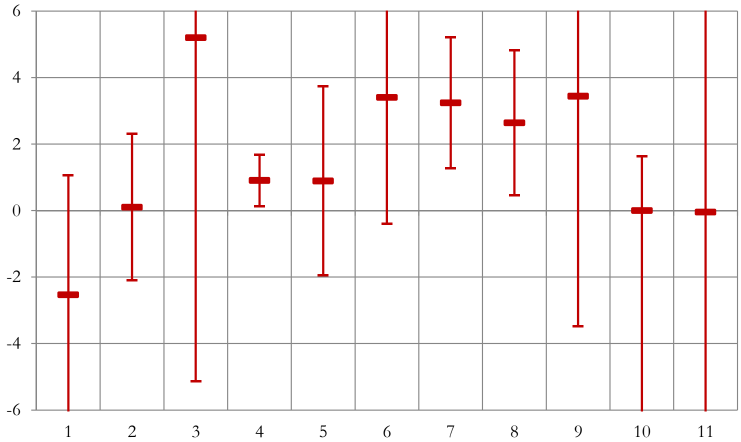

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries in Switzerland. We use the translog cost function as functional form.

Figure 3: Morishima elasticities, Swiss manufacturing industries

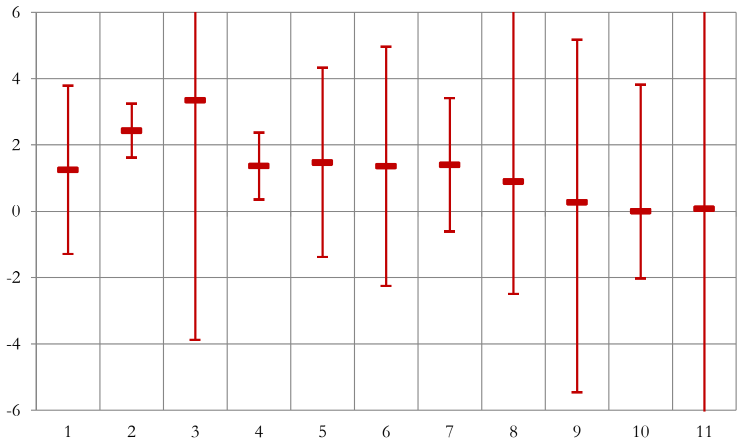
(a) $\text{MES}_{LE}$




(b) $\text{MES}_{EL}$



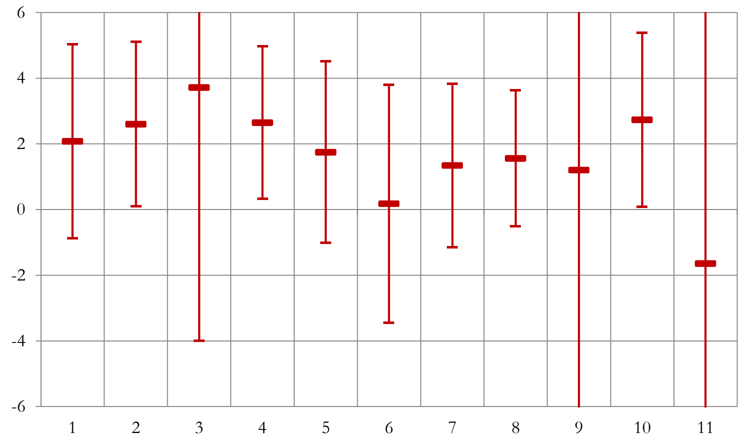

(c) $\text{MES}_{KE}$



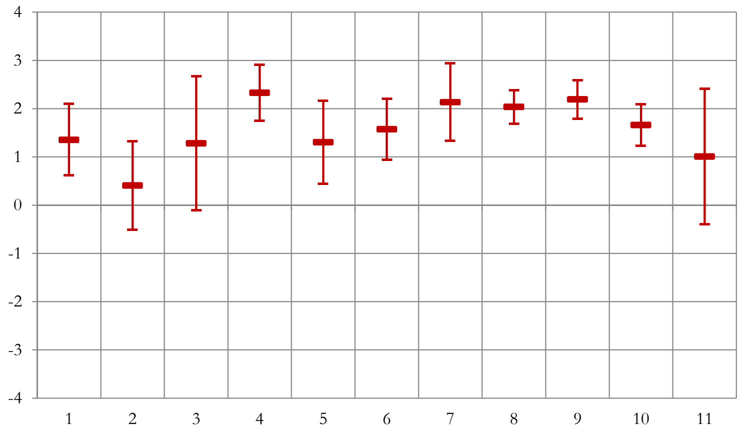

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries in Switzerland. We use the translog cost function as functional form.

Figure 3: Morishima elasticities, Swiss manufacturing industries (cont.)

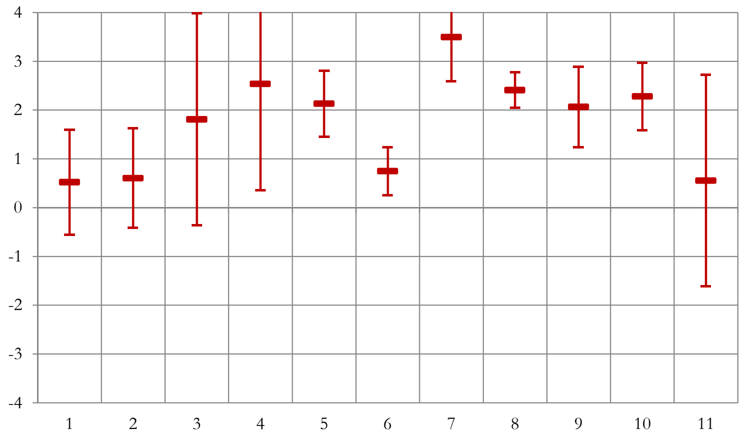
(d) $\text{MES}_{EK}$




(e) $\text{MES}_{KL}$




(f) $\text{MES}_{LK}$

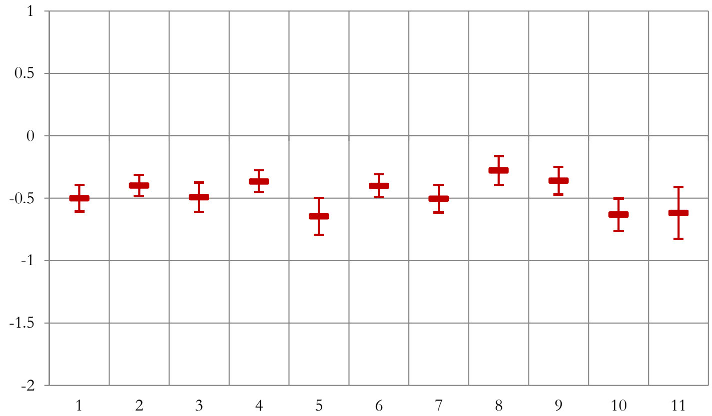



Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries in Switzerland. We use the translog cost function as functional form.

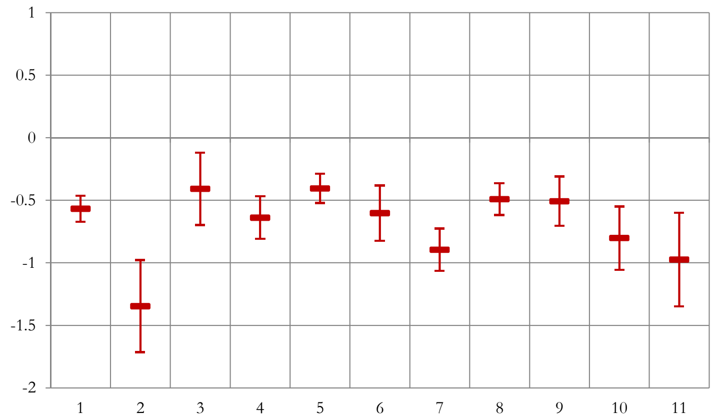
## A.6 Estimation Results: Pooled OECD Countries

Figure 4: Own-price elasticities, pooled across OECD countries

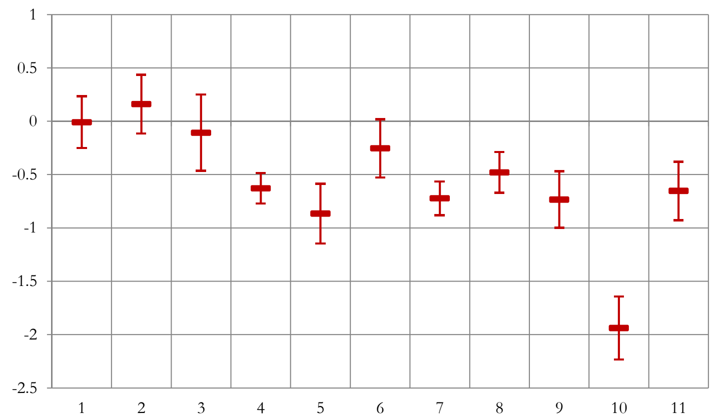
(a) $\text{CPE}_{LL}$




(b) $\text{CPE}_{KK}$



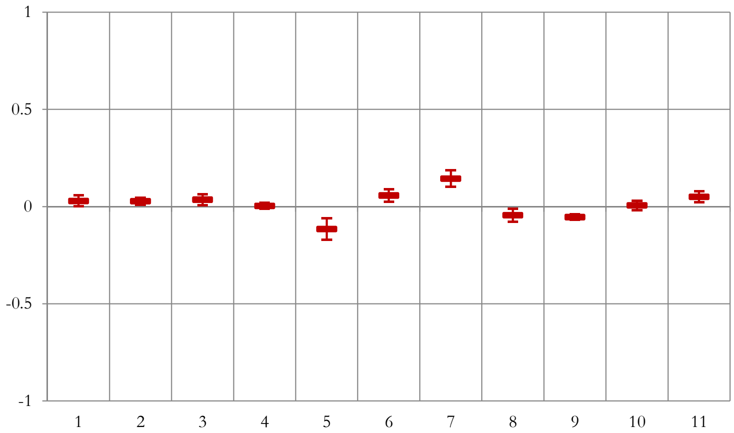

(c) $\text{CPE}_{EE}$



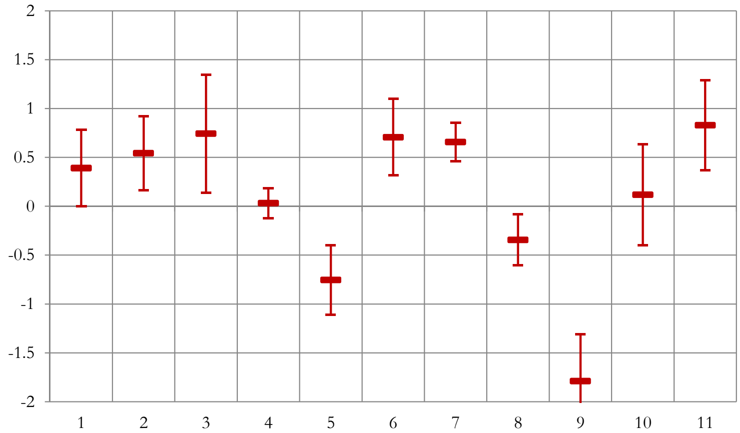

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries. Pooled estimates with fixed effects, translog cost function.

Figure 5: Cross-price elasticities, pooled across OECD countries

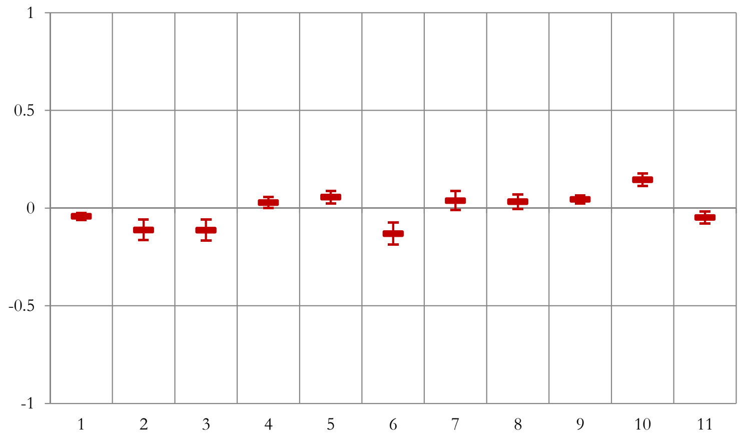
(a) $\text{CPE}_{LE}$




(b) $\text{CPE}_{EL}$



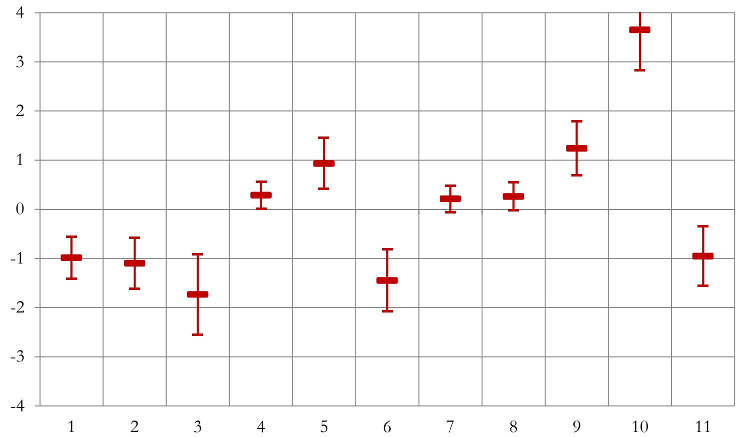

(c) $\text{CPE}_{KE}$



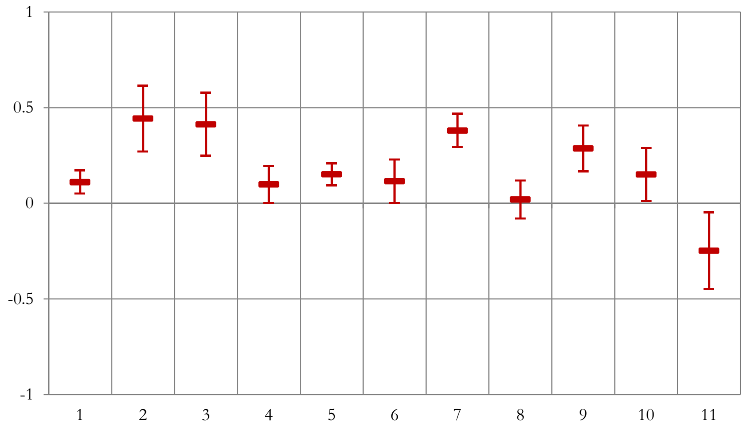

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries. Pooled estimates with fixed effects, translog cost function.

Figure 5: Cross-price elasticities, pooled across OECD countries (cont.)

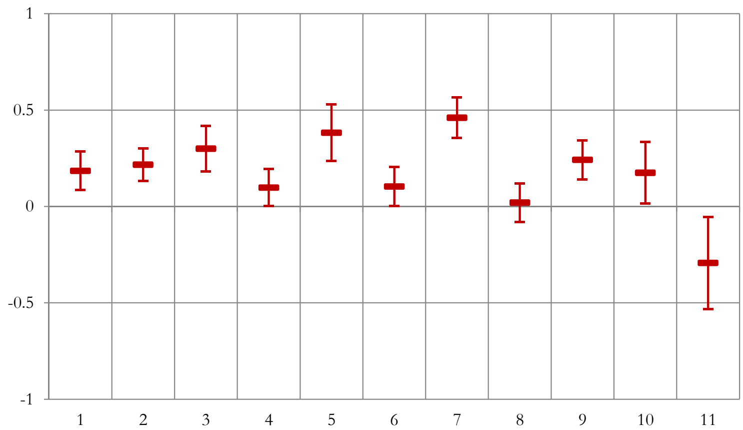
(d) $\text{CPE}_{EK}$




(e) $\text{CPE}_{KL}$



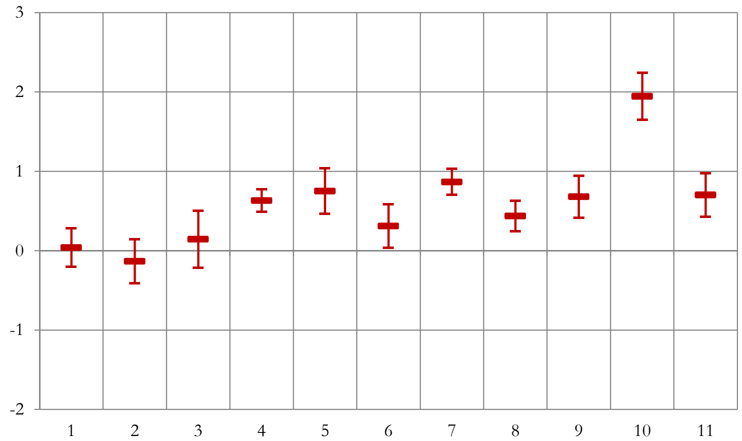

(f) $\text{CPE}_{LK}$



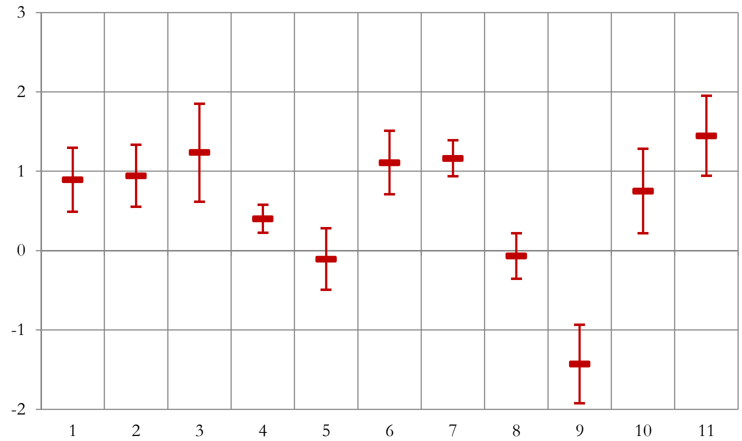

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries. Pooled estimates with fixed effects, translog cost function.

Figure 6: Morishima elasticities, pooled across OECD countries

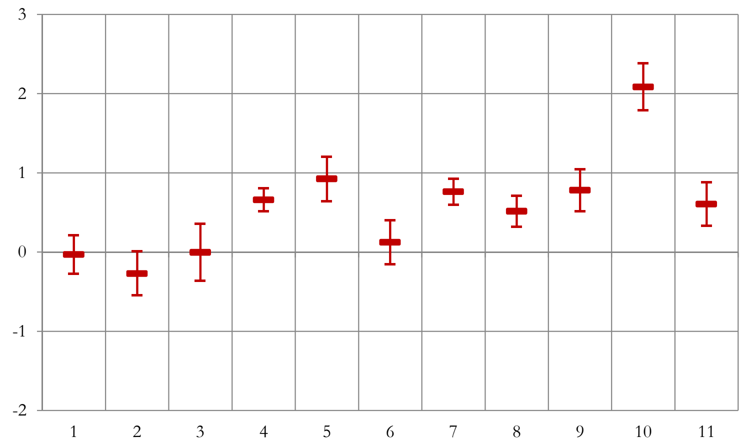
(a) $\text{MES}_{LE}$




(b) $\text{MES}_{EL}$



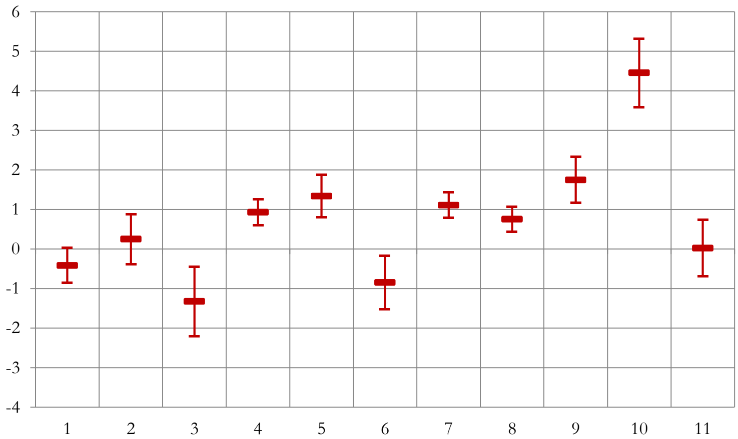

(c) $\text{MES}_{KE}$



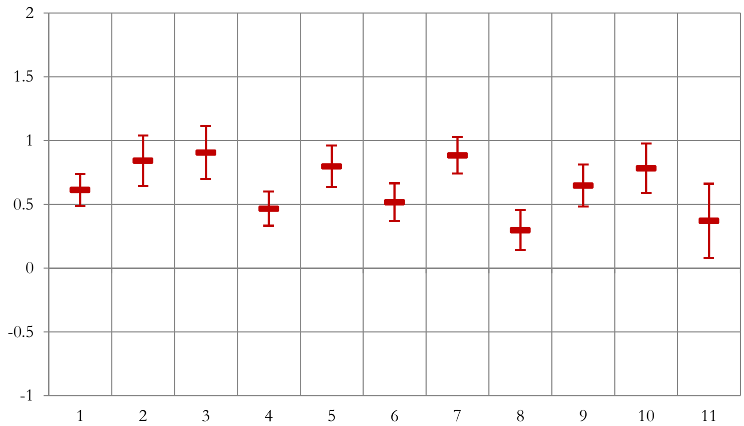

Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries. Pooled estimates with fixed effects, translog cost function.

Figure 6: Morishima elasticities, pooled across OECD countries (cont.)

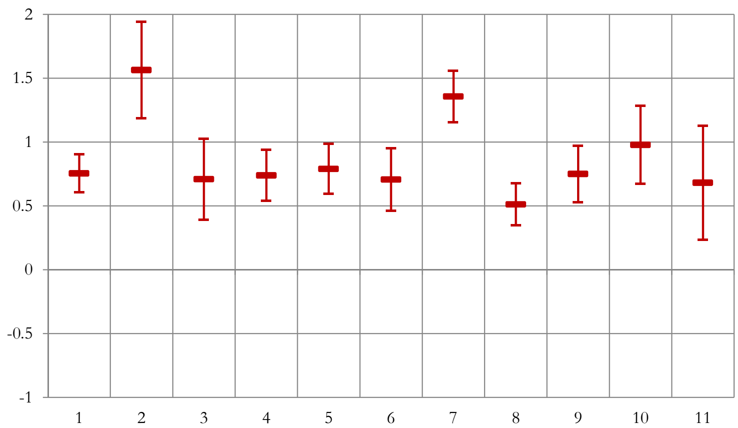
(d) $\text{MES}_{EK}$




(e) $\text{MES}_{KL}$




(f) $\text{MES}_{LK}$




Notes: Figures display the point estimates and the 95% confidence interval of the elasticities for the 11 manufacturing industries. Pooled estimates with fixed effects, translog cost function.

# Curriculum Vitae

Daniel Urs Müller was born in Lausen (Baselland, Switzerland) on September 19th 1980. He is the first child of Herbert Müller, a therapist and managing director of a social rehabilitation facility, and Gertrud Fischer-Müller, both originally from Germany. He has two brothers and one sister. From 1987 to 1992, Daniel attended the elementary school in Langenbruck, Baselland. Afterwards, he went to the secondary school in Oberdorf, Baselland, until 1996. He completed his Matura in 2000 at the Gymnasium in Liestal, Baselland, with a Major in Mathematics.

After one year in the private sector, Daniel began his studies of Economics at the University of Basel in 2002. He successfully completed his Bachelor's Degree in Business and Economics in 2005, and his Master's Degree in Business and Economics with a Major in Markets and Institutions in 2007. He received awards for the best master's degree and the VBÖ award for the best master's thesis in 2007. From 2007 to 2013, Daniel was employed as a doctoral assistant at the Department of Economic Theory at the University of Basel. In 2010, he successfully completed the full program for beginning doctoral students at the Study Center Gerzensee, founded by the Swiss National Bank.

From 2008 to 2012, Daniel worked part-time for the Institut für Wirtschaftsstudien Basel AG as an economist, where he contributed in numerous studies. In 2013, Daniel will take up a one-year post-doctoral position as a visiting researcher, funded by the Swiss National Science Foundation, at the Federal Reserve Bank in St. Louis.