Dissertations in Economics 2015: 1

Stockholm University

Taxes, Nudges, and Conformity

Essays in Labor and Behavioral Economics

Johan Egebark

Doctoral Thesis in Economics at Stockholm University, Sweden 2015

Taxes, Nudges, and Conformity

Essays in Labor and Behavioral Economics

Johan Egebark

ISSN 1404-3491
ISBN 978-91-7649-093-8
Printed in Sweden by Universitetsservice US-AB, Stockholm 2015
Distributor: Department of economics, Stockholm University

Till Anna-Karin och Iris

Acknowledgements

It would not have been possible for me to finish this thesis without the help and encouragement of a great number of people. I want to express my deepest gratitude to everyone who has been part of this journey.

First of all I want to thank my advisor Peter Fredriksson. His sharp insights, excellent guidance, and patient support have been instrumental in the completion of this thesis. I am particularly grateful that he supported my intentions to work on very different topics. I also want to thank him for always being generous with his time, and for always giving fast and invaluable feedback. I am also grateful to my co-advisor Jonas Vlachos, whose advice and support have been most valuable, especially during the writing of the second essay in this thesis.

I am grateful to Mathias Ekström, who has been amazing, not only as a co-author but also as a friend. I am privileged to have had him as my companion throughout this adventure; he has made me realize that conducting research can be truly fun and exciting. I hope that we can continue working (and skiing!) together in the future. I am grateful to Niklas Kaunitz, who has been a great co-author and a great friend. We have shared so many memorable moments during these years, memories that I will look back on with a smile on my face. I think we should be proud of ourselves, for patiently continuing working on the topics that we found relevant and important. Luckily it is now starting to pay off, and I hope that we can continue working (and tasting wine!) together also in the future.

I want to thank Yves Zenou, for always being positive and supportive, and for always being generous with his time. Thanks to Yves I got the opportunity to spend one year as a visiting scholar at Haas, Berkeley. For that I am very grateful. I also want to thank everyone at IFN, especially Magnus Henrekson for his advice, encouragement and support. Many fellow PhD students made these years so much more fun and interesting. I am particularly grateful to Patrick Augustin, Charlotta Boström, Johannes Breckenfelder, Sara Fogelberg, Manja Gärtner, Daniel Knutsson, Lisa Laun, Laurence Malafry, Kiflu Molla, Martin Olsson, Erik Prawitz, Eric Sjöberg, and Anders Österling.

There are also people outside the world of academia that I want to express my gratitude to. Linus Hasselström, thank you for your friendship and support. You are my role model in many aspects of life. Jan Schnitzler, your great spirit is a true source of inspiration. Thank you for welcoming me in Mainz, New York and Amsterdam. I also want to thank Maria Dahlqvist and Petter Göransson for great travel experiences. Thanks to all the players in Atlético Stockholm, and especially to Anders Björk, for offering a way for me to escape my studies. Thanks to my extended family, Mattias, Albin, Astrid, Inga-Lill, Lars, Gustaf with family, and Johan with family, for all your love and support.

Lastly, I would like to thank my closest circle. Mom and dad, thank you for your love, for your endless support, and for always encouraging us to pursue our dreams. Thank you Erik and Ylva for your love and support. Thank you Anna-Karin, my love. Words can never describe how patient and supportive you have been throughout these years. Thank you for making me realize what is important in life. Thank you for all the things we have shared. I love you. Last, but not least, thank you Iris for bringing perspective and true happiness.

Stockholm, February 2015

Johan Egebark

Contents

Introduction

This thesis consists of four self-contained essays. The first two essays, even though being independent of each other, are closely related. They both examine the effects of government policies implemented to address the growing threat of youth unemployment. Broadly speaking they deal with two main questions: Do employers respond to tax changes that suddenly make it cheaper to hire a young worker? Do tax reductions affect a young individual's decision to be self-employed? The remaining two essays concern questions within the field of behavioral economics. One presents the results from a natural field experiment that was set up to measure the causal effects of two resource conservation programs. The other one explores the micro-level foundations of behavioral contagion. Below, I provide a brief introduction to the various topics.

Essay 1: Do Payroll Tax Cuts Raise Youth Employment?

High and persistent youth unemployment is a major challenge for many developed economies. In the OECD as a whole, unemployment for individuals below 24 years of age has been twice as high as for those aged 25–64 since the beginning of the 1990's. In recent years, in the wake of the 2008 financial crisis, young people's employment opportunities have worsened even further. The current situation in many countries (Spain and Greece, for example, have youth unemployment rates of up to 50 percent) has spurred a wide and lively debate on what policies should be undertaken to effectively alleviate the youth unemployment problem.

The aim of the first essay is to examine whether targeted payroll tax reductions are an effective policy for raising youth employment. We make use of a Swedish reform, implemented in two steps in 2007–09, which suddenly reduced the payroll tax for employers of young workers. The reform created substantial variation in tax rates across cohorts, and thus offers a good opportunity to study the causal effect of payroll taxes on youth employment. By contrasting individuals below the treatment-defining age threshold to those just above, we find that lowering payroll taxes for young workers has a significant, but small, impact on employment. For 20-25 year-olds, the (relative) employment increase was around 2.5 percent in 2007 and 1.4 percent in 2008; for individuals close to the treatment defining cutoff, the effect was around 1.4 percent, both in 2007 and in 2008. We find no evidence of any additional effect on employment in 2009–2010, i.e. in the midst of recession, even though there was an additional cut in the tax rate these years. This is an important finding, not the least from a policy perspective, as it suggests that even large tax reductions do not counteract the negative impact of economic slowdowns. We estimate the gross cost per created job for 20–25 year-olds to SEK 0.8 to 1.6 million ($100,000 to $225,000). Since this corresponds to more than four times the cost of hiring the same number of workers at the average wage, we conclude that targeted payroll tax reductions are an expensive way to boost employment for young individuals.

Essay 2: Effects of Taxes on Youth Self-Employment and Income

High youth unemployment, in Sweden and in many other OECD countries, could reflect the fact that young individuals have few options in the formal sector, due to for example their lack of work experience or social connections. One way for them to exit unemployment could therefore be to start their own business. Despite the potential role that self-employment could play, there are basically no (credible) evaluations of the effectiveness of different policies to stimulate self-employment (OECD, 2012, 2013). The second essay provides hard evidence of the effect of taxes on young individuals' decision to run a business. I make use of the same Swedish reform as in the first essay. In addition to reducing the payroll tax for employers of young workers, the reform reduced the self-employment tax paid by young business owners. By using a Difference-in-Differences design that contrasts individuals on either side of the treatment-defining age cutoff, I show that youth self-employment is insensitive to tax reductions. As in the first essay it is striking to see that the reduced taxes have no impact in the recession years 2009–10. I also consider the effect of the tax reductions on the intensive margin. For those that are defined as self-employed I find large positive effects on income from self-employment, and negative effects on income from wage employment. Taken together, these findings suggest that the lower taxes caused the self-employed to reallocate time from regular work to self-employment.

Essay 3: Can Indifference Make the World Greener?

The aim of the third essay is to add to the growing body of research that uses so called nudging to affect decisions (see, e.g., Thaler and Sunstein, 2003; Allcott and Mullainathan, 2010; Sunstein and Reisch, 2014). We do so by evaluating the effect of two interventions aimed at reducing people's consumption of paper. The activity that we consider, document printing, consumes a vast amount of inputs every year. Estimates suggest that U.S. office workers use roughly five million metric tons of paper annually, amounting to around 20 million metric tons of wood. If this amount could be reduced by only five percent, roughly six and a half million trees (or 6,500 acres of forest) would be saved; this, in turn, would prevent the equivalent annual greenhouse gas emissions of 140,000 cars.

When sending a document to a printer a user can typically choose whether to print on both sides of a sheet of paper (duplex) or to print on only one side (simplex). Duplex printing reduces the number of sheets that is used, and is thus less resource intensive (i.e., greener). We use this functionality in a natural field experiment at a large Swedish university in order find out which type of behavioral interventions work and which do not. Our baseline intervention consisted of an e-mail campaign that actively tried to convince people to cut back on printing in general, and to use duplex printing whenever possible. The second intervention constitutes a more passive approach as it exploits people's tendency to stick to pre-set alternatives: at random points in time we changed the printers default settings, from simplex to duplex printing.

The effect of the two interventions differed sharply. The random subset of employees that was subjected to the first intervention—the moral appeal—displays no sign of changed behavior, not even on the day the message was communicated. In sharp contrast, we document a substantial and immediate effect of changing the default printer setting. On average, daily paper consumption dropped by 15 percent due to the changed settings, and this reduction occurred already on the very day that we introduce the intervention. We find some indications that the effect is larger for men, and for older subjects.

Our study makes important contributions by looking into some important, but often neglected, aspects of default rules. While many behavioral interventions, such as feedback and social comparison, often have significant effects in the short run, there is still limited evidence on the long-term impacts. By studying printing behavior more than six months after the intervention, we show that default rules can be influential also in the somewhat longer run. Second, we show that printing demand (measured by the number of printed pages and documents) is independent of the pre-set alternative. This finding is important as it indicates that changing defaults avoids unintended adverse effects.

Essay 4: The Origins of Behavioral Contagion: Evidence from a Field Experiment on Facebook

This essay explores the details of contagion dynamics. We use a field experiment on the networking site Facebook to examine how small changes in the size of the influencing group, and the introduction of social ties between the source and the target, affects the decision to conform. Members of Facebook express positive support to content on the website by clicking a Like button. Making use of people's actual accounts, we study whether users are more prone to support content if someone else has done so before.

We expose the subjects to three different treatment conditions: (1) one unknown individual has Liked the update, (2) three unknown individuals have Liked the update and (3) one user with a central position in the network has Liked the update. The results from this exercise are striking: whereas the first treatment condition left subjects totally unaffected, both the second and the third more than doubled the probability of Liking an update. This shows that the behavior of a single individual spreads, but only among friends within a network. In addition, once a sufficient number of in-group members have adopted the behavior, they start to affect people outside the network. An important contribution of the experiment is that subjects act in their natural environment, and are unaware of the fact that they are part of an experiment (Al-Ubaydli and List, 2012). This leads us to conclude that decades of social influence research from lab settings, including Asch's (1955) influential study and more recent contributions such as Goeree and Yariv (2010), cannot be dismissed as an artefact shaped by suspicious subjects, strange environments, or influential experimenters.

References

Al-Ubaydli, O. and J. A. List (2012). On the generalizability of experimental results in economics. NBER Working Paper Series.

Allcott, H. and S. Mullainathan (2010). Behavior and energy policy. *Science 327*(5970), 1204–1205.

Asch, S. (1955). Opinions and social pressure. *Scientific American 193*(5), 31–35.

Goeree, J. K. and L. Yariv (2010). Conformity in the lab. Revise and resubmit Economic Journal.

OECD (2012). Policy Brief on Youth Entrepreneurship: Entrepreneurial Activities in Europe. Technical report, OECD, Paris.

OECD (2013). Self-employment among the youth and seniors: Entrepreneurship at a Glance 2013. Technical report, OECD, Paris.

Sunstein, C. R. and L. A. Reisch (2014). Automatically green: Behavioral economics and environmental protection. *Harvard Environmental Law Review 38*, 128–158.

Thaler, R. H. and C. R. Sunstein (2003). Libertarian paternalism. *American Economic Review: Papers and Proceedings 93*(2), 175–179.

Do Payroll Tax Cuts Raise Youth Employment?*

Johan Egebark† Niklas Kaunitz‡

Abstract

This article examines whether targeted payroll tax reductions are an effective means to raise youth employment. In 2007, the Swedish employer-paid payroll tax was cut on a large scale for young workers, substantially reducing labor costs for this group. Using the variation in payroll taxes across cohorts, we estimate a significant, but small, impact both on employment and on wages. Our employment and wage estimates in combination imply that the firms' elasticity of demand for young workers in Sweden is at -0.37. Since the estimated cost per created job is at more than four times that of directly hiring workers at the average wage we conclude that payroll tax cuts are an inefficient way to boost employment for young individuals.

Key words: Youth unemployment; Payroll tax; Tax subsidy; Labor costs

JEL classification: H25, H32, J23, J38, J68

*We thank Anders Björklund, David Card, Mathias Ekström, Peter Fredriksson, Helena Holmlund, Markus Jäntti, Lisa Laun, Assar Lindbeck, Matthew Lindquist, Erik Mellander, Martin Olsson, Per Skedinger and Björn Öckert for helpful comments. Seminar participants at IFAU, Uppsala, and SOFI, Stockholm, as well as participants at the 24th annual EALE Conference in Bonn and The 3rd National Conference of Swedish Economics in Stockholm, have also provided valuable suggestions. We thank Nina Öhrn for excellent research assistance. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

†Department of Economics, Stockholm University and the Research Institute of Industrial Economics (IFN). E-mail: johan.egebark@ne.su.se

‡Swedish Institute for Social Research (SOFI), Stockholm University. E-mail: niklas.kaunitz@sofi.su.se

1 Introduction

High and persistent youth unemployment is a major challenge for many developed economies. In the OECD as a whole, unemployment for individuals below 24 years of age has been twice as high as for those aged 25–64 since the beginning of the 1990's. In addition, young people's employment opportunities have worsened even further in the wake of the 2008 financial crisis. Since labor market difficulties encountered in early working life are known to have lasting consequences, an increasing number of young people risk ending up in long-term unemployment.[1] Consequently, there is a wide and lively debate on what policies should be undertaken to improve young individuals' labor market prospects.

We examine whether targeted payroll tax reductions are an effective means to raise youth employment. Payroll taxes in Sweden are proportional to the employee's gross wage and are paid by the employer. In 2007–09, the tax rate for employers of young workers was reduced on a large scale in two steps. The first reduction, in effect 2007–08, lowered the payroll tax rate with 11 percentage points for employees who at the start of the year had turned 18 but not 25 years of age. In 2009, the reduction was extended to encompass all individuals who at the start of the year had not yet turned 26 years of age; at the same time, the rate was reduced with an additional 6 percentage points for the eligible individuals. Using this variation in payroll tax rates across cohorts, we investigate the causal effect of payroll taxes on youth employment.

We use Difference-in-Differences (DiD) to identify the effects of the payroll tax reductions, pitting individuals in the target group against slightly older individuals who were not subjected. Identification is, however, complicated by the fact that individuals of different ages tend to experience different employment cyclicality, with younger workers displaying larger cyclical variations. We deal with this problem—which essentially constitutes a threat to the identification assumption of parallel trends—by including a large number of covariates in the DiD model. We estimate the effect for the entire target group as well as for different subgroups, such as foreign-born and the unemployed. As a special case, we consider treatment-control pairs that are defined at a very small

[1] See, e.g., Gregg (2001), Nordström Skans (2004) and Gregg and Tominey (2005) for studies on the so-called scarring effect of early unemployment.

bandwidth around the treatment-defining age threshold; this resembles a regression discontinuity design, but with controlling for pre-reform discontinuity.

We find that lowering payroll taxes for young workers has a significant, but small, impact on employment. For the whole target group, the relative employment increase was around 2.5 percent in 2007 and 1.8 percent in 2008, whereas for individuals close to the treatment defining cutoff, the effect was around 1.4 percent, both in 2007 and in 2008. We find some support for the existence of substitution effects, implying that the reform may have created jobs for one group of individuals at the expense of another. Importantly, the presence of substitution effects also means that the absolute effect on employment is potentially smaller than what our estimates suggest. A striking finding is that there is no additional effect on employment of the 2009 extended reduction; this suggests that even large tax cuts cannot counteract the negative impact of economic slowdowns. Finally, our results show that even though the reform created a relative price wedge that induced employers to hire (or to keep) a young worker, it did not lead to any permanent increase in the likelihood that this individual is employed.

When it comes to explaining the modest impact, we point at certain observations that help us interpret the results. First, since wages did not adjust, shifting of the incidence of the tax burden to higher wages cannot explain the small employment effects. Second, since the tax cut had no impact at all for foreign-born youths, nor for individuals registered as unemployed, we argue that labor supply constraints are not the main issue. The question then arises why the demand elasticity of firms is so low. We argue in favor of demand constraints: for the group of uneducated, unexperienced young workers, labor costs are still too high—even with the payroll tax reduction in place.

Our employment and wage estimates in combination imply that the firms' elasticity of demand for young workers in Sweden is at around -0.37. Using a different metric: the estimated gross cost per created job for 20–25 year-olds was SEK 0.8 to 1.8 million ($100,000 to $225,000). Since this corresponds to more than four times the cost of hiring the same number of workers at the average wage, we draw the conclusion that targeted payroll tax reductions are an inefficient way to boost employment for young individuals.

The rest of the paper is organized as follows. Section 2 gives a brief overview of the previous literature. Section 3 presents some of the institutions specific to the Swedish setting. Section 4 describes the data and section 5 the methodology we apply. Section 6 gives the results, which are further analyzed in section 7. Section 8 provides a discussion and section 9 concludes.

2 Previous literature

Previous evidence on the effects of payroll tax cuts typically concerns general reductions. The basic result for the U.S. is that of extensive shifting of the incidence of the tax onto workers; hence, there are, at most, marginal employment effects (see, e.g., Gruber, 1997; Anderson and Meyer, 1997, 2000; Murphy, 2007).[2] However, since these studies may suffer from endogeneity problems it is difficult to draw decisive conclusions. For example, Anderson and Meyer (1997, 2000) exploit firm, or industry, level variation in unemployment insurance (UI) taxes. Since the UI tax paid by the firm is determined by the firm's lay-off history, and thus is potentially endogenous, it is not clear that the estimates can be interpreted as the causal effect of the UI tax.

More convincing evidence is found in studies that evaluate selective payroll tax reforms. Examples include Bohm and Lind (1993), Bennmarker et al. (2009) and Korkeamäki and Uusitalo (2009) who evaluate reductions targeted towards specific regions in Sweden or Finland. None of these studies find any effects on employment. However, compared to the U.S., the degree of shifting is small. Bennmarker et al. (2009) find that a 1 percent reduction in wage costs increased wages by 0.32 percent, whereas in Korkeamäki and Uusitalo (2009) the increase was 0.6 percent.

Besides the above-mentioned literature, there are some studies that focus on workers who display poor labor market outcomes. Kramarz and Philippon (2001) examine the impact of changes in total labor costs on employment of low-wage workers in France between 1990 and 1998. Their results suggest that a 1 percent increase of the labor cost leads to a 1.5 percent increase in the probability of transiting from employment to non-employment, whereas lower labor costs had no impact on transitions from non-employment to employment. Since payroll tax cuts were offset by

[2] Gruber (1997) studies manufacturing firms in Chile and finds that the incidence of payroll taxation is fully on wages, with no effect on employment.

rising minimum wages it is difficult, however, to distinguish between the effect of changes in payroll taxes from that of changes in minimum wages. Finally, Huttunen et al. (2013) study a Finnish hiring credit targeted at the employers of older, full-time, low-wage workers. They find no effects on employment or wages of the eligible groups, but a small increase in working hours among those who were already employed.

To the best of our knowledge, the only other study that examines payroll tax reductions explicitly aimed at young workers is Skedinger (2014). Skedinger looks at the same reductions as we do and studies the effects for the Swedish retail industry. He finds small or no effects on job accessions, separations, hours worked and wages. The most important difference between our study and Skedinger's is that he only considers one industry. Thus, he cannot assess the overall employment effect in the economy since he cannot separate new labor market entrants from movements between sectors. In addition, since we are using much more detailed data, we are able to study treatment effect heterogeneity with respect to immigration status and unemployment status.

3 Institutional background

3.1 Youth unemployment in Sweden

Official records show that youth unemployment in Sweden is currently high. Unemployment for 15–25 year-olds was roughly at 24 percent in 2013, which is three times higher than overall unemployment (Statistics Sweden, 2014). In 2007 and 2008, which are the years that we mainly focus on in this study, youth unemployment was somewhat lower, at 20 percent. In 2009–10, when the Swedish economy was fully hit by the financial crisis, it increased to 25 percent. It is sometimes argued that these (official) figures exaggerate the problem of youth unemployment in Sweden, mainly due to the fact that a large number of the unemployed participate in different types of education. Excluding those who study full-time lowers unemployment for 15–25 year-olds to 12 percent in 2013. However, it is not obvious that this adjustment makes sense: many might chose to study since it is difficult to find a job, even though they rather would be working.

We complement these figures with two other measures to provide some further understanding of the problem in the Swedish case. First, about 10 percent of all 20–24 year-olds were not employed and not in any education or training in 2013 (i.e., they belong to the so called NEET category). In 2007–08, the corresponding figure was 12 percent, and in 2009–10 roughly 13 percent (Statistics Sweden, 2014). A second measure looks at registrations at the unemployment office. The data that we use for the analysis below contain yearly information on job search activity, and so we can observe those that are registered as looking for a job. 21 percent of all 20–24 year-olds were registered at the unemployment office at some point during 2007–08, and 8 percent were registered for more than 100 days. During the recession years, these figures increased to 24 percent and 12 percent, respectively.

3.2 Swedish payroll tax reductions

Swedish payroll taxes are proportional to the employee's wage bill and, in contrast to e.g. the U.S., fully paid by the employer. The tax consists of seven mandatory fees, financing welfare services such as pensions, health and disability insurances, and other social benefits. Up until the beginning of the 1980's the payroll tax rate was the same for all employers in Sweden, but over the last 30 years there have been some exceptions. First, firms in so called regional support areas (RSA) in the northern parts of Sweden were twice subjected to reductions of roughly 10 percentage points in efforts to boost employment in these areas.[3] Second, besides these regional reductions, payroll taxes were cut for small firms in all of Sweden between 1997 and 2008.[4]

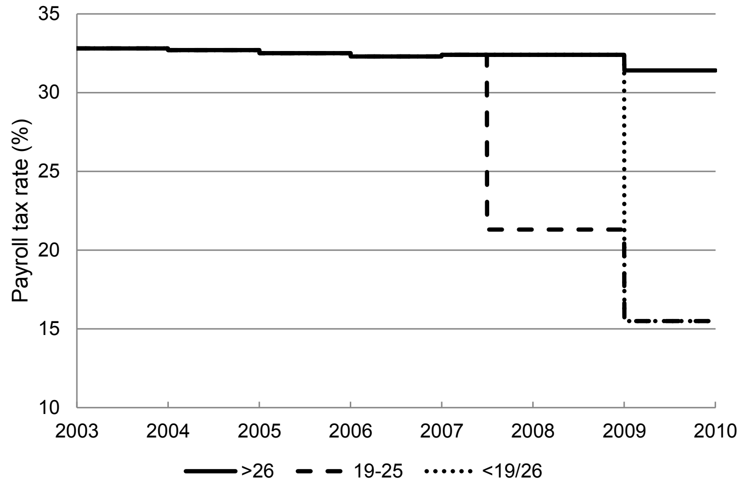
We study reductions targeted explicitly at young workers. Figure 1 provides a graphical illustration of the changes in the tax rate. On July 1, 2007, the payroll tax was cut by around 11 percentage points for individuals who at the start of the year had turned 18 but not 25 years of age. Six out of seven mandatory fees were halved, reducing the tax rate from 32.42 to 21.32 percent.[5]

[3] Neither Bohm and Lind (1993), who study reductions implemented between 1984 and 1999, nor Bennmarker et al. (2009), evaluating reductions introduced in 2002, find any employment effects.

[4] Firms with up to three employees were allowed a 5 percent reduction for wage sums up to around SEK 750,000 ($95,000) per year. Thus, this cut was relatively small, both in magnitude and comprehension. To the best of our knowledge, this reduction has not been evaluated.

[5] The date July 2007 is first mentioned in a press release from the ministry of Finance in October 2006. This date was confirmed when the new policy was ratified in the parliament on 15 March 2007. The only fee that was left unchanged was the pension fee. Individuals who are self-employed pay *egenavgifter*, roughly equivalent to payroll taxes paid by employers. These fees were also cut by about 10 percentage points, in order to avoid distortionary effects

Figure 1: The payroll tax reductions

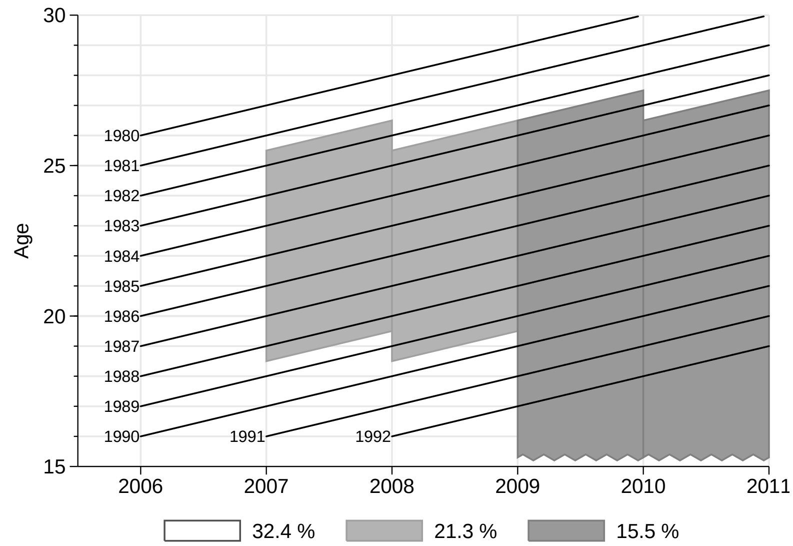


On January 1, 2009, the reform was modified in two ways. First, the tax reduction was extended to encompass all individuals who at the start of the year had not yet turned 26 years of age, i.e., the target group was extended at both ends. Second, the payroll tax reduction was increased, down to 15.52 percent. The payroll tax reductions were automatically implemented via the tax system, i.e., the employers did not have to send in an application to benefit from the lower tax rates. Figure 2 illustrates how different cohorts are subjected to the payroll tax reductions. In 2007, the target group consists of individuals born 1982–88 whereas in 2008 it consists of those born 1983–89. For simplicity, hereafter an age group a denotes all individuals who turn a during the year. With this terminology, the target group of the 2007 reform is referred to as "individuals aged 19–25", and the target group of the 2009 reform as "individuals aged 26 or below".

The group of 19–25 year-olds comprised around 10 percent of the labor force aged 15–64 in 2007, implying that the number of individuals directly affected by the new regime was substantial. Since the reductions applied also to existing employments, the cost of the reform was sizable. Yearly foregone tax revenues was SEK 9 billion (around $1.1 billion) in 2007 and SEK 9.9 billion in 2008 (around $1.2 billion), corresponding to about 1 percent of the fiscal budget in these years. These

with respect to choice of occupation. Besides the statutory payroll tax, collective-bargaining agreements require most employers to pay around 10 percent of gross wages to finance job search support, retraining and severance payments when employees are laid off. As these fees are not legislated, they were unaffected by the tax reduction.

Figure 2: Evolution of treatment status across cohorts



figures increased substantially when the reductions were extended, resulting in foregone revenues at SEK 17 billion ($2.1 billion) in 2009 and SEK 18 billion ($2.3 billion) in 2010.

3.3 Other relevant labor market reforms

With the purpose of increasing employment, both in general and for specific groups, several labor market reforms were introduced in Sweden during 2007. First, temporary subsidies for firms that hire individuals who have been unemployed or have received sickness or disability benefits, *New Start Jobs* (NSJ), were introduced on January 1, 2007. In 2007–08, individuals aged 20–24 could apply for the subsidy after six months of non-employment, whereas those who had turned 25 could apply only after twelve months of non-employment—thus, in contrast to the payroll tax cut, it was the exact age that mattered. In 2009, this cutoff was modified so that those who at the start of the year have turned 20 but not 26 were eligible after six months.[6] Consequently, in 2007–08 the target groups overlapped, and from 2009 onwards they completely coincide. In principle, this raises a

[6]When introduced, the subsidy was equal in size to the payroll tax amount. In 2009, the size of the subsidy increased to twice the payroll tax. The subsidy is given for a period equally long as the earlier non-employment spell and up to 5 years.

concern that the employment estimates of the payroll tax reduction will be contaminated. It turns out, however, that the number of applications for NSJ (available in our data) was comparatively low, at about 0.5 percent of the ages 20–26, and the difference in shares between 20–25 year-olds and 26-year-olds—the potential bias of our estimates—is around 0.1 percentage points. We thus conclude that this is not a source of concern.

Second, income tax deductions were introduced in Sweden on January 1, 2007, with the purpose of increasing labor supply in general. These deductions apply to all workers, regardless of age, but we cannot rule out that there is heterogeneity in labor supply effects with respect to age. If younger workers' labor supply responded differently, we risk misestimating the effect of the payroll tax reductions. Edmark et al. (2012) show that it is difficult to evaluate this deduction scheme due to the lack of unaffected comparison groups; hence, we do not know exactly how different age groups responded. In this study we assume that the response was similar for individuals close in age.

Finally, a third reform concerns employment protection legislation. Loosening of regulation in 2007 made it easier for employers to use fixed-term contracts. As temporary work is relatively more widespread among young workers, employment (and wages) may have been affected more for younger workers. However, Skedinger (2012) reports that only 1.4 percent of all temporary workers were employed with the new regulations in 2008. The reform, thus, had little impact in practice.

3.4 Wage formation in Sweden

Wage setting in Sweden has traditionally been characterized by a high degree of central bargaining. Over the last 10–15 years, there has been a substantial move toward the decentralization of negotiations, but many workers still have centrally agreed wages and this is likely to be more common for young workers.[7] In 2007, between April and July, central agreements covering 75 percent of all workers were renegotiated—i.e., before the implementation of the 2007 reform but after its passing

[7]Union density was at 80 percent in 1990 and 79 percent in 2000, and the share of workers covered by collective-bargaining agreements is even higher. The influence given to the local bargaining parties varies by sector. The private sector, to which most young workers in Sweden belong, has a higher degree of central wage setting than the public sector. See Fredriksson and Topel (2010) for a detailed discussion of the Swedish labor market.

in the parliament in March 2007 (National Mediation Office, 2007). New agreements were not made until 2010, one year after the implementation of the new extended reductions.

Another institutional feature specific to the Swedish labor market is the fact that minimum wages are negotiated, not legislated as in most other OECD countries. Collective-bargaining agreements differentiate wages based mainly on age, experience and levels of skill. This means that younger workers are more likely to have wages bound by the minimum wage level.

4 Data

The data are collected by Statistics Sweden (SCB) and contain yearly information on employment and demographical characteristics for all individuals living in Sweden at or above 16 years of age in 2001–10 (the Louise and Rams data sets). The employment data contain, for each individual and year, start and end months as well as total taxable income from each employment source during the year. From this information we can deduce, for each individual and month, total monthly income from paid work. In addition, we have access to detailed information on employment characteristics for a subsample of all employees (measured between August and November each year), containing data on actual monthly wages, work rate, industry affiliation of workplace, etc. For public sector employers, the total population is surveyed through official registers, while firms in the private sector are sampled using a stratification scheme.[8] This subsample, in addition to being used in the wage analysis, is combined with the income data from the tax registers to create monthly measures of employment for all individuals.

Our employment measure is constructed in the following way. Starting out from the reduced sample of employed workers, for all individuals working at least 25 percent of full-time, we partition the sample in cells defined by all unique combinations of age, gender, three groups of education, firm sector (local/central public, blue-collar/white-collar private), and year. For each cell, we calculate the 10th percentile of actual, full-time equivalent wage; these values are to be used as cutoff values, serving as an income criterion for full-time employment. These monthly cutoff values are matched to

[8]The stratification is based on six firm size classes and 54 industry groups, giving a total of 324 strata. Stratification weights are supplied with the data and used for table 1 and in the analysis of wages.

the tax register data on all individuals. For each month that an individual's taxable income exceeds the appropriate cutoff value, she is, thus, classified as being full-time employed. Our employment measure uses the quarter of these income cutoffs to arrive at a measure of working *at least* 25 percent of full-time, for a particular month.[9]

It should be noted that our employment measure is likely to be misleading when comparing specific months within a given year: the income cutoffs used for deducing employment status are computed on a yearly basis, while wages tend to rise continuously over time. Moreover, information on employment spells are only available separately for each year. This means that, e.g., for an employment stretching from December 2007 to April 2008 we have the exact income for December, but a 4-month average for January to April. We therefore use an annual measure of employment, taking the average of monthly employment status for each year.[10] Note that this method, in conjunction with our estimation method, handles most forms of remaining measurement errors. Only an error that evolves differently over time for different age groups, and that is uncorrelated to all control variables, would result in a bias in our DiD estimates.

Table 1 shows summary statistics divided by age, both for the full population (panel A) and for the smaller subsample (panel B). The table highlights some of the large differences in background characteristics across ages. For example, only 8.6 percent of the 20-year-olds have some form of education above high school, whereas among 27 year-olds this figure is 44.6 percent. Moreover, while foreign-born constitute 12.4 percent of the 20-year-olds, the same figure for 27-year-olds is 18.3 percent. These differences are unlikely to be stable over time since they depend on, e.g., the state of the economy, demographical changes and fluctuations in immigration. Panel B characterizes

[9]In practice, the procedure is slightly more complicated: as cells with ten or fewer individuals (about two percent of all cells) cannot be used (otherwise we would overestimate the 10th percentile), the cutoffs for these cells are instead estimated. We predict the (log of) wage cutoffs using the other cells in a linear regression, controlling for all interactions of female-age-year, and female-age-year-education. In other words, we impute the wage cutoffs for the small cells through linear interpolation. When an individual has multiple income sources for a particular month, the largest income source is used for sector matching. We have tested using the 20th percentile instead of the 10th percentile when defining full-time employment; although raising the cutoff point, by definition, lowers all employment *levels*, the dynamics are essentially the same and, thus, this does not significantly change our results. Further, we have experimented with using different work rate conditions for the outcome variable, such as 10 or 50 percent of full-time employment. Again, the results are not much affected (see section 6).

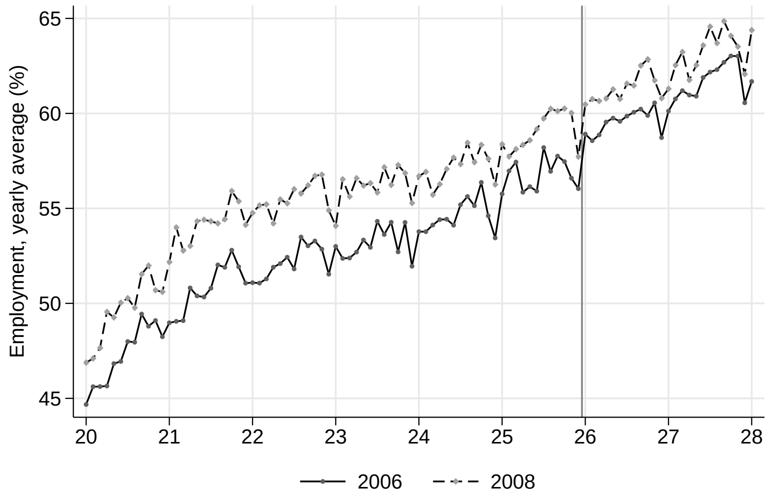
[10]Our measure differs from the official ILO definition of employment, according to which an individual is considered to be employed if working at least one hour per week (ILO, 1983). For our purpose, this is too lax a restriction; we are interested in employments that actually have an economic impact for an individual.

Table 1: Summary statistics, year 2006 (percentages)

	Age cohort, 2006				
	20	23	25	27	30
	Panel A: Full sample				
Employed, quarter-time	47.3	53.2	56.8	61.7	65.6
Employed, full-time	15.7	25.0	31.0	37.8	40.7
Educ. below high school	14.4	12.5	11.8	13.1	8.2
Educ. high school	77.0	53.5	46.2	42.3	46.1
Educ. above high school	8.6	34.1	42.0	44.6	45.7
Female	48.7	48.8	49.1	49.0	49.0
Foreign-born	12.4	16.6	17.7	18.3	19.0
N	112,618	105,303	108,174	110,202	112,582
	Panel B: Employed subsample				
Wage, full-time eq. (SEK)	18,428	19,776	21,028	22,205	23,972
Work rate (mean %)	86.3	90.1	92.7	93.7	93.7
Tenured	60.3	67.3	69.8	75.2	80.1
Public sector	15.1	20.4	23.3	25.8	26.9
Educ. below high school	8.1	10.8	6.4	9.5	4.5
Educ. high school	83.8	58.6	50.4	44.4	48.7
Educ. above high school	8.1	30.6	43.2	46.1	46.8
Female	44.4	45.7	45.1	45.6	44.7
Foreign-born	8.2	10.2	10.8	11.1	11.6
Sum of weights	46,150	48,740	61,664	64,875	75,815
N	22,621	27,393	35,836	38,834	46,073

Notes: The employment measure is constructed as described in section 4. For the employed subsample, the sum of stratification weights indicates population size.

Figure 3: Employment rates by age, 2006 and 2008

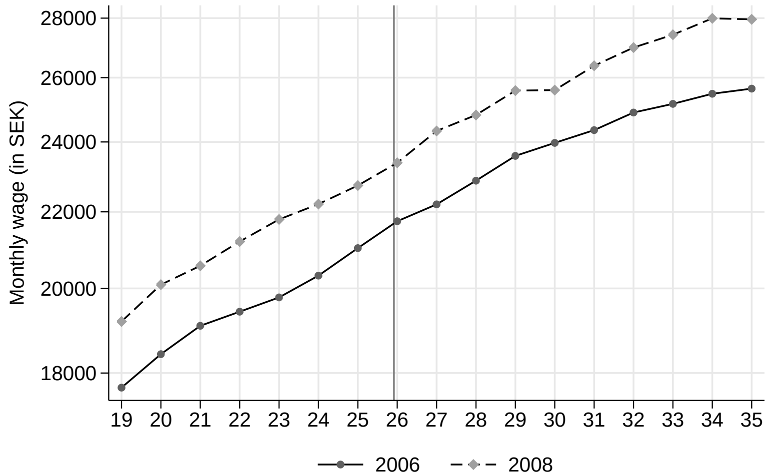


Notes: Employment is defined as working at least quarter−time. The vertical line indicates the age cutoff for the 2007 reform.

the subsample of employed individuals, conditional on working at least a quarter of full-time. As expected, both (full-time equivalent) monthly wage and the work rate tend to increase in age. Older workers are also increasingly tenured, public-sector employed, higher educated and foreign-born. By comparing the two panels, we can deduce that, e.g., those with low education, women and foreign-born have lower employment than other groups.

Finally, we take a look at the evolution of employment and wages over time. Figure 3 gives the age distribution of employment before and after the 2007 payroll tax reduction. There are two things to notice in the figure. First, there is a relative employment increase for 20–25 year-olds in 2008. Second, within the target group, workers at age 21–22 seem to have gained the most. This suggests that the reduction did have an impact on employment, and that this impact decreases in age. However, we know that, in general, younger workers perform better in economic expansions, so the relative increase in employment may simply be a result of the growing Swedish economy in 2006–08. This problem is further discussed in the next section. In figure 4, we depict the corresponding distributional change in wages. As seen, there is no clear-cut evidence of larger wage growth for younger workers.

Figure 4: Average wage by age, 2006 and 2008 (log scale)



Notes: Sample conditional on working at least quarter−time. For those working less than full−time, wage is scaled to its full−time equivalent. The vertical line indicates the age cutoff for the 2007 reform.

5 Identification

5.1 Modelling the counterfactual outcome

We rely on the Difference-in-Differences (DiD) framework to estimate the effects of the payroll tax cuts. While, *prima facie*, using a regression discontinuity design on the 25–26 age threshold might appear attractive, it is clear from figure 3 that such a strategy is not viable. There are systematic discontinuities at each cohort boundary in 2006, before the tax reduction was implemented.[11]

In its simplest form, DiD uses the evolution of the control group over time as a measure of how the treatment group would have evolved, had the intervention not taken place. This results in the identifying assumption

$$E\left[y^0_{i,t} \mid \mathrm{Tr}=1\right] = E\left[y^0_{i,t} \mid \mathrm{Tr}=0\right] + \alpha, \tag{1}$$

where $y^0_{i,t}$ is the no-treatment outcome for individual i at time t. In other words, the counterfactual

[11]This pattern has its main cause in the fact that it is year of birth that determines when a child starts school in Sweden; see Fredriksson and Öckert (2014). With a DiD design, we assume that these cohort discontinuities are constant over time, for each age pair.

outcome of the treatment group is identical to the actual outcome of the control group, except for a constant α. Figure 5 demonstrates that, in the present context, this is too strong an assumption. Inspecting the evolution of employment in the period before the reform (2001–06), it is clear that individuals of different ages differ in the degree of employment cyclicality, with younger workers tending to display larger cyclical variations.[12] As 2007 coincided with an economic expansion, comparing, say, 20-year-olds to 26-year-olds would result in an upward-biased reform estimate: even in absence of a reform, a relative employment increase for 20-year-olds would have been expected solely due to this group's higher employment cyclicality. In addition to this systematic age heterogeneity, there are idiosyncratic differences between cohorts (e.g., due to temporary waves of immigration).

In order to model the counterfactual outcome of the treatment group we supplement the basic DiD model with a large number of covariates. The estimated specification is

$$y_{i,t} = \delta_t \cdot D(i,t) + \mathbf{x}_{i,t}'\boldsymbol{\beta} + \varepsilon_{i,t}, \tag{2}$$

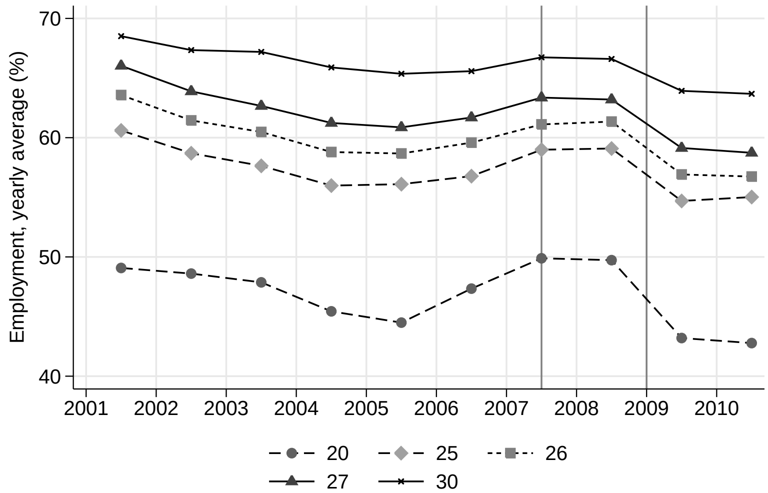
where $y_{i,t} \in [0,1]$ is average employment status in year t, $D(i,t)$ is a treatment indicator for individual i in year t, δ_t is the DiD estimate for year t, and $\mathbf{x}_{i,t}$ is a vector of control variables, capturing a multitude of factors that may influence the probability of being employed. These include dummy variables for year, age, county of birth (including indicator for foreign-born), gender, geography, and whether the parents immigrated into Sweden. For foreign-born, we also control for country of birth and years since immigration into Sweden.

5.2 Absolute versus relative effects

An implication of the DiD identifying assumption of parallel trends is that the control group must not be affected by the intervention. If such treatment spillovers exist, we will not measure the difference between the reform outcome and the counterfactual outcome, but the difference to the

[12] This heterogeneity is caused by, among other things, differences in labor market attachment, educational attainment and social situation. See Hoynes et al. (2012) for an extensive treatment of employment cyclicality for the U.S. labor market.

Figure 5: Employment trends for different age groups



Notes: Employment is defined as working at least quarter−time. The two vertical lines indicate the reform years.

control group deviation from its counterfactual outcome. In other words, we obtain a measure of the *relative* rather than the absolute effect of the reform. In the present case, there are strong reasons to suspect that the tax reduction had an indirect impact also on individuals not in the target group. The treatment spillover takes the form of substitution and scale effects. As a way of illustration, consider individuals at 25–26 years of age. The 2007 payroll tax reduction increases the cost of 26-year-old labor relative to 25-year-old labor. If firms consider 25-year-olds and 26-year-olds as substitute inputs they will, all else equal (i.e., holding output constant), lower demand for the latter group of workers, resulting in a negative substitution effect for 26-year-old labor. The magnitude of the negative substitution effect on non-treated individuals should depend on their similarity to individuals in the target group. Hence, the effect should decrease in age.

The scale effect tends to work in the opposite direction to the substitution effect. A factor input price drop results in a downward shift of the firms' cost functions, potentially causing them to expand output. Similar to income effects in consumer theory, the sign of the scale effect can be either positive or negative, but for normal factor inputs, demand is increasing in output. If employers prefer older, more experienced, workers, the scale effect increases in age. Nonetheless, this scale effect asymmetry, if it exists, is likely to be small, especially if we use treatment-control pairs that are close in age. Hence, the substitution effect bias is, arguably, the bigger problem.

5.3 Choice of comparison groups

The previous discussion suggests that there is an element of trade-off involved when choosing comparison groups: decreasing the age interval around the cutoff should get us closer to estimating a causal, albeit relative, treatment effect, but the estimate is unlikely to be generalizable to the target group as a whole. With this in mind, we evaluate the effects of the payroll tax reduction both for age-groups close to the cutoff and for 20–25 year-olds. The reason for excluding 19-year-olds is that they turn out to be substantially different in terms of employment cyclicality, thus invalidating the use of DiD. Most likely, this is explained by the fact that the majority of 19-year-olds are in their final year of high school for the first half of the year.

The parallel trends assumption is, by definition, not testable since it concerns counterfactual outcomes. A common convention is to consider the evolution of the treatment and control groups prior to the intervention, thus getting an indication on whether the assumption is likely to hold (or rather, when it is not likely to hold). While this procedure does not guarantee unbiased estimates, as is clear from the above discussion of treatment spillover effects, we consider parallel pre-treatment trends a minimal condition. This constrains us to use control group individuals close to the treatment cutoff, mainly 26-year-olds. As discussed above, these individuals are probably negatively affected by the reform and, thus, we interpret the estimations as upper bounds of the employment effect for the target group. As a special case, we consider individuals within a small bandwidth just around the treatment cutoff, comparing 25-year-olds born in January–March with 26-year-olds born in October–December. This specification has elements of a regression discontinuity design, but with controlling for the pre-reform discontinuity. While heterogeneous cyclicality should no longer be an issue, with comparison group so close in age, this comes at a cost: similar to RD designs in general, the estimates risk being only locally valid.

In theory, we should expect stronger treatment effects for younger workers since the remaining available treatment years (the treatment dose) is decreasing in age. Estimating effects for individuals close to the cutoff may, for this reason, underestimate the average treatment effect on the treated. Additionally, since the treatment and control groups are defined in terms of age groups, they are each year redefined in terms of cohorts. Consequently, an estimate based on single age groups is more sensitive to cohort heterogeneity, showing up as year shocks. In contrast, when using a treatment group of multiple ages, this heterogeneity is averaged out.[13] (Another way of dealing with this issue is to estimate pooled treatment effects for two years at a time, e.g., the 2007–08 effect. Such an approach averages out cohort offsets, but at a loss in temporal resolution.)

5.4 Repeated treatment and the 2009 extension

A difficulty with our method of evaluation is that, with time, it gets increasingly difficult to find individuals who have not been previously subjected to the payroll tax reduction. This makes it hard to identify the reform effect for the later years in our sample. Essentially, the problem of lagged treatment exists whenever employment spells extend from one year to the next. Figure 2 in section 3.2 illustrated how different cohorts are subjected to the payroll tax reductions. In 2007, the target group consists of individuals born 1982–88. Their natural control group consists of individuals that are slightly older, i.e., those born 1981. In 2008, individuals born 1983–89 are in the target group, and those born 1982 constitute the control group. Arguably, the employment estimate for 2007 is best identified since there is no earlier intervention, for any age group. Already in 2008, the control group may be affected by earlier treatment. For example, comparing 25-year-olds to 26 year-olds implies that our control group in 2008 (those born 1982) was in the target group the year before. One way to handle this is to use 27-year-olds instead of 26-year-olds as control, when possible. In the analysis below we experiment with altering the control group in this way.

Figure 2 also shows why it is difficult to evaluate the 2009 extension. Since 26-year-olds are included in the redefined target group, the youngest age group that can be used as a control group is now 27-year-olds, and they are not comparable—in terms of parallel pre-treatment trends—to

[13] Insofar as this cohort heterogeneity consists of compositional differences in dimensions that we observe, our control variables should take care of the problem. However, a *constant* offset for, say, the cohort of 25-year-olds in 2007 would bias the estimate of the reform effect. Cohort heterogeneity in the control group remains a potential problem since we, in most cases, cannot extend the age-interval upwards.

any age group below 24. We are thus restricted to studying the effects of the 2009 extension only for 24–26 year-olds. Those 24–25 years of age transition from 2007 treatment to 2009 treatment, while 26 year-olds transition from no treatment straight to 2009 treatment. Note, however, that for the 2009 extension we can only study individuals who have been previously treated, as apparent from figure 2.

In addition to these issues, the fact that the global financial crisis had its largest impact on Swedish employment in 2009–10—disproportionally affecting employment for younger workers—makes identification for these years even more difficult. When considering the 24–25 year-olds, the 2009 estimate will measure the impact of an extended reduction in the wake of the financial crisis. For 26-year-olds, we, correspondingly, get the effect of introducing a payroll tax reduction in the wake of an economic depression. Hence, both of these specifications could be seen as testing how the payroll tax reductions fare when labor market conditions worsen.

5.5 Estimating wage effects

The impact on employment depends on how much of the tax cut is shifted onto workers in the form of higher wages. In the long run, wages may adjust to counteract the effect of a payroll tax change. In the extreme case of full shifting, the payroll tax decrease will be fully cancelled out by wage increases, resulting in unchanged net labor costs for employers and, consequently, no employment effects. In the present case, with targeted reductions and a target group that has little attachment to the labor market, it is difficult, *ex ante*, to predict whether shifting will occur.[14]

Wage effects can appear through two channels: individual bargaining and union bargaining. In the latter case, there is a possibility that unions seek to make sure that all workers benefit so that the payroll tax reductions resulted in general shifting. This gives rise to a problem similar to when estimating employment effects: the $\boldsymbol{\delta}$ in equation 2 captures only the relative wage effect. However, the primary question we are interested in is not whether shifting occurred *per se*; rather, our focus is on whether relative wage increases around the cutoff can explain (the lack of) relative changes

[14]Some guidance may be found in Kolm (1998), who considers a two-sector (general equilibrium) model where market competitiveness differs between sectors, and where a general payroll tax cut would be fully shifted to workers. The model shows that taxing the less competitive sector more reduces unemployment.

in employment. Finally, it is important to stress that we only study the immediate impact on wages. If wage adjustments appear in the longer run, we will underestimate the long-term general equilibrium consequences of the payroll tax cuts.

6 Results

6.1 Main findings

Table 2 presents the main results for the 2007 reduction. The outcome variable is yearly average employment status, ranging from zero to one. All treatment effects are relative to the reference period 2001–04. The first two rows show whether the comparison groups move in parallel prior to the 2007 reform: significant pre-treatment effects for 2005 or 2006 would indicate that the control group is invalid.

The first column studies the effect at the treatment cutoff, comparing the three oldest birth month cohorts (born in January–March) of the 25-year-olds to the three youngest birth month cohorts (born in October–December) of the 26-year-olds. There is a statistically significant, albeit small, positive employment effect, both in 2007 and in 2008. The sudden change in relative employment at the cutoff is most likely caused by the reform since the point estimates for both pre-treatment years are insignificant and close to zero. From the local estimation we conclude that the lower payroll taxes increased the employment rate with 0.8 percentage points, corresponding to a rise in employment of around 1.4 percent.[15] Column 2 looks at the whole target group, excluding 19-year-olds. The treatment effect is substantially larger than in column 1: for 2007, the point estimate corresponds to a rise in employment of roughly 2.5 percent, while for 2008 the increase is around 1.8 percent. The larger effect for younger individuals is consistent with treatment dose effects: younger individuals have longer expected exposure to the reduced payroll tax. However, this difference may also depend on labor force composition. For example, if low-skilled jobs are affected more by lower payroll taxes and younger individuals to a larger extent are low-skilled, we would expect the treatment effect to decrease in age. As in column 1, the insignificant pre-treatment

[15]The percentage increase is relative to the counterfactual outcome. It is, thus, obtained as $\beta/(\overline{y}_{\mathrm{TG}} - \beta)$.

Table 2: Employment effects of the 2007 reduction, main results

	Alt. age interval		Alt. control group	
	Local	20–25 vs.26	24–25 vs.26	24–25 vs.27
DD 2005	0.001	–0.003’	0.001	–0.003
	(0.003)	(0.001)	(0.003)	(0.002)
DD 2006	0.000	0.002	–0.000	–0.000
	(0.004)	(0.002)	(0.001)	(0.002)
DD 2007	0.008**	0.014***	0.008**	0.005***
	(0.003)	(0.003)	(0.003)	(0.002)
DD 2008	0.008*	0.010***	0.006***	0.008***
	(0.003)	(0.003)	(0.002)	(0.001)
$\overline{y}_{\mathrm{TG}}$	0.63	0.58	0.61	0.61
N	419,153	6,015,905	2,588,746	2,606,207
R^2	0.12	0.10	0.11	0.12

Notes: ***$p < 0.1\%$, **$p < 1\%$, *$p < 5\%$, ’$p < 10$ %. Outcome is average employment status during the year (ranging from 0 to 1). ‘Local’ compares 25-year-olds born in Jan–Mar to 26-year-olds born in Oct–Dec. $\overline{y}_{\mathrm{TG}}$ denotes treatment group average employment in the treatment period. All treatment effects are relative to the reference period 2001–04. Fixed effects included for year and demographic characteristics. Standard errors are cluster-robust w.r.t. local labor markets.

estimates support a causal interpretation of the employment increase.[16]

We cannot rule out that the 2008 estimate in column 2 is downward biased due to the treatment effect the previous year (those 26 years of age in 2008 were treated in 2007). As mentioned in section 5.4, one way to handle this issue is to use 27-year-olds instead of 26-year-olds as the control group. Unfortunately, due to significant pre-treatment effects, we cannot include individuals older than 26 years in the control group when studying the whole target group. What we can do, however, is to look at 24–25 year-olds and alternate between using 26-year-olds and 27-year-olds as the group of comparison. The result of this exercise is presented in columns 3–4 of table 2. As is seen, the 2008 effect is slightly larger when using 27-year-olds; this could imply that specifications that use 26-year-olds as the control group suffer from biased estimates for years later than 2007. (We note that the difference between the 2008 point estimates is small, and so this may not be a crucial issue.)

As discussed in section 5, due to treatment spillovers in the control group, DiD measures the relative effect of the payroll tax reductions. While estimating the size of the treatment spillover is important in itself, we have previously explained why this cannot be done in any straightforward way using the method at hand. However, to get an idea of the magnitude of the substitution effect we again use the strategy of changing the control group, from 26-year-olds to 27-year-olds. We note that the magnitude of the negative substitution effect on non-treated individuals should decrease in age (since individuals just above the cutoff are closer substitutes to those in the target group). Hence, if the estimated treatment effect is larger when using 26-year-olds, this would indicate that substitution occurred. Comparing the 2007 point estimates in columns 3–4 of table 2 shows that the effect is slightly larger when using 26-year-olds, which is consistent with the existence of a substitution effect. We want to stress that we are reluctant to draw any decisive conclusions, since the effect, if it exists, appears to be small (at least in absolute terms). However, even indications that substitution occurred is an interesting finding, not the least from a policy perspective. It shows that the reform may have created jobs for one group of individuals at the expense of another. Note that we focus on the 2007 estimate in the discussion about substitution effects; as showed above,

[16]The 2005 estimate is significant at the 10%-level. There are no significant pre-treatment effects when considering slightly smaller age-intervals, as seen in columns 2–5 of table 3.

the 2008 estimate may suffer from a bias when 26-year-olds are used as controls.

Next, we examine age heterogeneity in more detail. Since we cannot compare single age groups (except the oldest ones) to any age groups above the cutoff (since we face significant pre-treatment effects), we use the strategy to expand the treatment group step by step. As shown in table 3, there are statistically significant treatment effects irrespectively of how we define the treatment group. In columns 1–5 the magnitude of the effect grows smoothly as we gradually include younger individuals; this is what would be expected considering that the period of remaining treatment is decreasing in age. In the last column, however, the effect appears to decrease again, especially if we consider year 2008. A possible explanation is that labor force participation is decreasing for the youngest individuals, which means that a large number of 20-year-olds are not, in practice, eligible for the payroll tax reduction. Another interpretation is that not even the substantially higher treatment dose—20-year-olds have a lower relative price for 5 years—can compensate for the lower expected productivity.

For any specific year in table 3, the reported treatment effect estimate is the sum of the treatment effect for the treatment group and the negative substitution effect for the control group. Consequently, we can use the 25–26 estimates in column 1 as an upper bound for the negative substitution effect for the 26-year-olds, and, hence, as an upper bound for the substitution effect bias affecting the other estimations. (Since we use the same control group when going from 25 year-olds to 20–25 year-olds the substitution effect bias is held constant.) This implies that the absolute employment increase for 20–25 year-olds is *at least* 0.8 percentage points in 2007 and 0.6 percentage points in 2008.

In 2009, the Swedish economy was hit by the financial crisis (as is evident from figure 5). By considering the 2009–10 time period we can thus examine whether reduced payroll taxes counteract the negative effects of an economic downturn. From 2009 and onwards there were two slight changes in the original payroll tax reform. First, the 2007 target group was subjected to an additional five percentage points cut in the tax rate. Second, 26-year-olds, who were not previously included in the target group, were now also subjected to the reduced payroll taxes. Table 4 shows yearly treatment effects for different age groups up until 2010. As 26-year-olds are part of the target group from

Table 3: Employment effects of the 2007 reduction, age heterogeneity

	25 vs.26 (1)	24–25 vs.26 (2)	23–25 vs.26 (3)	22–25 vs.26 (4)	21–25 vs.26 (5)	20–25 vs.26 (6)
DD 2005	0.002	0.001	0.000	0.001	0.000	–0.003’
	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
DD 2006	–0.002’	–0.000	0.001	0.003	0.003	0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
DD 2007	0.006**	0.008**	0.011***	0.014***	0.015***	0.014***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
DD 2008	0.004*	0.006***	0.009***	0.013***	0.013***	0.010***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
$\overline{y}_{\mathrm{TG}}$	0.63	0.61	0.60	0.60	0.59	0.58
N	1,735,836	2,588,746	3,438,874	4,291,748	5,148,083	6,015,905
R^2	0.11	0.11	0.11	0.10	0.10	0.10

Notes: ***$p < 0.1\%$, **$p < 1\%$, *$p < 5\%$, ’$p < 10$ %. Outcome is average employment status during the year (ranging from 0 to 1). $\overline{y}_{\mathrm{TG}}$ denotes treatment group average employment in the treatment period. All treatment effects are relative to the reference period 2001–04. Fixed effects included for year and demographic characteristics. Standard errors are cluster-robust w.r.t. local labor markets.

2009 and onwards, we have switched to using 27-year-olds as the frame of reference.

Column 1 studies 24–25 year-olds, who transition from 2007 treatment to 2009 treatment (as before, these are the youngest age groups that we can consider). We first confirm what we saw in table 2, column 4: there are significant employment effects both in 2007 and 2008, but not before. In contrast to the two years when the economy was expanding, the recession years 2009–10 display a somewhat jumpy pattern. The point estimate for 2009 is small and insignificant at conventional levels, whereas in 2010, when employment levels no longer fell dramatically, the effect is again significant and similar in magnitude to 2007–08. Next, in column 2 of table 4, we study 26-year-olds—the age group that were subjected to reduced payroll taxes for the first time in 2009. Strikingly, for this age group there is no effect of the lower payroll taxes in any of the recession years 2009–10 (there is even a small negative effect in 2009).[17] It seems reasonable to interpret the results in table 4 as evidence against any additional employment effect of the 2009 extended reduction; if anything, the effect is even lower than in the preceding years. This finding is important as it suggests that (substantial) payroll tax cuts are even less effective in economic downturns.

Column 2 of table 4 points to another interesting finding. From a welfare perspective, it is important to understand how lasting the effect is for an individual who is no longer eligible for the lower payroll tax, but who was previously treated. For 26-year-olds we do not expect any impact in 2007 since they were not part of the target group. In 2008, however, we may expect an impact depending on whether an effect persists to years without treatment. We have included column 3 to make comparisons easier; as we know from before, 25-year-olds display significant effects both in 2007 and 2008. Comparing the 2007 estimate in column 3 to the 2008 estimate in column 2 shows that the treatment effect vanishes quickly when an individual transitions from treatment to no treatment. (We note that 26-year-olds are treated in 2009, as disussed above.) Ultimately, this shows that even though the reform created a relative price wedge that induced employers to hire, or to keep, a young worker, it did not lead to any permanent increase in the likelihood that this individual is employed.

Our employment measure uses the quarter of the income cutoffs to arrive at a measure of working

[17] There is a significant, but small, pre-treatment effect in 2005. However, both in 2006 and in 2007 there are no significant pre-treatment effects.

at least 25 percent of full-time (for a particular month). We have also experimented with alternative definitions of employment. A stricter definition—going from working at least 25 percent to working at least 50 percent of full-time—produces somewhat smaller treatment effects, while, on the other hand, relaxing the employment restriction to 10 percent of full-time does not change the estimates. The latter suggests that it is not the case that we fail to account for part of the employment effect by choosing a too strict employment definition. Using an outcome measure of full-time employment, or less than 10 percent of full-time, is not viable since we then face significant pre-treatment effects.[18] In summary, there seem to have been positive, but small, employment effects of the 2007 payroll tax reduction. This holds irrespective of whether we study a small interval around the treatment cutoff, or examine the whole group of 20–25 year-olds. For the 2009 extended reduction, there is no evidence of any additional effect.

6.2 Treatment effect heterogeneity

We next turn to the subsample of young immigrants, in columns 1–2 of table 5. This group, which constituted about 15 percent of the age group 20–25 in 2007–08, is characterized by weak attachment to the Swedish labor market. Their employment rate is about 20 percentage points lower than for the whole population of young workers, as reported in the bottom rows of tables 2 and 5. Strikingly, there is no evidence that the payroll tax cut had any impact at all for young foreign-born. Importantly, the lack of treatment effects is not the result of noisy estimates due to a smaller number of observations.[19]

In theory, an explanation for the small general employment effects could be labor supply constraints. For the age group 20–25, many are taking part in higher education, and it is perhaps not reasonable to expect a strong employment response for this group. We examine this hypothesis by studying previously unemployed 25–26 year-olds—defined here as those individuals who were not taking part in education but registered unemployed at the unemployment office for at least 100 days

[18] The results for the other definitions of employment are available from the authors upon request.

[19] Since the sample of foreign-born is far from homogenous, we have also used finer subdivisions of region of birth, as well as disregarding newly arrived immigrants. Eastern Europeans is the only group for which we find a positive effect; the magnitude is similar to that of Swedish-born. These results are available from the authors upon request.

Table 4: Employment effects of the 2009 extension

	24–25 vs. 27	26 vs. 27	25 vs. 26
DD 2005	–0.003	–0.004**	0.002
	(0.002)	(0.002)	(0.002)
DD 2006	–0.000	0.000	–0.002’
	(0.002)	(0.002)	(0.001)
DD 2007	0.005***	–0.003	0.006**
	(0.002)	(0.002)	(0.002)
DD 2008	0.008***	0.002	0.004*
	(0.002)	(0.001)	(0.002)
DD 2009	0.002	–0.003*	
	(0.002)	(0.001)	
DD 2010	0.009***	0.001	
	(0.002)	(0.002)	
$\overline{y}_{\mathrm{TG}}$	0.59	0.63	0.63
N	3,305,579	2,224,207	1,735,836
R^2	0.12	0.14	0.11

Notes: ***$p < 0.1\%$, **$p < 1\%$, *$p < 5\%$, ’$p < 10$ %. Outcome is average employment status during the year (ranging from 0 to 1), $\overline{y}_{\mathrm{TG}}$ denotes treatment group average employment in the treatment period. All treatment effects are relative to the reference period 2001–04. Fixed effects included for year and demographic characteristics. Standard errors are cluster-robust w.r.t. local labor markets.

during the previous year. (In 2007, this group amounted to around 38 percent of all 25–26 year-old registered, and around 9 percent of the full cohorts.) For this group, labor supply constraints should be less of a problem: by definition they are not taking part in education, and the fact that these individuals are attending the unemployment office at least signals a willingness to take a job. As column 3 of table 5 shows, there is no indication that the effect for unemployed 25-year-olds were larger than in the general case. While this result does not completely rule out the labor supply story, it indicates that labor demand is the more important factor.

Table 5: Employment effects for subgroups

	Foreign-born		Unempl.
	25 vs.26	20–25 vs.26	25 vs.26
DD 2005	0.002	–0.001	–0.000
	(0.003)	(0.001)	(0.005)
DD 2006	–0.002	–0.001	–0.005
	(0.003)	(0.003)	(0.005)
DD 2007	0.003	0.005’	0.007
	(0.003)	(0.003)	(0.007)
DD 2008	–0.006’	–0.002	0.002
	(0.004)	(0.004)	(0.007)
$\overline{y}_{\text{TG}}$	0.39	0.35	0.45
N	291,125	890,911	153,931
R^2	0.19	0.18	0.11

Notes: ***$p < 0.1\%$, **$p < 1\%$, *$p < 5\%$, ’$p < 10$ %. Control variables include region of birth, year since immigration into Sweden, among others. Previously unemployed defined as having been registered at the unemployment office at least 100 days during the previous year. See also notes for table 2.

6.3 Wage effects

We next examine whether part of the payroll tax cut was passed on to employees as higher wages. The outcome measure is now the log of monthly, full-time equivalent, wage for those employed at least quarter-time (in symmetry with our main employment definition used above). Table 6 gives the impacts of both the 2007 initial cut and the 2009 extension. Starting with the 2007 reduction, there is no effect around the cut-off; the point estimates for 25-year-olds are small in economic terms, and insignificant. This implies that, close to the cutoff, there are significant employment effects but no wage effects. For 20–25 year-olds there is, however, a small relative wage increase, slightly above one percent both in 2007 and in 2008, which could indicate that some of the younger workers of the target group did take home a small fraction of the tax cut given to employers. Notably, the wage increase shows up already in 2007.[20] Comparing 24–25 year-olds to 27-year-olds allows us to study the evolution of wages into the 2009 extension. Since there is no additional wage effect in 2009–10 we conclude that wages did not adjust more in the somewhat longer run.

Understanding these wage effects requires making a few observations. To start with, there are indications that the unions and the employer organizations in 2007 agreed on letting minimum wages increase faster than general wages (National Mediation Office, 2007). Thus, we are potentially picking up negotiated minimum wage increases. It is, however, an open question whether these increases were the result of the reform or part of a long-term trend. (As mentioned in section 3, wages were renegotiated at the central level just after the passing of the 2007 reduction in the parliament. Hence, since both the unions and the employer organizations were aware of the forthcoming tax reduction there is, in principle, a possibility that the wage response came *before* the actual implementation.) What speaks against the minimum wage increase explanation is the evidence of wage effects even for age groups that typically have wages strictly above the minimum wage level.[21] Another potential explanation is that shifting works through individual wage bargaining; such an impact, if it exists, is likely to be more immediate than union-negotiated wage increases. Having said this, we conclude that given the small size of the wage increase, shifting cannot by itself explain the modest employment effects we have found.

7 Cost-benefit analysis

In the following, we present some further metrics for evaluating the payroll tax reduction, with emphasis on 2007 where we have the most credible identification. It is important to stress that these derived measures are likely to be overly optimistic. First, the substitution effect bias causes us to overestimate the treatment effect and, consequently, to overestimate the demand elasticity and underestimate the cost per job. Second, it is by no means clear that the target group employment

[20]For each of the age groups that we consider, we have tested for heterogeneity with respect to private or public sector, for blue collar or white collar workers, and for new or tenured employees. The results for these subgroups are similar to the general case.

[21]Forslund et al. (2012) report that young workers' wages in the private sector are often higher than the negotiated minimum wages, even for workers as young as 19 years old.

Table 6: Wage effects of the 2007 reduction and the 2009 extension

	25 vs. 26	20–25 vs.26	24–25 vs.27
DD 2005	–0.006	0.004	0.000
	(0.004)	(0.003)	(0.001)
DD 2006	–0.005	0.006	–0.003
	(0.004)	(0.004)	(0.003)
DD 2007	0.004	0.012***	0.009*
	(0.003)	(0.003)	(0.004)
DD 2008	0.004	0.013*	0.014*
	(0.005)	(0.005)	(0.006)
DD 2009			0.011***
			(0.003)
DD 2010			0.009*
			(0.003)
$\overline{y}_{\text{TG}}$	21,670	20,540	22,030
N	537, 619	1,485,391	981,757
R^2	0.25	0.25	0.27

Notes: $^{***}p < 0.1\%$, $^{**}p < 1\%$, $^{*}p < 5\%$, $'p < 10$ %. Outcome is the log of monthly full-time equivalent wage (trun- cated below to 0), $\overline{y}_{\text{TG}}$ denotes treatment group average outcome in the treatment period, in non-log form. All treatment effects are relative to the reference period 2001–04. Fixed effects included for year and demographic characteristics. Standard errors are cluster-robust w.r.t. local labor markets.

increase reflects a *net* increase of jobs in the economy. Rather, a part of this increase may be at the expense of older workers in the labor force. Although this will not affect the elasticity estimate—which is defined as being with respect to *young* labor—it will further bias the measure of cost per job, as job losses for older workers are not taken into account. This is discussed further in section 8.

7.1 Elasticities

We can combine the employment and wage estimates to get estimates of the elasticity of demand for young workers with respect to labor costs. For 20–25 year-olds, the 2007 employment increase is 2.5 percent, and the 2007 wage increase is 1.2 percent. Hence, we arrive at a labor demand elasticity at about -0.37.[22] Table 7 shows the corresponding figure for each age group.[23] Although some of these numbers may appear small, previous literature typically finds no employment effects of targeted payroll tax reductions. In particular, employment was unaffected by regional reductions in the Nordic countries, and by reductions targeted at the employers of older, full-time, low-wage workers in Finland (see Bohm and Lind, 1993; Bennmarker et al., 2009; Korkeamäki and Uusitalo, 2009; Huttunen et al., 2013).

7.2 How much money was spent on each job?

The gross cost of the payroll tax reductions—the sum of foregone payroll taxes, disregarding potentially increased revenues due to, e.g., higher profits—can be straightforwardly calculated since total taxable income is available to us in the tax registers. Figure 6 shows the cost broken down by age for the years 2008 and 2009, thus demonstrating the effect of the 2009 extension. The figure illustrates that incomes are markedly higher for the older individuals of the target group, as they

[22] Note that the employment effect is estimated in absolute numbers while the wage estimate is in log form. In addition to wage level and payroll tax, labor cost also includes a union negotiated fee at around 10 percent. Thus, labor demand elasticity is obtained as

$$\epsilon = \frac{\beta_{\text{empl}}/(\overline{empl}_{\text{TG}} - \beta_{\text{empl}})}{(e^{\beta_{\text{wage}}} - 1) - 0.111/(1 + 0.3242 + 0.10)}.$$

[23] The point estimates used to calculate the figures for 20–21 year-olds in table 7 come from regressions without pre-treatment years. This means that, for the youngest age groups, we are less convinced that the figures reflect the true elasticities. All elasticities are calculated using the wage estimate for 20–25 year-olds.

Table 7: Elasticities for 2007, by age

	ϵ
Age 20	-0.27
Age 21	-0.54
Age 22	-0.57
Age 23	-0.39
Age 24	-0.26
Age 25	-0.15
Age 20–25	-0.37

Notes: All elasticities are calculated using the wage estimate for 20–25-year-olds.

both have higher average wages and work more hours. As a consequence, the cost of the reductions increases in age. The figure also shows that the cost increased dramatically in 2009, by simultaneously increasing the size of the reduction and targeting a larger age group. The total gross cost increased from SEK 9.9 billion ($1.1 billion) in 2008 to 17 billion ($2.1 billion) in 2009. These high numbers reflect the fact that all employments were subsidized, not only new ones.

We can also deduce the total number of new jobs created each year by the payroll tax reduction. For 20-25 year-olds, a 95 percent confidence interval gives an estimate of 5,300 to 13,300 new jobs (with a point estimate of 9,300). In combination with the gross cost, we now get an estimate of the cost per created job, depicted in figure 7. For 20–25 year-olds, the cost for each job is SEK 0.8 to 1.8 million ($100,000 to $225,000), with a point estimate at SEK 1.1 million ($140,000). Notably, this is more than four times the hiring cost, assuming that the created jobs had the average annual income for this group.[24] Since the gross cost increases in age and, additionally, the number of new jobs decreases in age, it is not surprising that the cost per job soar as we move closer to the treatment age cutoff. For 25-year olds, the cost amounts to more than eight times the average hiring cost.

Finally, we note that these numbers apply only to the first tax reduction. In 2009, the payroll

[24]When calculating the hiring cost, we take the average income of those employed at least quarter-time, adding the cost for payroll taxes and the union-negotiated fee of (around) ten percent (in total 42.42 percent). We have chosen not to consider payroll taxes as government revenues, as they are, mostly, linked to insurances and social benefits for the employed. However, the figures change only marginally if we instead subtract increased payroll tax revenues from the cost figures (taking into account both new jobs and the wage increase associated with the tax reduction).

Figure 6: Gross cost per age group, 2008 and 2009

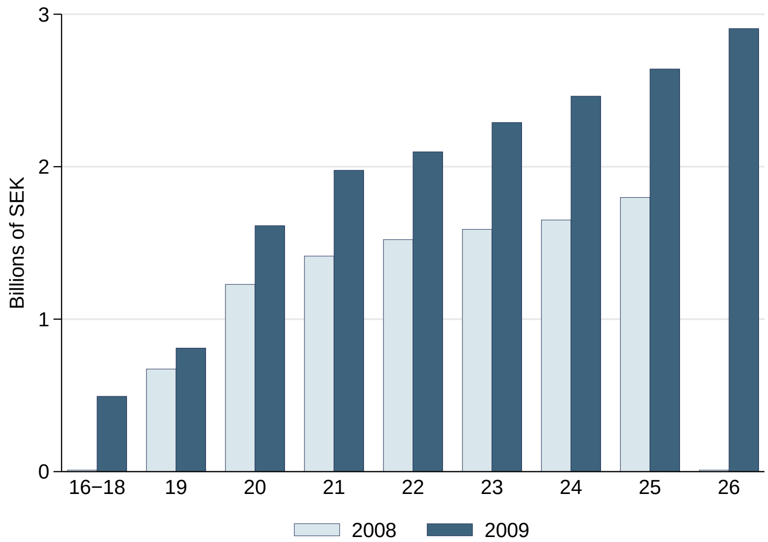
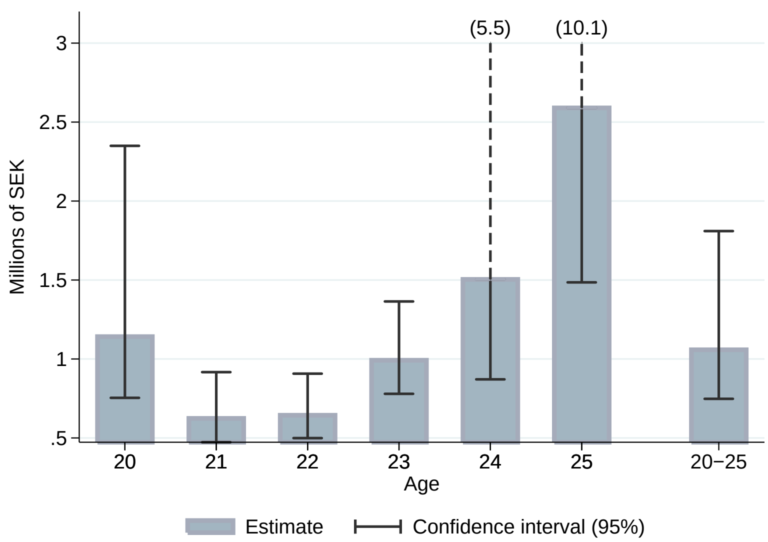


Figure 7: Estimated cost per new job of the 2007 reduction



tax reduction was both increased in magnitude and extended to encompass all individuals under 26 years of age. Although we have no useful employment estimates for this period, we know that the gross cost almost doubled in 2009. Thus, if our results are indicative also for the employment response of the 2009 extension, the cost per job is likely to be significantly higher for this period than for the 2007 original tax reduction.

8 Discussion

The previous sections have painted a picture of the 2007 and 2009 payroll tax cuts as being rather unsuccessful—the impact on youth employment was small, and the cost per created job was high. This may seem puzzling at first glance: wages should be rigid in the short run, so we might at least have expected a temporary employment boost. Indeed, the wage regressions demonstrate that there were no extensive wage adaptations that could explain the meager impact on employment. This raises the question of why employers do not increase their hiring of young workers, despite the latter now being significantly less expensive. In discussing potential answers to this question, we will consider labor supply constraints and labor demand constraints, in that order.

It is, in principle, possible that the lack of employment response is caused by low labor supply. There are many alternatives to employment for young individuals in Sweden. Many are taking part in higher education, others spend a couple of years after high school travelling the world. It is also possible that some of those who are formally applying for a job are actually quite satisfied with the comfortable life of receiving unemployment benefits while living with their parents, thus stifling the willingness to work. These speculations are, to some extent, tested in our regressions for the subsample of previously unemployed 25-year-olds. By restricting the sample to those registered at the unemployment office, we disregard both the unemployed students and the globe trotters. While the fundamental issue of weak economic incentives remain, we should diminish its importance by studying 25-year-olds—for individuals at this age there is a strong social stigma both of being unemployed and of living with one's parents (thus the economic incentives kick in stronger as well). The null effect for the unemployed in table 5 indicates that labor supply is not the main problem.

We thus conclude that the weak employment response is more likely to be a consequence of low demand elasticity.

Turning to labor demand, we discuss a number of alternative explanations. First, it is unlikely that employers were unaware of the new rules since the reform was covered rather extensively in the media, both when it was ratified and later on. (The payroll tax reductions were also criticized by the political opposition in Sweden and, therefore, rather intensely debated.) It is also unlikely that employers were reluctant to take any action in the short run because they were uncertain about how persistent the new rules would be. The reform was implemented shortly after the 2006 elections, meaning that employers should have anticipated the new rules to be in place for at least one length of office, which is four years in Sweden. To be sure, the extension of the payroll tax reductions in 2009 should clearly signal that this was not a temporary policy, but even here, we find small, or no effects.

Another possible explanation is linked to short-term capital rigidity. Since increasing output may require long-run capital investment, the scale effects are not allowed to work to its full extent in the short run. Thus, if firms were capacity constrained when the lower taxes were implemented, they could not immediately make the capital investments to accommodate more labor. The fact that the 2007 reduction was implemented in a booming economy speaks for this explanation. But this explanation is, at the very most, plausible only for the very short run—if this were true, we would see increasing effects at least at the end of the period under study. Furthermore, Skedinger (2014) finds small effects also in the Swedish retail industry, where firms should be less capacity constrained. Indeed, for this industry it is during a boom that employers should be most willing to hire young workers, also in the short run.

A third possible explanation for the lack of large employment effects is that the wage cost for the typical young worker is too high in relation to her productivity, even after the tax cut. That is, the labor cost reduction does not compensate for the risk premium of hiring a young, untrained, and unexperienced worker. This corresponds to a situation where, for many firms, factor demand for young labor is at a corner solution, at zero demand. In such a scenario, any cost-reducing measure that does not push labor costs below the hiring threshold will have zero effect on the firm's labor demand—i.e., the demand elasticity will be locally zero. This explanation is corroborated by the fact that for previously unemployed, where labor costs should correspond even less to productivity, we find no effects at all.

It is important to stress that the figures reported in this study may not reflect net effects on the labor market as a whole. In section 5 we describe how control group substitution induces a substitution effect bias in all of our estimates. But negative substitution is likely to affect also older workers in the economy—if they are similar to the target group in terms of labor market characteristics. Thus, the larger employment increase for 20–25 year-olds, compared to 25-year-olds, can be the result of increased substitution with older workers. In other words, while we do find an *absolute* employment increase for the target group, this may not reflect a *net* increase in the economy as a whole. The share of the employment increase that is associated with a net creation of jobs corresponds to the relative share of the scale effect, which, unfortunately, we cannot quantify. However, it should be noted that if factor inputs are close to perfect substitutes (e.g., low-skilled labor at different ages), there may be large substitution effects even though the scale effect is small. As a consequence, it is likely that our estimates overestimate the number of new jobs created: partly because the estimates overestimate the actual employment increase (due to control group treatment spillover), partly because the actual employment increase may have been at the expense of older workers in the economy. Similarly, the estimated cost per job, reported in the previous section, is bound to underestimate the true cost.

9 Conclusion

We study whether large-scale payroll tax reductions for employers of young workers is an effective means to raise youth employment. In 2007, the Swedish employer-paid payroll tax was cut on a large scale for young workers, substantially reducing labor costs for this group. We estimate the short-run effect of this tax cut to be, at most, an employment increase at 2.5 percent. We find no additional employment effect of an extension of the original reductions, implemented in 2009. Shifting of the tax cut onto workers in the form of higher wages cannot explain the modest employment effect: the size of the wage adjustments in the wake of the reform is small, at roughly one percent.

The employment and wage estimates in combination imply that the short-run elasticity of demand for young workers in Sweden is at around -0.37. Using a different metric, the estimated cost per created job for 20–25 year-olds is at more than four times the cost of directly hiring workers at the average wage. We conclude that targeted payroll tax cuts are an expensive way to boost employment for young individuals.

References

Anderson, P. M. and B. D. Meyer (1997). The effects of firm specific taxes and government mandates with an application to the u.s. unemployment insurance program. *Journal of Public Economics 65*(2), 119–145.

Anderson, P. M. and B. D. Meyer (2000). The effects of the unemployment insurance payroll tax on wages, employment, claims and denials. *Journal of Public Economics 78*(1-2), 81–106.

Bennmarker, H., E. Mellander, and B. Öckert (2009). Do regional payroll tax reductions boost employment? *Labour Economics 16*(5), 480–489.

Bohm, P. and H. Lind (1993). Policy evaluation quality : A quasi-experimental study of regional employment subsidies in sweden. *Regional Science and Urban Economics 23*(1), 51–65.

Edmark, K., C.-Y. Liang, E. Mörk, and H. Selin (2012). Evaluation of the swedish earned income tax credit. Working Paper Series 2012:1, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Forslund, A., L. Hensvik, O. Nordsröm Skans, and A. Westerberg (2012). Kollektivavtalen och ungdomarnas faktiska begynnelselöner. Working Paper Series 2012:19, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Fredriksson, P. and B. Öckert (2014). Life-cycle Effects of Age at School Start. *The Economic Journal* (forthcoming).

Fredriksson, P. and R. H. Topel (2010). Wage determination and employment in sweden since the early 1990s: Wage formation in a new setting. In R. B. Freeman, B. Swedenborg, and R. H. Topel (Eds.), *Reforming the welfare state : recovery and beyond in Sweden*, pp. 540–559. Chicago: University of Chicago Press.

Gregg, P. (2001). The impact of youth unemployment on adult unemployment in the ncds. *The Economic Journal 111*(475), 626–653.

Gregg, P. and E. Tominey (2005). The wage scar from male youth unemployment. *Labour Economics 12*(4), 487 – 509.

Gruber, J. (1997). The incidence of payroll taxation: Evidence from chile. *Journal of Labor Economics 15*(3), S72–101.

Hoynes, H. W., D. L. Miller, and J. Schaller (2012). Who suffers during recessions? Working Paper 17951, National Bureau of Economic Research.

Huttunen, K., J. Pirttilä, and R. Uusitalo (2013). The employment effects of low-wage subsidies. *Journal of Public Economics 97*(0), 49 – 60.

ILO (1983). Thirteenth International Conference of Labour Statisticians, Resolution Concerning Statistics of the Economically Active Population, Employment, Unemployment and Underemployment. *Bulletin of Labour Statistics* (1983-3), xi–xv.

Kolm, A.-S. (1998). Differentiated payroll taxes, unemployment, and welfare. *Journal of Public Economics 70*(2), 255 – 271.

Korkeamäki, O. and R. Uusitalo (2009). Employment and wage effects of a payroll-tax cut – evidence from a regional experiment. *International Tax and Public Finance 16*, 753–772.

Kramarz, F. and T. Philippon (2001). The impact of differential payroll tax subsidies on minimum wage employment. *Journal of Public Economics 82*(1), 115–146.

Murphy, K. J. (2007). The impact of unemployment insurance taxes on wages. *Labour Economics 14*(3), 457–484.

National Mediation Office (2007). *Avtalsrörelsen och lönebildningen år 2007.* Medlingsinstitutet, Stockholm.

Nordström Skans, O. (2004). Scarring effects of the first labour market experience: A sibling based analysis. Working Paper Series 2004:14, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Skedinger, P. (2012). Tudelad trygghet. In A. Teodorescu and L.-O. Pettersson (Eds.), *Jobben kommer och går : behovet av trygghet består*, pp. 114–135. Stockholm: Ekerlid.

Skedinger, P. (2014). Effects of Payroll Tax Cuts for Young Workers. *Nordic Economic Policy Review* (forthcoming).

Statistics Sweden (2014). *Arbetskraftundersökningarna*. SCB, Stockholm.

Effects of Taxes on Youth Self-Employment and Income*

Johan Egebark†

Abstract

I study the link between taxes and youth self-employment. I make use of a Swedish reform, implemented in 2007–09, which suddenly made the payroll tax and the self-employment tax vary by age. The results suggest that youth self-employment is insensitive to tax reductions, both in the short run and in the somewhat longer run. I also study the effect of the tax reductions on income. For those that are defined as self-employed, I find positive effects on income from self-employment, and negative effects on income from wage employment. This finding suggests that the lower taxes caused the self-employed to reallocate time from employment to self-employment.

Key words: Youth unemployment; Self-employment tax; Tax subsidy; Self-employment

JEL classification: H25, H32, J23, J38, J68

*I wish to thank Peter Fredriksson, Helena Holmlund, Niklas Kaunitz and Jonas Vlachos for valuable comments. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. All errors are my own.

†Department of Economics, Stockholm University, and the Research Institute of Industrial Economics (IFN). E-mail: johan.egebark@ne.su.se

1 Introduction

High and persistent youth unemployment is a major concern for many developed economies. Different policies have been tested to address this problem. In the 1990's the UK government launched large-scale active labor market programs (the *New Deal for the Young and Unemployed*) to improve young individuals' labor market opportunities (Blundell et al., 2004). More recently, as a response to the 2008 financial crisis, many countries, including France and Spain, have initiated different types of targeted hiring credits in attempts to boost employment among the young (Cahuc et al., 2014; Ferran, 2015).

High youth unemployment could reflect the fact that young individuals have few options in the formal sector, due to for example their lack of work experience or social connections. This suggests that one way for them to exit unemployment is to start their own business. In fact, means to stimulate self-employment is increasingly seen as part of a strategy to handle the youth employment challenge. The European Commission's *Youth Employment Package*, launched in 2012, states explicitly that the member states should support job creation by "promoting and supporting self-employment /.../ and business start-ups" and should focus its efforts on "business start-ups by unemployed and people from disadvantaged groups" (European Comission, 2013). One concrete example comes from Spain, where the government recently launched new initiatives to increase self-employment rates among young adults. Measures include lower social contributions and the possibility of extending unemployment benefits for young people that register a business.

Despite the potential role that self-employment could play, there are few studies on how to increase the number of businesses run by young individuals. OECD has suggested two main interventions: entrepreneurship education and financial support. However, since there are basically no (credible) evaluations of the effectiveness of these policies, it is difficult for policymakers to know what approaches actually work (OECD, 2012, 2013).[1]

In this paper, I provide hard evidence of whether reducing taxes is an effective way to increase

[1] There are studies on the effects of entrepreneurship education (see, e.g., Oosterbeek et al., 2010). However, these studies do not focus on disadvantaged groups: participants are often college graduates since the evaluated training is part of some higher education. Furthermore, as it is often the participants' intentions that are studied, the findings do not say whether training actually leads to self-employment.

youth self-employment. I make use of a Swedish reform, implemented in two steps during 2007–09, that introduced substantial variation in tax rates across age groups. In July, 2007, the employer-paid payroll tax was cut by 11 percentage points for workers who, at the start of the year, had turned 18 but not 25 years of age. These age groups were, at the same time, allowed a 10 percentage points reduction in the self-employment tax. After 18 months, the initial reduction was modified: both the payroll tax and the self-employment tax was cut additionally, and the target group was extended so that the reductions now encompassed all individuals who at the start of the year had not yet turned 26 years of age. The two reductions suddenly made taxes vary across cohorts, and, hence, offer a good opportunity to study the causal effects of taxation on a young person's decision to run a business.

I use Differences-in-Differences (DiD) to identify the effect of the tax changes, contrasting individuals below the treatment defining age cutoff to those just above. I consider the effects on the (overall) self-employment rate, and on transitions between self-employment and wage employment (i.e., on occupational choice). By studying each of the two reductions separately, I examine whether the effect varies depending on the state of the economy. Furthermore, by using different age intervals I am able to uncover whether the effect varies across ages.

The results suggest that youth self-employment is insensitive to tax changes. Both the 2007 cut and the 2009 cut left self-employment completely unaffected. The lack of treatment effects is precisely estimated and is robust to a battery of sensitivity tests. For example, none of the subgroups that I consider—e.g., men, women, natives, foreign-born, or those with vocational training—display any effects. The fact that there was no impact in 2009–10, i.e. in the midst of recession, suggests that (large) tax cuts have no role to play even in times of economic slowdowns.

I proceed by looking at intensive margin responses. I first show that for the young self-employed that faced the lower tax rate—essentially because they were lucky—income from self-employment increased by up to 20 percent on average. I then compare the estimated income effect to the predicted mechanical effect (due to the lower tax rate) and find that, in each year 2007–09, the estimated effect is greater.[2] I argue that a potential explanation for this pattern is that the tax

[2]The mechanical effect is the positive effect that exists irrespective of behavioral adjustments (Chetty et al., 2013).

cut caused self-employed individuals to allocate more time to self-employment—either by reducing time in leisure or in wage employment. I find some support for the reallocation of time explanation: for those that are defined as self-employed, income from regular employment decreases due to the tax cut. Since the income effect is temporary, however, I conclude that the reallocation of working hours did not pertain to years when an individual no longer faces the lower tax rate. While intensive margin responses to tax changes have proven difficult to detect in the past, recent work shows that such adjustments can be substantial (Chetty et al., 2013).

The rest of the paper is organized as follows. Sections 2 and 3 offer some background. Section 4 briefly describes the conceptual framework. Sections 5 and 6 handle the data and the identification strategy. Section 7 gives the results and section 8 concludes.

2 Institutional framework

2.1 Youth unemployment in Sweden

Official records show that youth unemployment in Sweden is currently high. Unemployment for 15–25 year-olds was roughly at 24 percent in 2013, which is three times higher than overall unemployment (Statistics Sweden, 2014). In 2007 and 2008, which are the years that I mainly focus on in this study, youth unemployment was somewhat lower, at around 20 percent. In 2009, when the Swedish economy was fully hit by the financial crisis, it increased to 25 percent.

It is sometimes argued that these (official) figures exaggerate the problem of youth unemployment in Sweden, mainly due to the fact that a large number of the unemployed participate in different types of education. Excluding those who study full-time lowers unemployment for 15–25 year-olds to about 12 percent in 2013. However, it is not obvious that this adjustment makes sense: many might chose to study since it is difficult to find a job, even though they rather would be working.

I complement these figures with two other measures to provide some further understanding of the problem in the Swedish case. First, about 10 percent of all 20–24 year-olds were not employed and not in any education or training in 2013 (i.e., they belong to the so called NEET category). In

2007–08, the corresponding figure was 12 percent, and in 2009 it was roughly 13 percent (Statistics Sweden, 2014). A second measure looks at registrations at the unemployment office. The data that I use for the analysis below contains yearly information on job search activity, and so I can observe those that are registered as looking for a job. 21 percent of all 20–24 year-olds were registered at the unemployment office at some point during 2007–08, and 8 percent were registered for more than 100 days. During the recession year 2009, these figures increased to 24 percent and 12 percent, respectively.

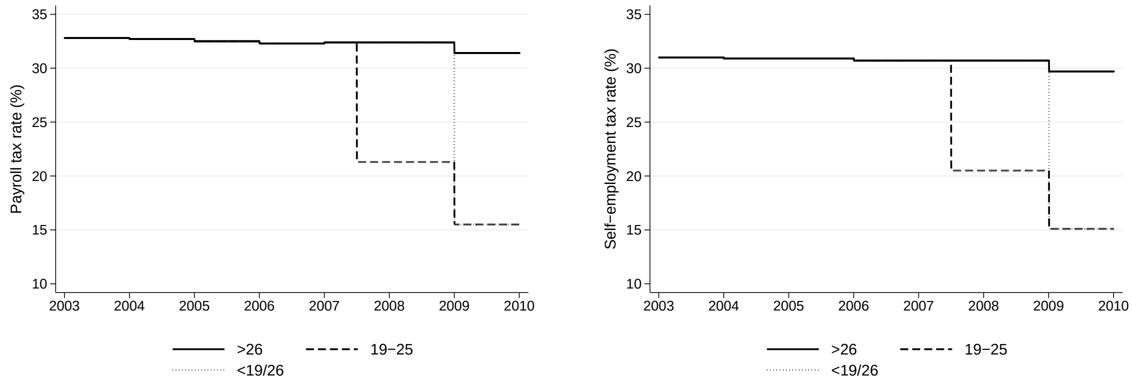
2.2 The 2007–09 tax reductions

Swedish employers finance welfare services for employees, such as pensions and health and disability insurances, through payroll taxes. Payroll taxes are proportional to the employee's wage bill, and consist of seven mandatory fees. Those who are self-employed finance their own welfare through a mandatory self-employment tax. This is essentially a tax on the surplus that the business has generated during the year. Between 2007 and 2009, the payroll tax for employers of young workers was cut in two steps. During the same period, young business owners were twice allowed reductions in the self-employment tax. Figures 1 (a) and 1 (b) provide a graphical illustration of these reductions.

On July 1, 2007, the payroll tax was cut by roughly 11 percentage points for (employers of) workers who, at the start of the year, had turned 18 but not 25 years of age. Six out of seven mandatory fees were halved, reducing the tax rate from 32.42 to 21.32 percent. Individuals within the same age interval were, from the same date, allowed a reduction in the self-employment tax. The rate was cut by 10 percentage points from 30.71 to 20.45 percent.[3] On January 1, 2009, the reform was modified in two ways. First, the target group was extended at both ends so that the reductions now encompassed all individuals who at the start of the year had not yet turned 26 years of age (i.e., the upper cutoff was changed and the lower cutoff was abolished). Second, the payroll tax was lowered additionally to 15.52 percent, and the self-employment tax additionally to 15.07

[3] July 1, 2007, is first mentioned in a press release from the ministry of Finance in October 2006. This date was confirmed when the new policy was ratified in the parliament on 15 March 2007.

Figure 1: The reductions in the payroll tax (a) and in the self-employment tax (b)

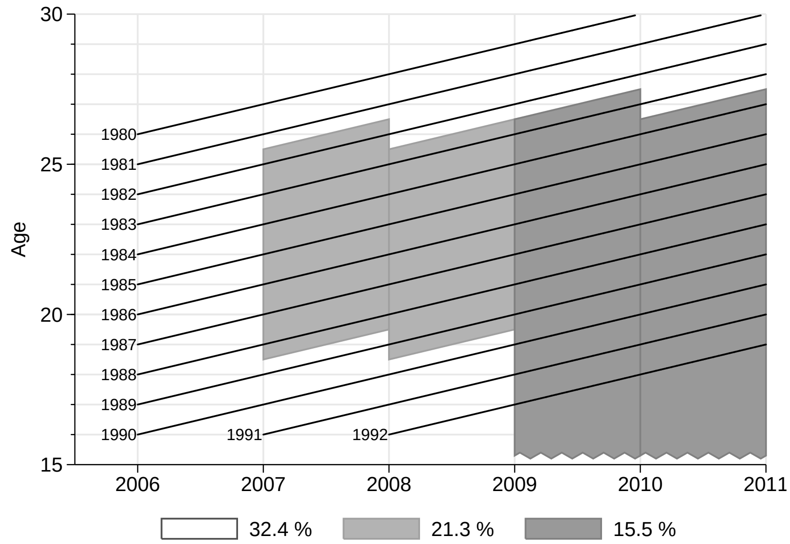


(a) (b)

percent. Figure 2 illustrates how different cohorts are subjected to the payroll tax reductions (the reductions in the self-employment tax apply to the same cohorts). In 2007, the target group consists of individuals born 1982–88 whereas in 2008 it consists of those born 1983–89. For simplicity, hereafter an age group a denotes all individuals who turn a during the year. With this terminology, the target group of the 2007 reform is referred to as "individuals aged 19–25", and the target group of the 2009 reform as "individuals aged 26 or below". The group of 19–25 year-olds comprised around 10 percent of the labor force aged 15–64 in 2007, and so the number of individuals directly affected by the new regime was substantial. The tax reductions were automatically implemented via the tax system, meaning that neither employers nor the self-employed had to send in an application to benefit from the lower tax rates.

There are two types of businesses that you can run if you are self-employed in Sweden, an unincorporated or an incorporated firm. The absolute majority of young business owners run unincorporated firms. It is those who run unincorporated firms that pay the self-employment tax. Due to the reduced rate, a young person running an unincorporated firm pays a lower tax on her surplus, and hence has more money left at the end of the day (all else equal). She cannot use the surplus for dividends since there are no shareholders; what is left of the surplus after she has paid

Figure 2: Evolution of treatment status across cohorts



the self-employment tax is her (taxable) income that she reports to the tax authorities. Those who run incorporated firms pay payroll taxes, since they are seen as employees. (This organizational form has limited liability, but requires an initial capital investment of 50,000 SEK.) After the reform, a young person running an incorporated firm has a lower tax rate on his gross wage. Those who run incorporated firms can choose whether to use the surplus from the business as salary or to pay shareholders in the form of dividends. Dividends are taxed at the same rate (30 %), irrespective of age.

When analyzing the Swedish reform it is important to bear in mind that there is an asymmetry regarding tax incidence: whereas the payroll tax is levied upon employers, the self-employment tax falls directly on the person who is running the business. This means that the reform reduced one tax faced by employers, and one tax faced by the individual worker.[4] The main purpose with these tax reductions was to decrease youth unemployment in general. Egebark and Kaunitz (2013) looks at how employers responded to the reduced payroll taxes. They find that the 2007 reduction led to a 2.5 percent employment increase, and that the estimated effect on wages was small, at around 1

[4]Of course, one could argue that the payroll tax ultimately falls on employees in the form of lower wages, at least in the long run.

percent (resulting in a demand elasticity of -0.37). They also show that the extended payroll tax reduction, implemented in 2009, did not boost employment further. Skedinger (2014) evaluates the same payroll tax reductions, but focuses on the retail industry. His findings are in line with the results in Egebark and Kaunitz (2013).

2.3 Other labor market reforms

With the purpose of increasing employment, both in general and for specific groups, several labor market reforms were introduced in Sweden during 2007. First, temporary subsidies for firms that hire individuals who have been unemployed or have received sickness or disability benefits, *New Start Jobs* (NSJ), were introduced on January 1, 2007. In 2007–08, individuals aged 20–24 could apply for the subsidy after six months of non-employment, whereas those who had turned 25 could apply only after twelve months of non-employment; thus, in contrast to the payroll tax cut, it was the exact age that mattered. In 2009, this cutoff was modified so that those who at the start of the year have turned 20 but not 26 were eligible after six months. Consequently, in 2007–08 the target groups overlapped, and from 2009 onwards they completely coincide. In principle, this raises a concern that the estimates in this study will be contaminated. It turns out, however, that the number of applications for NSJ (available in the data) was comparatively low, at about 0.5 percent of the ages 20–26, and the difference in shares between 21–25 year-olds and 26-year-olds—the potential bias of the estimates—is around 0.1 percentage points. I thus conclude that this is not a source of concern.

Second, income tax deductions were introduced in Sweden on January 1, 2007, with the purpose of increasing labor supply in general. These deductions apply to all workers, regardless of age, but I cannot rule out that there is heterogeneity in labor supply effects with respect to age. If younger workers' labor supply responded differently, the estimates for the tax effect could potentially be biased. Edmark et al. (2012) show that it is difficult to evaluate this deduction scheme due to the lack of unaffected comparison groups; hence, we do not know exactly how different age groups responded. In this study I assume that the response was similar for individuals close in age.

Finally, a third reform concerns employment protection legislation. Loosening of regulation in

2007 made it easier for employers to use fixed-term contracts. As temporary work is relatively more widespread among young workers, employment (and wages) may have been affected more for younger workers. However, Skedinger (2012) reports that only 1.4 percent of all temporary workers were employed with the new regulations in 2008. The reform, thus, had little impact in practice.

3 Previous literature

Previous research on the link between taxation and self-employment has primarily focused on the effect of income taxes. The effect of income taxes on the decision to become a business owner is theoretically ambiguous. On the one hand, high taxation may cause lower levels of self-employment since the expected return from running a risky business venture decreases. On the other hand, higher taxes make it more attractive to underreport taxable income. Hence, since underreporting is easier for the self-employed, there may be a positive effect. In addition, since most countries grant (various types of) loss offsetting, a higher tax may encourage risk taking due to the fact that the government's share of the loss increases with the tax rate (Domar and Musgrave, 1944). The theoretically ambiguous effect has lead researchers to turn to empirical evidence. So far, however, there is no consensus on the direction (and magnitude) of the effect (see Bruce, 2002, and Hansson, 2012, for brief summaries). This is true also for later work that uses individual level data instead of aggregated time-series. For example, Schuetze (2000) and Cullen and Gordon (2004) both find a positive correlation between income taxes and self-employment, whereas Gentry and Hubbard (2003) find no statistically significant correlation, and Gentry and Hubbard (2004) find a negative relationship. A final example is Hansson (2012), who uses Swedish data to show that income taxes are negatively correlated with the probability of becoming self-employed.

More recent work has focused less on the level of the income tax and more on the tax structure. Bruce (2000, 2002) analyzed how different taxation of income from employment and self-employment in the U.S. affects the choice to enter and exit self-employment. He found that larger individual-specific differences in marginal tax rates in the two sectors reduce self-employment entry rates: those with higher wage-and-salary minus self-employment differences (in expected marginal tax rates) are less likely to become self-employed. Bruce stresses the role of tax avoidance and evasion as one potential reason for this seemingly counterintuitive result. A contrasting example is Stabile (2004), who used different tax treatments in Canada to study occupational choice. He exploited the unexpected introduction of a payroll tax faced by employees, but from which self-employed were exempt, as a natural experiment. In particular, he compared the region of Ontario in Canada, where the new tax was introduced, to three other regions without the new tax. Contrary to the findings in Bruce's studies, Stabile showed that the payroll tax levied on employees had a positive effect on the probability of starting a business.

One drawback with many of the existing studies on the link between taxation and self-employment is that they do not use exogenous variation in taxes to estimate treatment effects, and so it can be questioned whether they estimate causal effects. (The most credible study in this respect is Stabile, 2004.) Instead, they rely heavily on creating synthetic tax rates, by using lags and leads, to control for the potential endogeneity that arises because an individual's decision to move into self-employment affects her income tax rate (i.e., in order to deal with simultaneity). Furthermore, these studies rely on econometric techniques such as, e.g., including inverse Mills ratios as controls to address the so-called initial conditions problem, which is essentially a bias due to omitted variables. I argue that this study offers a more credible identification strategy by making use of a Swedish reform that suddenly made taxes vary by age. Consequently, since I have well-defined treatment and control groups that can be followed over time in a Difference-in-Differences design, I should come closer to estimating the causal effect of taxes on the decision to become self-employed.[5] Another advantage is that I use a tax change that is both immediate and substantial. Previous studies have mostly used fairly small variations in taxes seen over longer periods of time.

While this study ties in with the literature on taxes cited above, it also deals with a more specific question: how can employment opportunities for disadvantaged groups be improved? As pointed out by both the OECD and the European Commission, stimulating entrepreneurial activity could work as a way to lower unemployment for those who have a hard time finding a regular job, such as

[5] Since this study focuses on young individuals, one should bear in mind that the results may not generalize to the population in general. As pointed out by Bruce (2000), there are surprisingly few studies on youth self-employment. Important exceptions include Blanchflower and Meyer (1994) and Dunn and Holtz-Eakin (2000).

youths and young adults (European Comission, 2013; OECD, 2013).[6] So far there are few studies on how to increase self-employment among those with a weak attachment to the labor market, and so it is difficult for policymakers to know what approaches actually work (OECD, 2012, 2013). To the best of my knowledge there is no previous study on the effect of introducing tax reductions to support youth self-employment.

4 Conceptual framework

I study the effect on self-employment of a reform that simultaneously reduced the employer paid payroll tax and the tax paid by those who are self-employed. The purpose of the 2007–09 tax reforms was to reduce youth unemployment. The motivation for cutting the payroll-tax was to create an incentive for employers to hire young workers, by making the labor cost substantially lower (see, e.g., Gruber, 1997; Bennmarker et al., 2009; Huttunen et al., 2013). The reduced self-employment tax, on the other hand, suddenly made it more attractive to be self-employed than to be unemployed, for those 19-25 years of age. (The prediction is that the net outflow from unemployment to self-employment increases.) This is the most obvious reason for why we would expect self-employment to increase for those in the target group, relative to older individuals. For a concrete example, compare a 25-year-old self-employed to a 26-year-old self-employed. As of July 1, 2007, the younger business owner has a competitive advantage due to the possibility to charge a lower price (all else equal). Or, she could keep prices unchanged and use the extra money for other purposes, such as investments, private consumption etc. In either case, the tax cut has made life easier for the younger self-employed. A reasonable prediction is, therefore, that individuals in the target group have a lower probability of transitioning from self-employment to unemployment (and, vice versa, a higher probability of transitioning from unemployment to self-employment).

It should also be useful to consider transitions between occupations in more detail (i.e., transitions from employment into self-employment and vice versa). Predictions about how transitions are affected are not as clear-cut. We might think that since the individual faces the self-employment

[6] There are many studies on whether or not stimulating self-employment is a good way of reducing unemployment. I will not review this literature here.

tax himself, any change in this tax is more salient (Chetty et al., 2009). If the saliency of the tax is important, this may lead us to predict a reduced net outflow from self-employment to employment. On the other hand, if a large fraction of those that are self-employed are forced into this type of occupation, any change in the payroll tax that induces employers to start hiring may have the opposite effect on the net outflow (i.e., increased net outflow from self-employment). Since the direction of the effect appears to be ambiguous, this is in the end and empirical question. In the analysis below I therefore measure the effect on both the (overall) self-employment rate, and on occupational choice.

Income effects are, to some extent, more predictable. The reductions in the self-employment tax are implemented automatically through the tax system. Hence, all else equal, the size of the effect on income from self-employment is known. Any difference between this mechanical effect and the actual effect indicates the existence of behavioral adjustments.

5 Data

The data are collected by Statistics Sweden (SCB) and contain yearly information on employment status, income and demographical characteristics for all individuals living in Sweden who are at least 16 years of age, for the years 2002–09 (the LOUISE and RAMS data sets). The registers contain information on various types of income, including income from wage employment and from self-employment. Even though I have information on individuals as young as 16 years of age, I will not consider the youngest individuals of the target group in the analysis. The reason for excluding 19–20 year-olds is that they turn out to be substantially different in terms of cyclicality, thus making comparisons over time difficult. I will thus focus on 21–25 year-olds in the following.

Table 1 presents 2006 summary statistics for individuals 21–27 years of age. The table shows that younger individuals (21–25 year-olds) have a lower probability of being employed, and of being self-employed. I use the definition of self-employment that is used by Statistics Sweden. First, an individual is defined as self-employed if the income she earns (in November) comes exclusively from the own firm. Second, if the individual has income both from an own firm and from employment, she is defined as self-employed if the income from the business multiplied by 1.6 is greater than the income from employment.[7] (The definition of self-employment was slightly different before 2004.) Younger individuals further have lower incomes, both from employment and from self-employment. For both age groups, men are more likely to be self-employed; they also earn more than women, both as employed and as self-employed.

As is evident from table 1, it is only a small fraction of the young that run their own business. Table 2 follows three different groups of young individuals over time, and thus provides some understanding of who the young self-employed are. Columns 1–2 look at young employed, columns 3–4 at young self-employed, and columns 5–6 at somewhat older self-employed. Panel A considers unemployment risk. Compared to the young employed, the young self-employed have fewer days as registered unemployed, and a lower probability of being registered in the first place, both in 2006 and in 2008 (see columns 1–4). Looking at the two groups of self-employed (columns 3–6) we realize that there is basically no difference between them, neither in levels nor in trends. Since there are no changes across the two groups over time I conclude that the 2007 tax cut did not cause a compositional change in terms of unemployment risk.[8]

Panel B highlights some of the differences, and similarities, between the three groups in terms of industries. I have included the three industries where the employed mainly works, and the three industries where the self-employed mainly works. About 44 percent of the young employed works within the three industries listed first, i.e. within Health care, Manufacturing, or Retail. The corresponding figure for the young self-employed is instead 7–9 percent. For the next three industries it is more or less the opposite relationship that holds: 50 percent of the young self-employed works within Construction, Hair/Body or Hotel/Restaurant, as compared to 21 percent for the young employed. The most striking difference is found for Hair/Body: roughly 15 percent of the self-employed are found in this industry, where less than one percent of the employed works.

I complement these numbers with figure 3 (a), which includes more industries and thus provides a more detailed comparison. One thing that is clear is that there are several industries where the

[7]The higher weight on income from the business is due to the fact that, for given levels of income, the number of hours spent working as self-employed is typically greater than the number of hours spent working as employed.

[8]Regression results confirm that there are no statistically significant compositional changes, for any of the characteristics presented in table 2.

Table 1: 2006 summary statistics by age group

	21–25 year-olds		
	All	Men	Women
Employed	65% (N=536,859)	65% (N=274,716)	65% (N= 262,143)
Self-employed	1.5% (N=536,859)	1.9% (N=274,716)	1.0% (N=262,143)
Income from empl.[a]	129,800 (N=448,097)	150,300 (N=230,694)	108,000 (N=217,403)
Income from self-empl.[a]	66,740 (N=7,912)	69,200 (N=5,318)	61,600 (N=2,594)
	26–27 year-olds		
	All	Men	Women
Employed	69% (N=221,914)	70 % (N= 113,233)	67% (N=108,681)
Self-employed	2.7 % (N=221,914)	3.6% (N=113,233)	1.8% (N=108,681)
Income from empl.[a]	179,200 (N=185,715)	203,700 (N=95,892)	153,100 (N=89,823)
Income from self-empl.[a]	80,400 (N=6,044)	85,600 (N=4,131)	69,200 (N=1,913)

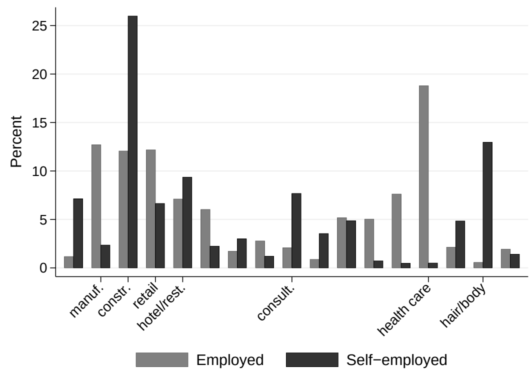
[a] Yearly income in SEK.

Table 2: Comparing levels and trends across three subgroups

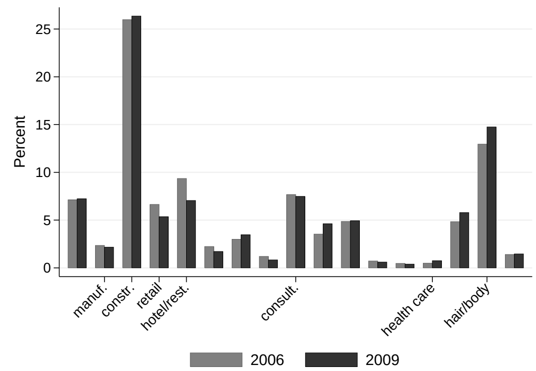
	21–25 employed		21–25 self-employed		26–27 self-employed	
	2006	2008	2006	2008	2006	2008
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Unemployment risk, previous year					
Days unempl.	23.4	14.5	15.3	8.5	15.6	8.2
	(52.1)	(42.2)	(45.0)	(33.9)	(51.7)	(38.0)
Prob. unempl.	0.28	0.17	0.17	0.10	0.14	0.08
	(0.45)	(0.37)	(0.38)	(0.30)	(0.35)	(0.27)
	Panel B: Industries					
Health care	19%	18%	0.5%	0.6%	0.9%	0.8%
Manufacturing	13%	13%	2.3%	2.1%	2.7%	2.6%
Retail	12%	12%	6.6%	4.6%	5.2%	4.1%
Construction	12%	13%	26%	26%	31%	32%
Hair/Body	0.6%	0.6%	13%	15%	11%	11%
Hotel/Rest	7.1%	7.2%	10%	7.5%	7.2%	7.2%
	Panel C: Parent characteristics					
Father's inc.[a]	277,000	296,000	251,000	262,000	234,00	240,000
Mother's inc.[a]	204,000	225,000	186,000	202,000	183,00	199,000

Notes: Unemployment risk refers to the previous year.[a] Yearly income in SEK.

Figure 3: Industry comparisons



(a) 21–25 year-olds in 2006



(b) Self-employed 21–25 year-olds

difference between the two groups is substantial. Columns 3–6 of panel B show that the two groups of self-employed are very similar also in terms of the industries where they are active. As before, since there are no changes across the groups over time, there is no compositional effect in this dimension either. Figure 3 (b) provides an illustration of how stable industry composition is over time.

Finally, panel C shows some slight differences with respect to parents' income. We notice that parents' income is higher for the employed, and that the parents of the young self-employed earn more than the parents of the older self-employed.

6 Identification

I use the Difference-in-Differences (DiD) estimator to capture the effect of the tax cuts on self-employment and income. I estimate the following model:

$$y_{i,t} = \delta_t \cdot D(i,t) + \mathbf{x}'_{i,t}\boldsymbol{\beta} + \varepsilon_{i,t} \tag{1}$$

where $y_{i,t}$ indicates whether individual i is self-employed in year t, $D(i,t)$ is a treatment indicator for individual i in year t, δ_t is the DiD estimate for year t, and $x_{i,t}$ is a vector of control variables. This vector includes dummy variables for year, age, gender and whether being foreign-born, and indicators for local labor market. When studying income, the only thing that changes is the outcome variable, $y_{i,t}$.

DiD uses the evolution of the control group over time as a measure of how the treatment group would have evolved, had the intervention not taken place. The key assumption is, hence, that the two groups would have moved in parallel in absence of treatment. This parallel trends assumption is, by definition, not testable since it concerns counterfactual outcomes. However, to get an indication of whether it is likely to hold it is important to confirm that the evolution of the treatment and control groups are similar before the intervention took place. In the analysis below I therefore estimate treatment effects also for the years before the tax reductions was in place.

An implication of the DiD identifying assumption is that the control group must not be affected by the intervention under study. If such treatment spillovers exist, DiD will not measure the difference between the reform outcome and the counterfactual outcome, but the difference to the control group deviation from its counterfactual outcome. Consequently, we obtain a measure of the relative rather than the absolute effect of the reform. In the present case I cannot rule out that age groups just above the treatment-defining cutoff are unaffected. For example, the tax cut gives 25-year-olds the opportunity to compete with 26-year-olds by reducing prices; this would certainly affect 26-year-olds ability to continue running their businesses. Hence, using 26-year-olds as the control group in the above DiD-model may overestimate the absolute effect. With this in mind I estimate the (relative) effect using both the 25–26 and the 21–27 age-intervals. (Using a larger bandwidth should also, at least to some extent, handle different types of cohort heterogeneity.)

A second issue with using DiD in this case is that it gets increasingly difficult to find a control group that has not been subjected to the tax reduction in the past. For example, comparing 25-year-olds to 26-year-olds implies that the control group in 2008 (those born 1982) was in the target group the year before (see figure 2 in section 2.2). Ultimately, this means that it is hard to identify the reform effect for the later years in the sample. One way to handle this issue would be to use 27-year-olds instead of 26-year-olds as the control group. In the analysis below I experiment with altering the control group in this way.

As described in section 2.2 above, the 2007 tax cuts were extended in 2009. If the initial reduction had an impact, it is, due to the problems described above, difficult to evaluate the 2009 reduction. However, in absence of any earlier impact, the parallel trends assumption is more likely to hold. The most straightforward way would then be to compare 25–26 year-olds to 27–28 year-olds in 2009: 26-year-olds transition from no treatment straight to 2009 treatment while 25-year-olds transition from 2007 treatment to 2009 treatment.[9] By considering the later years in the sample I will be able to uncover whether the effect differs as the economy is hit by the financial crisis. This is important since it helps shed light on whether financially stimulating self-employment, in the form of tax reductions, works better in times of economic slowdowns.

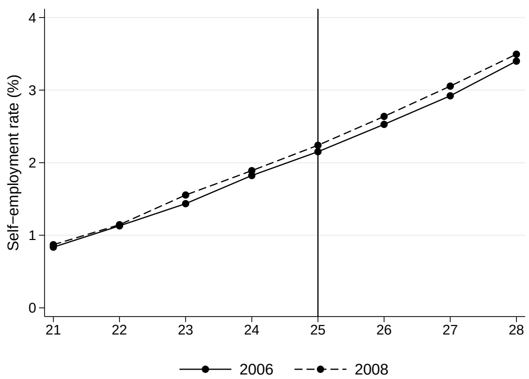
7 Results

7.1 Main findings

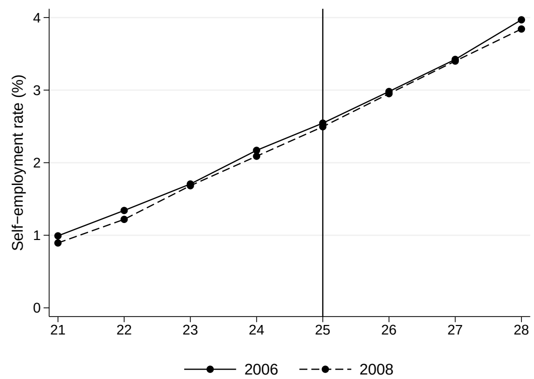
A first look at the data clearly indicates that youth self-employment was unaffected by the 2007 tax changes. Figure 4 (a) shows how the self-employment rate—here defined as the fraction of self-employed in the population as a whole—depends on age, before and after the 2007 reform. As is seen, the distribution looks very similar between 2006 and 2008, especially if we focus on ages 24–26. The fact that nothing seems to happen around the cut-off (indicated by the vertical line)

[9] This comparison hinges on the assumption that there are no effects of the initial 2007 reductions. Below I discuss whether this assumption appears to be valid empirically.

Figure 4: Self-employment rates before and after the 2007 tax cut



(a) Full sample



(b) Conditional on working in 2006

suggests that the propensity to be self-employed did not increase for the younger individuals. As described in section 4, it is also useful to consider transitions between occupations in more detail (i.e., transitions from employment into self-employment and vice versa).[10] A simple way to test whether the choice of occupation was affected is to narrow the sample to individuals that were either employed or self-employed in 2006, i.e., the year before the tax cut was implemented. Figure 4 (b) gives the age distributions before and after the reform, for the smaller sample. (Excluding those not employed or self-employed in 2006 increases the self-employment rate slightly.) Neither for this sample there are any visible changes around the cutoff. (If anything, the propensity to be self-employed seems to decrease for younger individuals relative to older ones.)

The main message from the graphical evidence is that youth self-employment was unaffected by the tax reductions up until 2008. This conclusion is supported by DiD-estimates in tables 3 and 4. Both tables report treatment effects for different age groups, using 26-27 year-olds as the control group. When I study self-employment in the following I chose to exclude years 2002–03. The reason for this is twofold. First, precision is higher when using the shorter time period, making it easier to

[10]For the self-employment rate in figure 4 (a) to be constant, any increase (decrease) in the net inflow from employment to self-employment has to be compensated (exactly) by increased (decreased) net outflow from self-employment to non-employment. Hence, already the result that the self-employment rate in figure 4 (a) is unaffected could be taken as an indication that there is no effect on transitions.

draw strong conclusions. In particular, focusing on a small time window around the time of the tax reform makes it easier to detect even small changes in the outcome; as will be clear, I do not find any significant effects even with the narrow window, and these null results are precisely estimated. A second argument for not using 2002–03 is that the definition of self-employment changes in 2004 (see section 5). Importantly, using the longer time period does not produce any substantially different results. To facilitate readability I have multiplied all the coefficients in the tables by 100, and so the point estimates represent percentage points.

Table 3 shows pooled treatment effects, for each of the two samples used in figures 4 (a) and 4 (b). (This is the simplest possible model in the sense that it groups years 2004–06 and 2007–08, respectively.) Evidently, the reductions had no effect on youth self-employment. First, there are no statistically significant effects for any of the samples, irrespective of what age-interval is used. Second, consider for example the full sample: we can be 95% confident that any (positive) effect for 21–25 year-olds is no greater than 0.06 percentage points (i.e., the upper bound for a 95%-confidence interval is 0.058 percentage points).

Table 4 gives a more detailed picture by reporting DiD-estimates (in percentage points) for each year 2006–08. The reason for including the treatment effect for 2006 is to examine whether the trends prior to the policy intervention are the same in treatment and control groups. The fact that the pre-treatment effects are statistically insignificant, and precisely estimated, lends credibility to the identifying assumption of parallel trends. In general, the results in table 4 support the findings from above.

Since there is no treatment effect for 25-year-olds in 2007, the issue of a lagged treatment effect for the control group in 2008 is less of a problem. Nevertheless, to address the bias discussed in section 6, I have also tried using only 27-year-olds as the control group in all of the above specifications (thus excluding 26-year-olds). As is seen in tables A.1 and A.2 in the appendix, all of the above results are robust to this change. As a second sensitivity test, I have tried changing the definition of self-employment in table A.3 in the appendix. Instead of using the dummy variable described in section 5 as outcome, I use a dummy variable that equals one if an individual has non-zero income from self-employment (and zero if income from self-employment is zero). I thus disregard the fact

Table 3: Pooled effects (in percentage points) by sample and age group

	Full sample		Conditional	
	TG: 25	TG: 21–25	TG: 25	TG: 21–25
DD 07–08	–0.002	–0.011	–0.003	–0.037
	(0.060)	(0.035)	(0.069)	(0.041)
$\overline{y}_{\text{TG}}$	2.2	1.5	2.5	1.7
N	1,091,071	3,804,593	909,945	3,170,954
R^2	0.003	0.005	0.004	0.006

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. The outcome is a dummy variable for being self-employed in a given year. Point estimates represent percentage points. $\overline{y}_{\text{TG}}$ gives the treatment group average (in percent) in the treatment period. Years 2004–06 constitute the reference period. While the definition of the treatment group (TG) varies, the control group consists of 26–27 year-olds. The conditional sample uses those who work in 2006. Fixed effects included for year, age and demographic characteristics (see section 6). Robust standard errors in parenthesis.

that an individual might have income from wage employment. With this alternative definition, I can examine whether the lower self-employment tax caused a young person to at least try to start a business (since filing a positive or negative income amount indicates that the individual has a business of some kind). Strikingly, even for this (very liberal) definition of self-employment, the lack of treatment effects is precisely estimated.

Finally, I have experimented with smaller sub-samples to find out if there exist any heterogeneous treatment effects (see table A.4 in the appendix). First, I have run separate regressions for men, women, natives, foreign-born, and those with vocational training. For vocational training I have considered both a broad group and a smaller, more homogenous, group.[11] None of these six groups display any significant effects. (For foreign-born and the vocationally trained, where samples sizes are smaller, the null results are less precise.) Second, previous research suggests that the response to tax changes depends, to a large degree, on how aware people are of new tax rules.

[11] I restrict the sample to individuals who, at age 22 at the latest, have finished three years of vocational training. Vocational training in Sweden is relatively broad; it includes for example musical and art training. The more homogenous group consists of, e.g., carpenters, painters, plumbers and electricians.

Chetty et al. (2013), for example, show that labor supply effects of the Earned Income Tax Credit vary substantially across neighborhoods: individuals in high-knowledge areas change wage earnings sharply to obtain larger EITC refunds relative to those in low-knowledge areas. This finding suggests that individuals with better knowledge about the 2007 tax reform could have responded differently. It seems reasonable to expect that individuals whose parents are self-employed are more aware of the tax change, and, to some extent, we may also expect those with high income parents to have better knowledge. Hence, I run separate regressions for those that have either a mother or a father who is self-employed, and for those with high income parents. As is seen in the bottom panel of table A.4, there are no significant effects for any of these subsamples either. While this could indicate that knowledge was less important in the present case, we should also note that for those with self-employed parents the estimates are less precise. In other words it is difficult to draw strong conclusions.

Table 5 studies the extended tax reductions that were implemented in 2009. As discussed in section 6, I use 25–26 year-olds as the treatment group, and 27–28 year-olds as the control group. The easiest way to examine whether the 2009 additional reductions had an impact is to contrast 2009 to 2004–08. This is done for each of the respective samples in columns 1 and 3. The fact that the 2009 estimate is statistically insignificant and precisely estimated clearly speaks against any impact. Since the comparison is based on the assumption that there are no effects of the 2007 initial tax reductions, I have also allowed for treatment effects for each year 2007–09 in columns 2 and 4. The estimates for the pre-treatment years are insignificant, but somewhat large (in absolute terms).[12] Even though the common trends assumption is somewhat less credible in table 5, it seems reasonable to conclude that there was no effect of the 2009 reduction. This result is important since it indicates that youth self-employment is insensitive to tax reductions also in times of economic slowdowns.

[12] The fact that the estimates appear to jump above and below zero depending on year could indicate that it is just random shocks.

Table 4: Yearly effects (in percentage points) by sample and age group

	Full sample		Conditional	
	TG: 25	TG: 21–25	TG: 25	TG: 21–25
DD 2006	0.016	0.023	–0.004	–0.060
	(0.079)	(0.047)	(0.092)	(0.055)
DD 2007	0.014	0.026	0.027	–0.019
	(0.080)	(0.047)	(0.092)	(0.054)
DD 2008	–0.007	–0.031	–0.025	–0.095
	(0.080)	(0.047)	(0.093)	(0.055)
$\overline{y}_{\mathrm{TG}}$	2.4	1.5	2.5	1.7
N	1 091 071	3 804 593	909 945	3 170 954
R^2	0.003	0.005	0.004	0.006

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. The outcome is a dummy variable for being self-employed in a given year. Point estimates represent percentage points. $\overline{y}_{\mathrm{TG}}$ gives the treatment group average (in percent) in the treatment period. Years 2004–05 constitute the reference period. While the definition of the treatment group (TG) varies, the control group consists of 26–27 year-olds. The conditional sample uses those who work in 2006. Fixed effects included for year, age and demographic characteristics (see section 6). Robust standard errors in parenthesis.

Table 5: Effects of 2009 extension (percentage points)

	Full sample		Conditional	
	25–26 vs. 27–28		25–26 vs. 27–28	
DD 2007		0.053		0.082
		(0.057)		(0.066)
DD 2008		−0.049		−0.029
		(0.057)		(0.067)
DD 2009	−0.005	−0.005	−0.006	0.004
	(0.055)	(0.058)	(0.065)	(0.068)
$\overline{y}_{\mathrm{TG}}$	2.6	2.6	2.8	2.8
N	2,647,469	2,647,469	2,647,469	2,647,469
R^2	0.004	0.004	0.005	0.005

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. The outcome is a dummy variable for being self-employed in a given year. Point estimates represent percentage points. $\overline{y}_{\mathrm{TG}}$ gives the treatment group average (in percent) in the treatment period. Years 2004-08 (2004-06) constitute the reference period in columns 1 and 3 (2 and 4). The conditional sample uses those who work in 2006. Fixed effects included for year, age and demographic characteristics (see section 6). Robust standard errors in parenthesis.

7.2 Income

The previous section provides convincing evidence against any effects on the extensive margin. This section uncovers adjustments along the intensive margin, by studying the effect on income. I examine both whether the lower taxes caused self-employed individuals to allocate more time to self-employment, and whether they reallocated time between wage employment and self-employment. I do this by simply comparing the estimated effect on income from self-employment to the predicted mechanical effect (i.e., the effect that exists irrespectively of behavior adjustments). While intensive margin responses to tax changes have proven difficult to detect in the past, recent work shows that such adjustments can be substantial (Chetty et al., 2013).

Figure 5 shows income from self-employment by age, for years 2006 and 2008, respectively. In 2006, just before the tax reform was implemented, income grows continuously with age (grey bars). In contrast, the 2008 distribution shows a sharp increase for those below the treatment defining age cutoff (black bars). While income is lower in 2008 than in 2006 for those above 25 years of age, this is clearly not the case for those below the cutoff. The income effect is substantial: close to the cutoff it pushes young individuals' earnings well above the level of older age groups.

Another striking result is given in figure 6, where I contrast 26-year-olds to 28–29 year-olds. (The reason for using 28–29 year-olds as control is that they are unaffected the whole period.) For 26-year-olds we do not expect any effect until 2009 when the target group was extended. In addition, we should expect a rather sharp increase this year, since for this age group the tax was cut by 15 percentage points at once. Figure 6 shows that, while relative income is (more or less) constant up until 2008, there is a sharp increase for 26-year-olds in 2009. Strikingly, the tax cut shoots income for the younger individuals above the level of the older ones.

The conclusions from the figures are strengthened by DiD estimations in tables 6 and 7. Table 6 gives pooled 2007–09 treatment effects for different age groups, using 28–29 year-olds as the control group. The pooled effects are highly significant and large in magnitude, irrespectively of whether I use small or large bandwidths. For 21–25 year-olds income has increased by as much as 20 percent.[13]

Pooling the income effect is useful for demonstrating the presence of a general treatment effect

[13] The estimated percentage increase is relative to the counterfactual outcome. It is, thus, obtained as $\beta/(\bar{y}_{\text{TG}} - \beta)$.

Figure 5: Income distributions before and after the 2007 tax cut

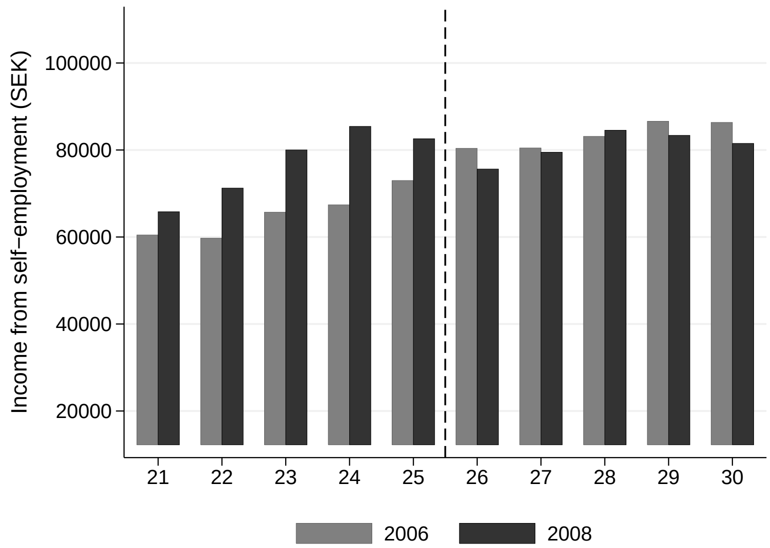


Figure 6: Income effect for 26-year-olds

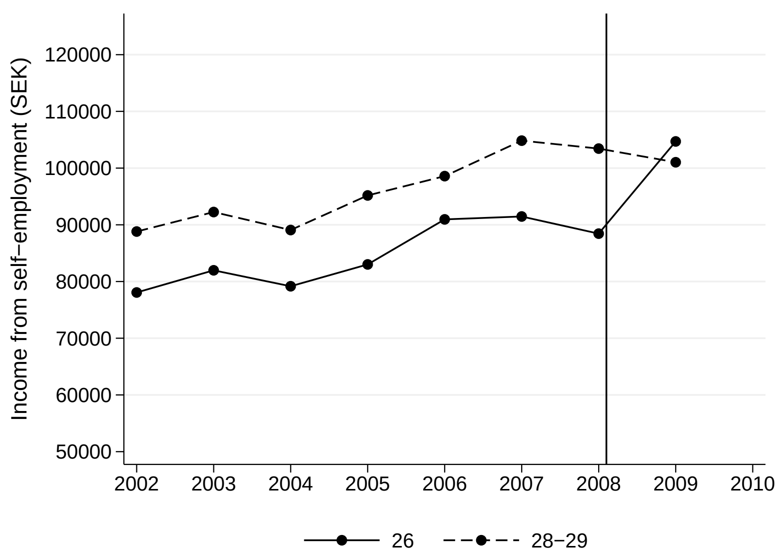


Table 6: Pooled income effects, by age group

	Income from self-employment (SEK)		
	TG: 25	TG: 23–25	TG: 21–25
DD 07–09	9,790***	12,150***	12,870***
	(1,800)	(1,350)	(1,230)
$\overline{y}_{\text{TG}}$	81,350	79,730	76,990
N	103,522	140,350	162,721
R^2	0.02	0.02	0.02
Effect size:	14%	18%	20%

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. The outcome is yearly income from self-employment in SEK. $\overline{y}_{\text{TG}}$ gives the treatment group average (in SEK) in the treatment period. Years 2002–06 constitute the reference period. While the definition of the treatment group (TG) varies, the control group consists of 28–29 year-olds. Fixed effects included for year, age and demographic characteristics (see section 6). Robust standard errors in parenthesis.

in the after period. However, yearly effects are hidden, and so we may miss out on interesting patterns in the data. Figure 1 (b) above shows that the first tax cut only applied to the second half of 2007 whereas in 2008 it applied to the whole year. In addition, there was an extended reduction in 2009 (from 20.45 to 15.07 percent). This means that we expect the income effect to increase in 2008 and in 2009, relative to 2007. Table 7, which gives yearly treatment effects, proves that the effect is in fact growing over time. In absolute terms, the 2009 effect is more than twice as large as the 2007 effect (as shown by columns 1–3). The positive effects on income is unlikely to be driven by age specific cyclicality, since they appear both when the economy was expanding in 2007–08, and when the economy was contracting dramatically in 2009. Column 4 of table 7 confirms what we saw in figure 6: 26-year-olds show a sharp increase in 2009, but not before. The last column is included as a simple placebo test; it shows that 27-year-olds, as we should expect, are unaffected during the whole period.

All of the pre-treatment point estimates in table 7 are insignificant and close to zero. This

Table 7: Yearly income effects, 2005–09, by age group

	Income from self-employment (SEK)				
	TG: 25	TG: 23–25	TG: 21–25	TG: 26	TG: 27
	(1)	(2)	(3)	(4)	(5)
DD 2005	1,720	290	890	−1,740	−130
	(2,500)	(1,850)	(1,700)	(2,320)	(2,490)
DD 2006	–370	63	910	760	−3,100
	(2,480)	(1,890)	(1,760)	(2,460)	(2,440)
DD 2007	3,065	5,390**	8,130***	−1,095	−1,490
	(2,800)	(2,120)	(1,930)	(2,670)	(2,740)
DD 2008	9,730***	13,150***	13,060***	1,840	1,130
	(3,000)	(2,250)	(2,000)	(2,850)	(2,850)
DD 2009	16,600***	17,240***	17,930***	13,300***	3,830
	(2,900)	(2,180)	(1,990)	(2,970)	(2,640)
$\overline{y}_{\mathrm{TG}}$	81,350	79,730	76,990	85,120	81,350
N	103,522	140,350	162,721	107,802	103,522
R^2	0.018	0.017	0.018	0.018	0.018

Notes: ***p < 1%, **p < 5%, *p < 10%. The outcome is yearly income from self-employment in SEK. $\overline{y}_{\mathrm{TG}}$ gives the treatment group average (in SEK) in the treatment period. Years 2002–04 constitute the reference period. While the definition of the treatment group (TG) varies, the control group consists of 28–29 year-olds. Fixed effects included for year, age and demographic characteristics (see section 6). Robust standard errors in parenthesis.

suggests that the parallel trends assumption holds, and so it is likely that the estimated income increases have a causal interpretation. However, the standard errors reported in the tables do not handle serial correlation in the error term (i.e., they are not clustered at any level). We may therefore worry that the significant income effects are the result of downward biased standard errors. Table A.5 in the appendix repeats the exercise of table 7, but with standard errors that are robust to two different types of clustering. Evidently, clustering does not change any of the conclusions. The reason that I prefer to use the robust standard errors in table 7 is that they are larger than any of the cluster robust errors (Angrist and Pischke, 2009).

So far we have seen that the 2007–09 tax reductions caused income from self-employment to grow substantially. What is important to understand at this stage is whether the estimated income effect exceeds the mechanical effect. The 2007 payroll tax reduction lowered the tax rate from 30.71 to 20.45 percent whereas the 2009 reduction lowered the rate down to 15.07 percent. All else equal, we thus expect income to increase by 7.4 percent in 2007, 14.8 percent in 2008, and by 22.6 percent in 2009, relative to the period before the changes.[14] Figure 7 shows the predicted percentage increases (grey bars) as well as the estimated percentage effects for 21–25 year-olds (black bars). Evidently, the estimated effect exceeds the mechanical effect each year 2007–09. This is important since it could indicate a labor supply adjustment along the intensive margin: the age groups that suddenly faced the lower tax may have increased the number of hours spent in self-employment—either by reducing leisure or time in wage employment. (I will come back to alternative explanations for the pattern in figure 7 later in this section.) It is easy to quantify the behavioral change. We note that the average yearly difference between the mechanical and the estimated effect, for 21–25 year-olds in 2007–09, is roughly SEK 3,250.[15] Furthermore, since the market wage for a typical 21–25 year-old worker is around SEK 100–125 per hour, the estimated adjustment corresponds to 26–32 hours per year. (The increase in working hours could be even bigger since the self-employed most likely earn less than the market wage.)

An interesting question from a policy perspective is how lasting the income effect is. We know that some cohorts, essentially because of luck, were allowed a lower tax rate for a limited period of time. The cohort that was born in 1982, for example, faced the lower tax rate for one year, in 2007. Does this mean that their income in 2008—when they are 26 and no longer eligible— increased as well? Figure 8 shows that this is not the case: the 1982 cohort experiences an income shock in 2007, but the effect disappears already the following year. In other words, the days with substantially higher income appears to have been temporary. This, in turn, means that any potential increase in working hours was short lived, i.e., there was no spillover effect to consecutive years when the tax rate went back to normal levels. In a broader perspective, this suggests that temporary tax cuts (even though they are substantial) will have no lasting effects on intensive margin labor supply.

[14]The 2007 effect is calculated assuming that income is evenly distributed across the first and the last six months of the year.

[15]In absence of a treatment effect, average income from self-employment for 21–25 year-olds in 2007–09 is SEK 76,990–12,870 = SEK 64,120 (see column 3 of table 6). The average yearly mechanical effect is 15 percent in 2007–09, which corresponds to SEK 9,618 (15 percent of SEK 64,120). Finally, the amount by which the estimated effect exceeds the mechanical effect is SEK 12,870–9,618 = SEK 3,252.

Figure 7: Differences between mechanical and estimated effects

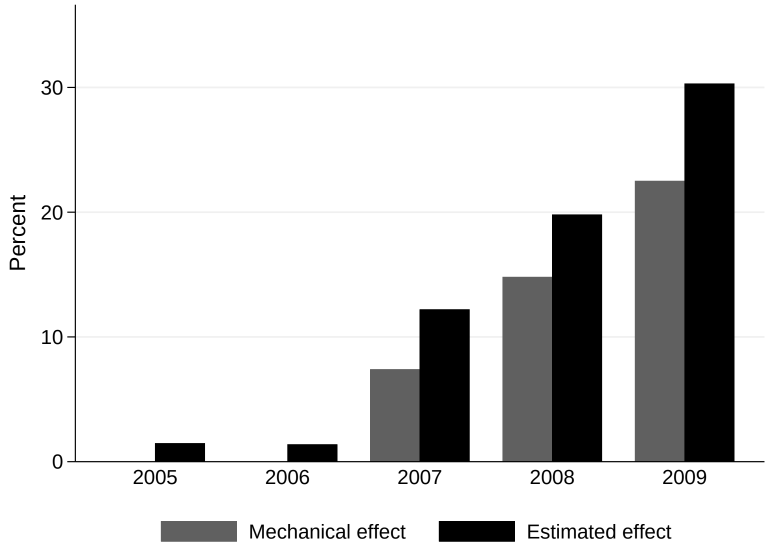


Figure 8: Persistence of 2007 effect

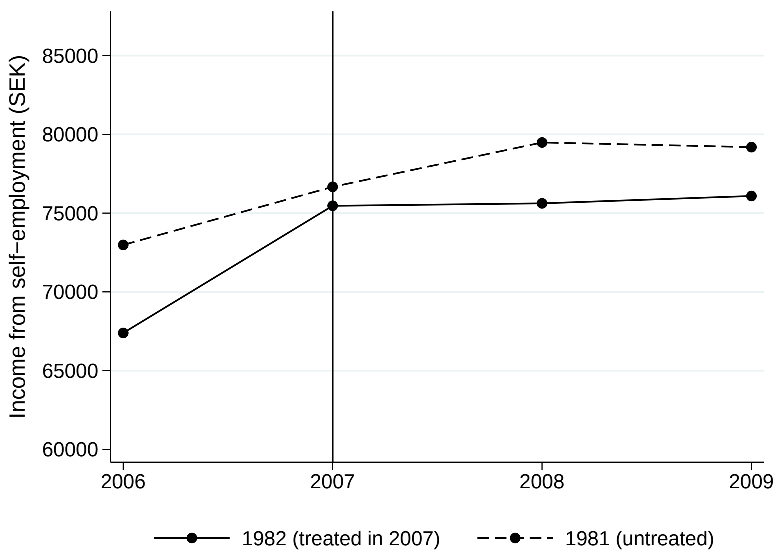


Table 8: Income effects for those self-employed in 2006

	Income from self-employment (SEK)			Income from wage employment (SEK)		
	TG: 25	TG: 23–25	TG: 21–25	TG: 25	TG: 23–25	TG: 21–25
	(1)	(2)	(3)	(4)	(5)	(6)
DD 07–09	18,840***	17,720***	$14,680$***	$-4,229$**	$-6,190$***	$-6,760$***
	(4,750)	(3,980)	$(3,730)$	$(2,230)$	$(1,620)$	$(1,510)$
$\overline{y}_{\mathrm{TG}}$	110,630	107,060	103,450	44,190	41,180	39,790
N	43,177	56,779	64,159	43,177	56,779	64,159
R^2	0.02	0.02	0.02	0.04	0.04	0.037
Effect size:	21%	20%	17%	9%	13%	15%

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. The outcome is yearly income from self-employment in SEK. $\overline{y}_{\mathrm{TG}}$ gives the treatment group average (in SEK) in the treatment period. Years 2002–06 constitute the reference period. While the definition of the treatment group (TG) varies, the control group consists of 28–29 year-olds. Fixed effects included for year, age and demographic characteristics (see section 6). Robust standard errors in parenthesis.

In principle, we may worry about endogeneity when studying income from self-employment above. If the tax cuts affected self-employment, it is not the same types of individuals that are self-employed after 2006, and, hence, there is a selection issue that potentially could bias the income measures.[16] Of course, since we saw rather convincing evidence above that nothing happened to self-employment this should not be a big concern. I nevertheless repeat the table 6 analysis for a smaller sample that keeps only those individuals who were self-employed in 2006 (i.e., one year before the reform). Columns 1–3 of table 8 are reassuring, as they show that there are significant and large income effects for the restricted sample as well.

A second advantage with using the restricted sample is that it allows me to look at income from wage employment. A large fraction (roughly one third in 2006) of the young self-employed has in fact wage income, even though it is in general low. Any effects on this margin would suggest that the tax cuts caused the self-employed to allocate time in regular wage work differently. Table 8 shows significant negative effects on income from wage employment. This means that the number

[16] Since it is not obvious what the selection looks like—it could be more, or less, able individuals that become self-employed—it is difficult to say in what direction the bias goes.

of hours spent in regular work has (most likely) decreased. Consider, for example, 23–25 year-olds in columns 2 and 5 of table 8 (for 23–25 year-olds, the percentage effects are somewhere between the percentage effects for the other two age groups). 23–25 year-olds earn around SEK 125 per hour in regular work, implying that time in wage employment has decreased by about 50 hours per year (SEK 6,190 divided by SEK 125). Furthermore, the average yearly mechanical effect, for 23–25 year-olds in 2007–09, amounts to SEK 13,340, which means that the behavioral adjustment corresponds to SEK 4,380 per year (SEK 17,720–13,340). If we assume that a self-employed individual earns around 60 percent of the market wage (i.e., SEK 80), time in self-employment has increased by more or less the same number of hours (SEK 4,380 divided by SEK 80).[17] Strikingly, based on these simple calculations it appears as if the self-employed did not increase total number of working hours in the wake of the tax reform.

While the results in table 8 suggest that self-employed individuals reallocated time from employment to self-employment, we cannot be certain that such substitution took place. It could also be that individuals just lowered their time in regular work, without increasing time as self-employed. In particular, since they suddenly have more money in the pocket they may have used the money to buy leisure by reducing time in regular work. But then the difference in the estimated and the mechanical effects in figure 7 still needs to be explained. While there are other potential reasons, for example decreasing tax evasion, it seems reasonable to interpret the results as evidence for reallocation of working hours across occupations.

I finally test for heterogeneity by using different subsamples in table A.6 in the appendix. We first conclude that there are no gender differences, as both men and women display the same percentage income increase (about 20 percent). For natives, on the other hand, the response seems to have been larger than in general (about 24 percent). The heterogeneity could be taken as support for the awareness story discussed above: since natives have higher expected knowledge about changes in the Swedish tax schedule, we expect them to respond stronger. For foreign-born the coefficient is close

[17]In absence of a treatment effect, average income from self-employment, for 23–25 year-olds in 2007–09, is SEK 107,060–17,720 = SEK 89,340 (see column 2 of table 8). The average yearly mechanical effect is 15 percent in 2007–09, which corresponds to SEK 13,341 (15 percent of SEK 89,340 is SEK). Finally, the difference between the mechanical effect and the estimated effect, which corresponds to the yearly behavioral adjustment, is SEK 17,720–13,341=SEK 4379. Statistics Sweden assumes that a self-employed individual earns about 60 percent of the regular wage when they define self-employment; I use this assumption as guidance.

to zero, and insignificant. This may appear strange at first glance, as we expect at least a mechanical effect. However, for this small sample, the comparison groups display different trends before 2007, making it difficult to identify any effects. Table A.6 also shows that those with high-income parents seem to have responded slightly less.

8 Conclusion

In this paper I examine the link between taxes and youth self-employment. I make use of a Swedish reform, implemented in 2007–09, which suddenly made the payroll tax and the self-employment tax vary by age. The results, based on DiD estimations, suggest that youth self-employment is insensitive to tax changes. Both a 2007 cut and a 2009 cut left self-employment completely unaffected. The lack of treatment effects is precisely estimated and is robust to a battery of sensitivity tests. The fact that there was no effect in 2009–10, i.e. in the midst of recession, suggests that (large) tax cuts have no role to play even in times of economic slowdowns. I also study intensive margin responses. I first show that for the young self-employed, that faced the lower tax rate, income from self-employment increases by up to 20 percent on average. I then compare the estimated income effect to the predicted mechanical effect (due to the lower tax rate) and find that, in each year 2007–09, the estimated effect is greater. I argue that a potential explanation for this pattern is that the tax cut caused self-employed individuals to allocate more time to self-employment—either by reducing time in leisure or in wage employment. I find some support for the reallocation of time explanation: for those that are defined as self-employed, income from employment decreases due to the tax cut. Since the income effect is temporary, however, I conclude that the (potential) reallocation of working hours did not pertain to years when an individual no longer faces the lower tax rate. The finding of adjustments of working hours is in line with recent work showing significant intensive margin responses to tax changes.

References

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics*. Princeton University Press.

Bennmarker, H., E. Mellander, and B. Öckert (2009). Do regional payroll tax reductions boost employment? *Labour Economics 16*(5), 480–489.

Blanchflower, D. and B. Meyer (1994). A longitudinal analysis of the young self-employed in australia and the united states. *Small Business Economics 6*(1), 1–19.

Blundell, R., M. C. Dias, C. Meghir, and J. van Reenen (2004). Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association 2*(4), 569–606.

Bruce, D. (2000). Effects of the united states tax system on transitions into self-employment. *Labour Economics 7*(5), 545 – 574.

Bruce, D. (2002). Taxes and entrepreneurial endurance: Evidence from the self-employed. *National Tax Journal 55*(1), pp. 5–24.

Cahuc, P., S. Carcillo, and T. Le Barbanchon (2014). Do Hiring Credits Work in Recessions? Evidence from France. IZA Discussion Papers 8330, Institute for the Study of Labor (IZA).

Chetty, R., J. N. Friedman, and E. Saez (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review 103*(7), 2683–2721.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *American Economic Review 99*(4), 1145–77.

Cullen, J. B. and R. H. Gordon (2004). Taxes and entrepreneurial activity: Theory and evidence for the us. NBER Working Paper 9015, NBER Cambridge.

Domar, E. D. and R. A. Musgrave (1944). Proportional income taxation and risk-taking. *The Quarterly Journal of Economics 58*(3), 388–422.

Dunn, T. and D. Holtz-Eakin (2000). Financial capital, human capital, and the transition to self employment: Evidence from intergenerational links. *Journal of Labor Economics 18*(2), pp. 282–305.

Edmark, K., C.-Y. Liang, E. Mörk, and H. Selin (2012). Evaluation of the swedish earned income tax credit. Working Paper Series 2012:1, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Egebark, J. and N. Kaunitz (2013). Do payroll tax cuts raise youth employment? Working Paper Series 2013:27, IFAU - Institute for Evaluation of Labour Market and Education Policy.

European Comission (2013). *Youth Employment Package*. European Comission, Brussels.

Ferran, E. (2015). Labor demand elasticities over the life cycle: Evidence from spain's payroll tax reforms. Job market paper, Columbia University.

Gentry, W. M. and G. Hubbard (2003). Tax policy and entry into entrepreneurship. mimeograph, Columbia University.

Gentry, W. M. and G. Hubbard (2004). Success taxes, entrepreneurial entry, and innovation. NBER Working Paper 10551, NBER Cambridge.

Gruber, J. (1997). The incidence of payroll taxation: Evidence from chile. *Journal of Labor Economics 15*(3), S72–101.

Hansson, A. (2012). Tax policy and entrepreneurship: empirical evidence from sweden. *Small Business Economics 38*(4), 495–513.

Huttunen, K., J. Pirttilä, and R. Uusitalo (2013). The employment effects of low-wage subsidies. *Journal of Public Economics 97*(0), 49 – 60.

OECD (2012). Policy Brief on Youth Entrepreneurship: Entrepreneurial Activities in Europe. Technical report, OECD, Paris.

OECD (2013). Self-employment among the youth and seniors: Entrepreneurship at a Glance 2013. Technical report, OECD, Paris.

Oosterbeek, H., M. van Praag, and A. Ijsselstein (2010). The impact of entrepreneurship education on entrepreneurship skills and motivation. *European Economic Review 54*(3), 442 – 454.

Schuetze, H. J. (2000). Taxes, economic conditions and recent trends in male self-employment: a canada us comparison. *Labour Economics 7*(5), 507 – 544.

Skedinger, P. (2012). Tudelad trygghet. In A. Teodorescu and L.-O. Pettersson (Eds.), *Jobben kommer och går : behovet av trygghet består*, pp. 114–135. Stockholm: Ekerlid.

Skedinger, P. (2014). Effects of Payroll Tax Cuts for Young Workers. *Nordic Economic Policy Review* (forthcoming).

Stabile, M. (2004). Payroll taxes and the decision to be self-employed. *International Tax and Public Finance 11*(1), 31–53.

Statistics Sweden (2014). *Arbetskraftundersökningarna*. SCB, Stockholm.

A Appendix

A.1 Additional extensive margin results

Table A.1: Pooled treatment effects, using 27-year-olds as control group

	Full sample		Conditional	
	TG: 25	TG: 21–25	TG: 25	TG: 21–25
DD 07–08	0.022	–0.0004	0.033	–0.022
	(0.062)	(0.049)	(0.071)	(0.056)
$\overline{y}_{\mathrm{TG}}$	2.2	1.5	2.5	1.7
N	1,093,515	3,257,921	914,908	3,170,954
R^2	0.004	0.005	0.004	0.006

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. See also notes for table 3.

Table A.2: Yearly treatment effects, using 27-year-olds as control group

	Full sample		Conditional	
	TG: 25	TG: 21–25	TG: 25	TG: 21–25
DD 2006	–0.005	0.012	–0.051	–0.083
	(0.082)	(0.066)	(0.095)	(0.076)
DD 2007	0.086	0.061	0.11	0.021
	(0.082)	(0.065)	(0.095)	(0.074)
DD 2008	–0.056	–0.054	–0.081	–0.12
	(0.083)	(0.067)	(0.096)	(0.076)
$\overline{y}_{\mathrm{TG}}$	2.2	1.5	2.5	1.7
N	1,093,515	3,257,921	914,908	2,714,310
R^2	0.004	0.005	0.004	0.006

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. See also notes for table 4.

Table A.3: Pooled treatment effects, using different definition of self-employment

	Full sample		Conditional	
	TG: 25	TG: 21–25	TG: 25	TG: 21–25
DD 07–08	0.0004	–0.026	–0.002	–0.077
	(0.070)	(0.041)	0.08175	(0.048)
$\overline{y}_{\mathrm{TG}}$	3.1	2.2	3.4	2.5
N	1,091,071	3,804,593	909,945	3,170,954
R^2	0.004	0.006	0.004	0.007

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. See also notes for table 3.

Table A.4: Treatment effects by own and parents' characteristics

Panel A: Own characteristics

	Full sample, 21–25 vs. 26–27					
	Men	Women	Natives	F-born	Voc/broad	Voc/narrow
DD 07–09	–0.0047	–0.015	–0.012	–0.090	–0.142	–0.087
	(0.057)	(0.041)	(0.039)	(0.080)	(0.080)	(0.151)
$\overline{y}_{\mathrm{TG}}$	1.9	1.1	1.6	1.6	2.0	1.6
N	1,943,792	1,860,801	3,168,643	635,950	1,04,631	291,316
R^2	0.004	0.002	0.005	0.003	0.005	0.006

Panel A: Parents' characteristics

	Full sample, 21–25 vs. 26–27			Conditional, 21–25 vs. 26–27		
	Self-empl.	Inc.> Md	Top qt inc.	Self-empl.	Inc.> Md	Top qt inc.
DD 07–09	0.027	–0.019	–0.030	–0.019	–0.041	–0.060
	(0.133)	(0.041)	(0.049)	(0.143)	(0.048)	(0.061)
$\overline{y}_{\mathrm{TG}}$	3.4	1.3	1.3	3.5	1.4	1.5
N	556,753	2,474,792	1,637,961	501,506	2,046,101	1,284,848
R^2	0.19	0.13	0.13	0.20	0.13	0.14

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. Voc/broad includes all individuals with vocational training. Voc/narrow consists of construction workers. See also notes for table 6.

A.2 Additional intensive margin results

Table A.5: Yearly treatment effects, 2005-2009. Cluster robust standard errors

	Income from self-employment (SEK)			
	TG: 25	TG: 23–25	TG: 21–25	TG: 26
DD 2005	1,720	290	890	–1,740
	(990)	(1,610)	(1,700)	(1,380)
	[640]	[680]	[840]	[1,190]
DD 2006	–370	63	910	760
	(1,340)	(1,640)	(1,760)	(1,640)
	[650]	[680]	[840]	[1,180]
DD 2007	3,065***	5,390***	8,130***	–1,095
	(960)	(1,480)	(1,930)	(2,670)
	[640]	[670]	[830]	[1,180]
DD 2008	9,730***	13,150***	13,060***	1,840
	(820)	(1,600)	(2,000)	(1,260)
	[650]	[680]	[840]	[1,190]
DD 2009	16,600***	17,240***	17,930***	13,300***
	(1,850)	(1,970)	(1,990)	(2,089)
	[650]	[680]	[850]	[1,190]
$\overline{y}_{\text{TG}}$	81,350	79,730	76,990	85,120
N	103,522	140,350	162,721	107,802
R^2	0.018	0.017	0.018	0.018
No. of clusters	(24) [16]	(40) [16]	(56) [16]	(24) [16]

Notes: $***p < 1\%$, $**p < 5\%$, $*p < 10\%$. Standard errors clustered w.r.t. $Age \times Year$ in parenthesis. Standard errors clustered w.r.t. $Treatment \times Year$ in brackets. Significance levels refer to standard errors in parenthesis. See also notes for table 7.

Table A.6: Income effects by own and parents' characteristics

Full sample, 21–25 vs. 28–29

	Own characteristics				Parents' characteristics		
	Men	Women	Natives	F-born	Self-empl.	Inc.> Md	Top qt inc.
DD 07–09	12,910***	11,217***	15,305***	–217	13,590***	12,700***	11,900***
	(1,600)	(1,870)	(1,360)	(2,910)	(2,510)	(1,530)	(1,860)
$\overline{y}_{\mathrm{TG}}$	79,460	72,300	79,000	67,330	78,790	77,540	73,063
N	108,384	54,337	137,847	24,874	42,051	93,764	64,103
R^2	0.02	0.01	0.02	0.02	0.02	0.02	0.02
Effect size:	19%	18%	24%	—	21%	20%	19%

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. See also notes for table 8.

Can Indifference Make the World Greener?*

Johan Egebark† Mathias Ekström‡

Abstract

We conducted a natural field experiment at a large public university in Sweden to evaluate the causal effect of two resource conservation programs. Our first intervention consisted of a campaign that actively tried to convince people to cut back on printing in general, and to use double-sided printing whenever possible. The second intervention exploited people's tendency to stick with pre-set alternatives. At random points in time we changed the printers' default settings, from single-sided to double-sided printing. Whereas the moral appeal treatment had no impact, not even in the short run, the default change cut daily paper use by 15 percent. Two important pieces of evidence complement the basic default rule result. First, paper consumption was still at the new lower level more than six months after the change, which shows that default effects can last for longer periods of time. Second, printing demand was completely independent of the pre-set alternative, suggesting that "green defaults" do not affect people's intrinsic motivation to save resources.

Key words: Resource Conservation; Default Option; Moral Appeal; Natural Field Experiment

JEL classification: C93; D03; Q50

*We want to express our gratitude to Stefano DellaVigna, Peter Fredriksson, Magnus Johannesson, Erik Lindqvist, Bertil Tungodden, and Robert Östling, as well as numerous conference and seminar participants, for helpful discussions and valuable comments. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. All errors are our own.

†Department of Economics, Stockholm University, and the Research Institute of Industrial Economics (IFN). E-mail: johan.egebark@ne.su.se

‡Department of Economics, Norwegian School of Economics, and the Research Institute of Industrial Economics (IFN). Email: mathias.ekstrom@nhh.no

1 Introduction

Depletion of natural resources, such as water, food and forests, constitutes a severe threat to the global environment, and to sustained economic growth (UN, 1992; EU, 2011). In order to address this threat, policy makers need to understand what type of actions will have an impact on resource conservation, and what actions are ineffective. In recent years, so called nudges—non-price based behavioral interventions that preserve choice—have gained increasing attention, both from academic scholars and among politicians (Thaler and Sunstein, 2003; Sunstein and Thaler, 2003; Shafir, 2012). It has been argued that nudges have a particularly important role to play within the environmental domain, not the least since regular price adjustments have been difficult to implement (Allcott and Mullainathan, 2010; Carlsson and Johansson-Stenman, 2012; Sunstein and Reisch, 2014). However, empirical evidence is lagging behind, and to the extent that it exists the results are mixed. Allcott and Rogers (2014), for example, show that using feedback to reduce energy consumption affects behavior, but only in the short run. Furthermore, Kallbekken et al. (2013) find that disclosing information on lifetime energy costs can affect purchases of durable goods, but only for certain product categories and when the information is coupled with training of the sales staff.

We add to this growing literature by studying the causal effect of two different behavioral interventions aimed at lowering the consumption of paper. The activity that we consider is universal, frequent and consumes a vast amount of resources every year: document printing. Estimates suggest that U.S. office workers use roughly five million metric tons of paper annually, amounting to around 20 million metric tons of wood. To grasp the potentials, reducing this amount by only five percent would save roughly six and a half million trees, free 6,500 acres of forest for other productive use, and prevent the equivalent annual greenhouse gas emissions of 140,000 cars.

When sending a document to a printer, a user can typically choose whether to print on both sides of a sheet of paper (duplex) or to print on only one side (simplex). Duplex printing reduces the number of sheets in the production of a document and is thus less resource intensive. We use this functionality in a field experiment at a large Swedish university to evaluate two different paper saving programs. In the active approach we test whether moral appeal increases environmental responsibility. In collaboration with the university we designed an e-mail campaign that encouraged employees to cut back on paper use in general, and to use duplex printing whenever possible. The second intervention constitutes a more passive approach. A common feature of most modern printers is the existence of a default option, i.e., the alternative people obtain when not actively making a choice. By randomly changing the default, from simplex to duplex printing, we test whether people's tendency to stick to the pre-set alternative can help save resources.

There are several motives for studying these two interventions in particular. We can see at least two major reasons why it is important to continue to study default rules. First, while seminal studies documented substantial effects on retirement savings (Madrian and Shea, 2001) and organ donation (Johnson and Goldstein, 2003), recent experiments have shown that default effects do not necessarily generalize to the environmental domain (Löfgren et al., 2012), nor to different types of samples (Coppen et al., 2005; Bronchetti et al., 2013). In fact, Löfgren et al. (2012) is (to the best of our knowledge) the only published paper that experimentally tests to what extent defaults influence people to act pro-environmentally. They find that when participants at a conference in environmental and resource economics are asked whether they want to offset travel-related CO_2-emissions or not, the pre-set alternative turns out to be completely irrelevant. This clearly shows that there is room for more evidence from experiments conducted in the field. Second, given that default options have been proposed as a potential policy tool (Sunstein and Reisch, 2014) it is important to understand the long-term impacts, and the possibility of unintended negative impacts. To illustrate the second point, Li et al. (2014) show that introducing a healthy default option in a restaurant setting could reduce the number of customers and sales.

The moral appeal treatment is included for two main purposes. First, it provides a natural baseline: it is the intervention first chosen by the university, and similar messages are commonly used in many everyday settings. Examples include appeals to hotel customers to reuse towels during their stay, and reminders about garbage collection in public parks. The common use of moral appeal messages suggests that policy makers perceive them as effective. The second reason for including the moral appeal treatment is that it will show to what extent a default effect is explained by implicit recommendation. While endorsement is sometimes mentioned as a potential explanation in previous studies on default effects, little is known empirically about its relative importance (for exceptions, see McKenzie et al., 2006; Altmann et al., 2013).

The effect of the two interventions differed sharply. We document a substantial and immediate effect of changing the default printer setting. On average, daily paper consumption dropped by 15 percent due to the changed settings, and this reduction occurred already on the day that we introduce the intervention. Put differently, the default option determined how one third of all documents were printed. In contrast, using moral appeal to encourage people to take responsibility had no effect, not even on the day the message was communicated. These results clearly show the importance of carefully choosing "no-action" options, and that passively taking advantage of people's (lack of) preferences can be much more effective than actively trying to change them. In a broader perspective, the failure of the appeal to affect actions speaks against strategies that rely on persuasive communication to make people save on resources (see, e.g., Costanzo et al., 1986; Stokes et al., 2012).

We continue by looking into some important, but often neglected, aspects of default rules. While many behavioral interventions, such as feedback and social comparison, often have significant short-run effects, there is still limited evidence on the impacts in the longer run (Allcott and Rogers, 2014). We show that default rules can be equally effective also in the (somewhat) longer run, as paper consumption was still at the new lower level more than six months after the intervention. Second, we show that printing demand (measured by the number of printed pages and documents) is independent of the pre-set alternative. This finding is insightful as it indicates that changing defaults avoids unintended adverse effects. A relevant and contrasting comparison is Catlin and Wang (2013), who report an increase in paper use when the possibility to recycle is introduced. Similarly, it has been shown that providing information about neighbors' energy use has the unintended consequence that low-consuming households respond by using more energy (Schultz et al., 2007; Ayres et al., 2009). Third, it is sometimes argued that default interventions could have a negative impact on people's welfare since they may have to spend time (and effort) on opting out of new settings they do not like. We show that the intervention in this case was in fact welfare enhancing: the fraction of users that opt out is larger when simplex is the default, implying that more people were hurt in the simplex regime.

Finally, to better understand what is driving the default effect in our study, and what explains the lack of treatment effects for the moral appeal, we recruited employees from the experiment to participate in an *ex post* survey. It is not the case that the substantial default effect is explained by employees not knowing how to opt in or out of the respective settings, since, when asked, 97 percent state that they know how to do this. A more compelling explanation, instead, is that many users have weak preferences over the two alternatives, and therefore stick with the pre-set alternative to avoid a tiny switching cost. This explanation is supported by the survey data, as 60 percent of the men and 50 percent of the women are indifferent between simplex and duplex printing. For the moral appeal treatment, we can reject that the null result is due to inattention since 83 percent of the respondents claimed that they had read the e-mail. What is even more intriguing is that employees do not seem to dislike the appeal: 75 percent say that they believe it will affect their colleagues' behavior, and 35 percent state that it will affect their own way of printing.

2 The experiment

2.1 Two interventions

We conducted a natural field experiment at a large public university in Sweden to measure the effect of two paper saving programs. The first intervention constitutes a passive approach to behavior change. Over a period of three months we changed the default option on printers at the university, from simplex to duplex printing, with random timing of the intervention. Theory predicts that consumption of paper is completely independent of whether simplex or duplex printing is set as the pre-set alternative, as long as the cost of opting out is small.[1] However, research has shown that people are biased towards the status quo (Samuelson and Zeckhauser, 1988); thus, a careful choice of the "no-action" alternative might nudge people to save resources.

The second intervention consisted of a message that actively tried to convince employees at the

[1] The print screens of the pop-up window included in the appendix show how easy it is to opt in and out of the default option.

university to reduce their use of paper. The moral appeal was communicated via e-mail and was signed by the environmental coordinator at the university.[2] It encouraged people to participate in the strive to reduce the university's impact on the environment by cutting back on printing in general and by using duplex printing whenever possible. The employees were also reminded of how easy it is to use the duplex printing mode (an English version of the e-mail is included in the appendix).

The moral appeal treatment serves two main purposes. First, since it is the standard tool in many everyday settings—one example is reminders about garbage collection in public parks—it is interesting to evaluate in its own right. Second, it helps provide a better understanding of the mechanisms behind a default effect in our setting. It is often argued that inertia is the main reason for why defaults are influential. However, there are many other potential explanations, and we want to address two of them. We cannot be sure that people are actually aware of the green alternative, i.e., the possibility to use duplex on their specific printer. Hence, a default effect could fully or partly be explained by ignorance. Furthermore, a pre-set default option may be interpreted as advice or as an implicit norm about how to behave (Madrian and Shea, 2001). Changing the default option to duplex might therefore affect people because it alters their notion about what is the right thing to do. The message treatment helps address these two mechanisms: if there is no behavioral change of the message, we can rule out both ignorance and norm compliance as drivers. The moral appeal treatment was not included to provide a deeper understanding of what type of messages is more effective than others (for readers interested in this topic we refer to Schultz et al. 2007; Goldstein et al. 2008; Fellner et al. 2013; Pruckner and Sausgruber 2013; Dwenger et al. 2014). We do note, however, that Fellner et al. (2013) report that a one-shot letter can have a sizeable impact on public good compliance, irrespective of the content in the letter.

[2]We deliberately had a person with some authority send the e-mail since previous evidence suggests that source credibility is important when using persuasive communication DellaVigna and Gentzkow (2010).

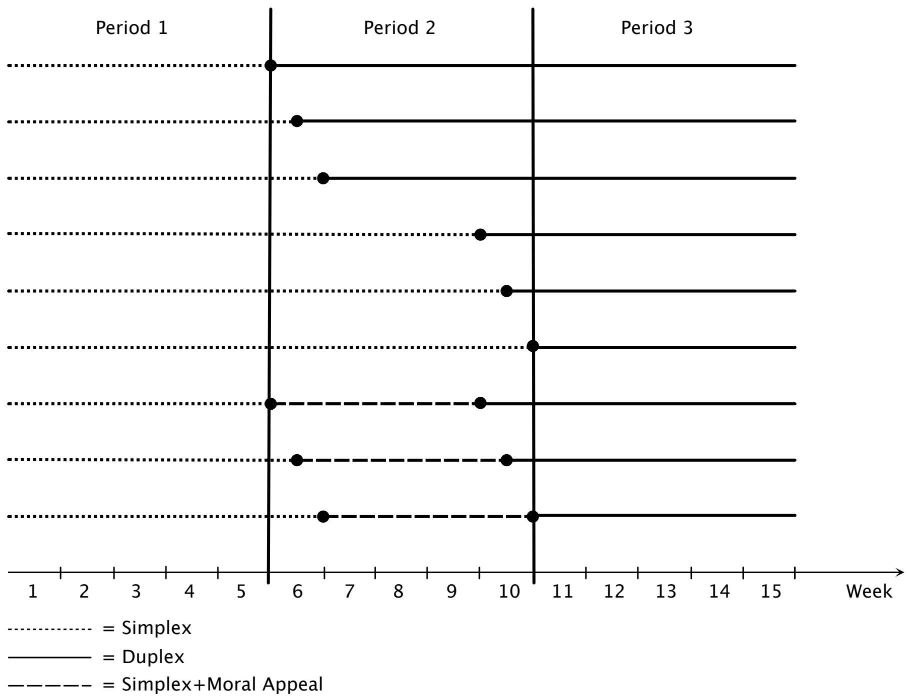
2.2 Design

To implement the default change intervention, we asked the IT-support at the university for a list of all printers that reported sufficient statistics online. This original sample consists of 54 printers distributed across 31 different departments. 19 of the departments had at least one printer that allowed for a change in the default settings, from simplex to duplex. We contacted the heads of these 19 departments directly, asking whether they would be willing to participate in the study. The heads were informed about the intention of the experiment and that participation required the default change to be communicated to staff from within the department. (The e-mail employees received about the change is found in the appendix.) We stressed the importance of not informing any of the employees about the experiment, as this could ruin the benefits of conducting a natural field experiment as defined by Harrison and List (2004). 18 out of 19 departments accepted these terms and became part of the study. The high compliance rate suggests that the departments had not chosen their default settings in a rational and informed way; rather, it seems as if they stuck to the default which was set when the printer was first installed and put to work. The 18 departments had in total 25 printers which we use in the experiment.

Figure 1 illustrates the experimental design. The experiment consisted of three periods spread evenly over 15 weeks. We started with a pre-treatment phase of five weeks when all printers had simplex as default, and we ended with a post-treatment phase of five weeks when all printers had duplex as default. Between these, there was a treatment phase of five weeks, in which the 18 departments were randomly assigned to one of six different pre-determined intervention dates.[3] A random subset of the departments was, prior to their default change, exposed to the moral appeal treatment. The message went out to all employees (300+) in these departments, and was sent on different days in the first week of the treatment phase, allowing for a four-week evaluation period.

[3]For departments that participate with more than one printer it would be strange to repeat the procedure, especially if there is little space in time between the interventions. For that reason we used the department, and not the printer, as the unit of randomization.

Figure 1: Experimental design



2.3 Identification strategy and data

The staggered passage of changed print settings means that our experiment constitutes a controlled event study: all printers are used for the default treatment, but the timing of the intervention is random. Given the design, we measure the default effect by using all 25 printers, normalize time to zero, and compare mean outcomes before and after the intervention. Identification is more credible if there is a clear and visible shift in the outcome that appears close in time to the intervention. Since treatment occurs at different points in time, depending on the printer, a potential treatment effect will not be confounded by other time events. The moral appeal treatment is studied in the same fashion, except that we use a smaller sample of printers.

Data were collected for each printer on a daily basis, implying that the unit of observation is printer and day. The two outcomes of primary interest are the ratio of sheets to pages, and the number of printed sheets. The first outcome measures how efficiently paper is used whereas the second measures how much paper is actually saved. We are also interested in the number of printed document pages to account for the possibility that the respective treatments affect printing demand. Table 1 gives some basic pre-treatment statistics, for all the departments combined, for the sample of departments that did not receive the message, and for the sample of departments that did receive the message. Looking at the full sample figures, we note that the average printer uses 170 sheets of paper to print 221 document pages on a typical day. This translates to a ratio of sheets to pages of 0.85, and so there is considerably more simplex printing than duplex printing in the pre-treatment period. In fact, less than one third of all printed sheets are duplex sheets. We also note that there are some slight differences between the samples in columns 2–3, which is expected considering the low number of departments in each sample (12 and 6, respectively).

2.4 Survey

After collecting the data, we decided to survey a random subset of the employees that participated in the experiment. The survey was carried out to collect information about people's expressed printing preferences, their general knowledge about default settings, and how they perceived the

Table 1: Five week pre-treatment averages

	Full sample	No message	Message
SP ratio	0.85	0.87	0.83
# Sheets	170	143	204
# Pages	221	180	274
# Duplex sheets	51	37	69
# Simplex sheets	119	107	135
N	585	337	248
# Printers	25	14	11
# Departments	18	12	6

moral appeal message. We contacted 249 employees in total and asked whether they wanted to participate in a web survey about printing habits. 119 employees finished the survey, which means that the response rate was 48 percent. To be able to include questions about the moral appeal message, we made sure that the environmental coordinator communicated the exact same message that we used in the experiment two days before we sent out the survey. Table A.1 included in the appendix shows the survey responses, for the entire sample and separately for men and women. We will refer the reader to the table when we elaborate on the results from the main experiment. We do note, however, that 97 percent state that they know how to opt in and out of the default print settings, and that 83 percent claimed that they read the e-mail containing the moral appeal message.

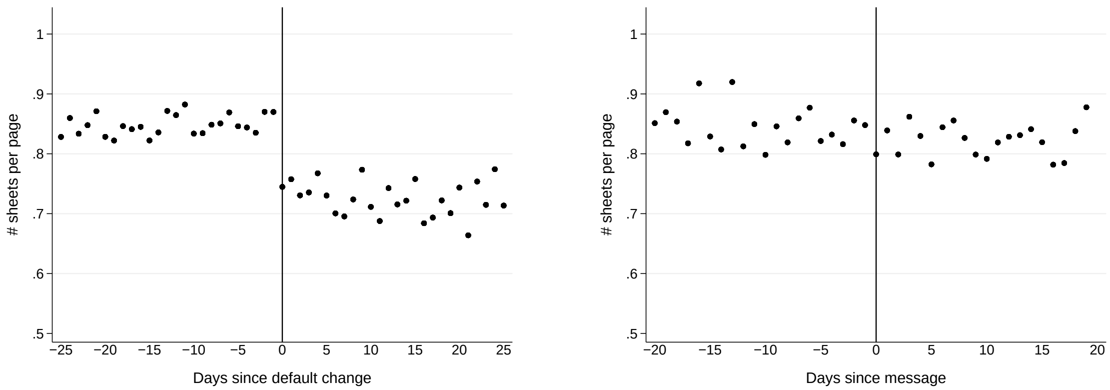
3 Results

3.1 Main findings

We begin by studying the effect on resource efficiency. Figure 2 shows daily means of the ratio of sheets to pages, before and after the respective treatments. The lower the ratio, the more duplex printing there is.[4] The leftmost scatter plot shows the immediate default effect. The very day that we change the default settings, the ratio drops from around 0.85 to 0.73, and once the new level is

[4]Documents with only one page will have a ratio of one irrespective of the printing mode. This implies that the lower bound is in practice strictly larger than 0.5 (document-level data from two printers in our sample suggest it is about 0.55). To ensure a balanced panel, we restrict the time window to five (four) weeks before and after the default change (moral appeal message).

Figure 2: The effect on resource efficiency



(a) Default treatment

(b) Moral appeal treatment

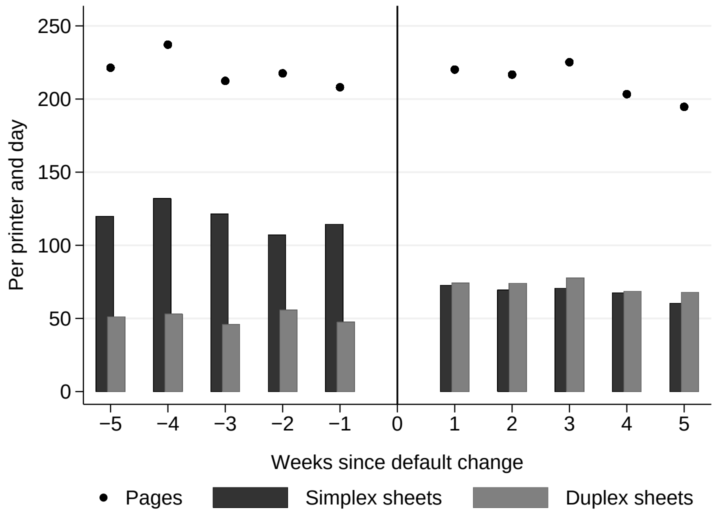
reached it holds constant. A regression of the outcome on an indicator variable for the treatment period confirms that the difference in daily means is highly significant: $t(1171) = 6.26, p = 0.000$.[5] In sharp contrast, there is no visible impact of the moral appeal—not even on the day the message was communicated. Regression confirms that there is no statistically significant effect in figure 2 (b): $t(388) = 1.45, p = 0.147$.

The change in the default settings clearly improved resource efficiency. However, this does not automatically imply that resources were saved due to this intervention, since people may adapt their behavior and start consuming more paper (Khazzoom, 1980). It turns out that the number of printed pages is completely unaffected by the new default settings, as seen in figure 3. In terms of a formal test, the difference in means before and after the intervention is far from significant: $t(1249) = 0.53, p = 0.593$. This finding is important as it shows that users did not respond by printing more (or less) document pages; we hence conclude that printing demand is independent of the default setting.[6] Instead, it is the relation between simplex and duplex printing that has

[5] The t-statistics in this section come from regressions of the respective outcomes on a dummy variable that equals one for the first 25 days after a printer's default change and zero for the 25 days prior to the change. Given that we have at most 25 clusters, we performed the regressions using both robust and cluster robust standard errors (by printer) and report t-statistics for the more conservative of the two.

[6] Estimates from the two printers in the sample that report document-level data confirm that the length of documents, and the number of documents are also unaffected.

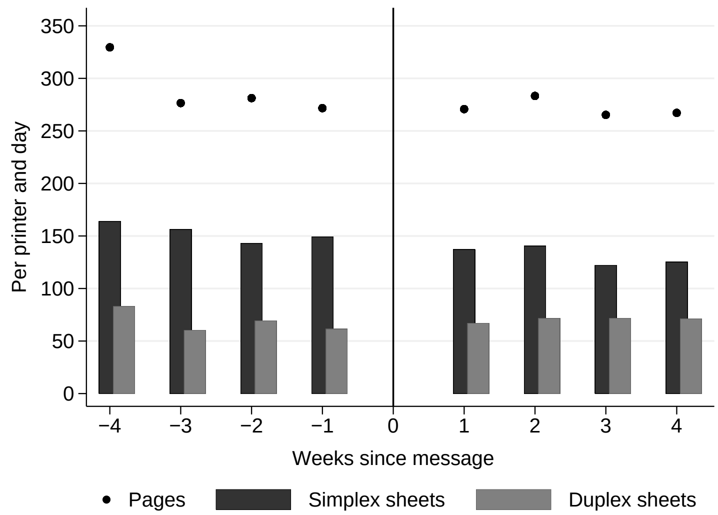
Figure 3: Default effect details



shifted (gray and white bars). Both the reduction in the number of simplex sheets ($t(1249) = 4.29, p = 0.000$) and the increase in the number of duplex sheets ($t(1249) = 2.75, p = 0.011$) are highly significant. We also note that there is still a substantial amount of simplex printing after the duplex default has been implemented, suggesting that it is easy for people to opt out if they wish. The increase in duplex printing implies that paper consumption is substantially lower after the settings were changed; the number of sheets that is used per day has dropped by 15 percent on average ($t(1249) = 2.45, p = 0.014$). Turning to the moral appeal treatment in figure 4, there is further evidence that asking people to use duplex printing has no impact on behavior. In particular, comparing bars on either side of the vertical line there is no drastic shift in the relation between duplex and simplex printing (the conclusion is confirmed by t-tests). This null result can be contrasted to the responses in the survey, where 35 percent of respondents said that the message would affect their own printing behavior, and 75 percent believed it would affect their colleagues at the university (see rows 2–3 in table A.1).

While many behavioral interventions, such as feedback and social comparison, often have large short-run effects, there is limited evidence on the impacts in the longer run (Allcott and Rogers,

Figure 4: Appeal effect details

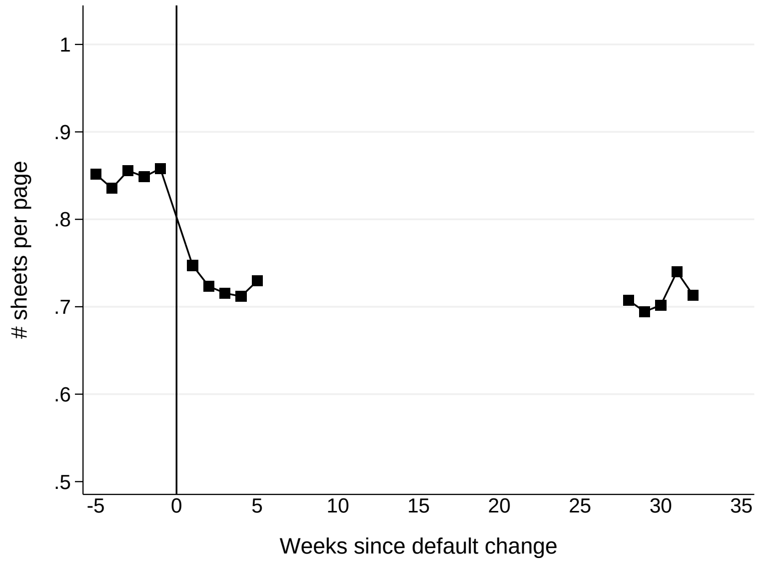


2014). An important question, therefore, is whether the default effect that we observe endures over time. To investigate this, we collected additional data from the same set of printers, during a second consecutive five-week period, commencing 28 weeks after the intervention. Figure 5 plots weekly averages of the ratio of sheets to pages. Strikingly, more than six months after the change, the number of sheets per page is at the same low level as in the very short run. The fact that people did not revert back to simplex printing shows that default rules can be powerful also in the longer run.

3.2 Effect heterogeneity

This section takes a closer look at default effect heterogeneity. We first look at differences with respect to the type of department. There are both academic departments and administrative departments in the sample, with a 50 percent share of both. Heterogeneity in this dimension may inform us about the generalizability of the results: if the effect is mainly driven by academic departments, the results may be confined to highly educated people only; if the effect, on the other hand, is mainly driven by administrative departments, the results may be of more relevance for,

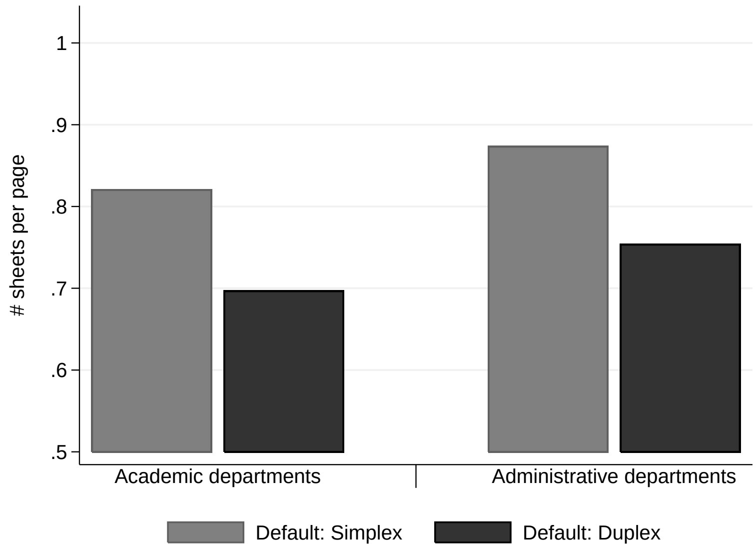
Figure 5: Long run default effect



say, government agencies. As seen in figure 6, the default effect is independent of the type of department (the point estimate from a difference–in–differences regression is 0.004, and it is not statistically significant: $t(1171) = 0.10, p = 0.918$). This result is consistent with a more general effect applicable to different types of workplaces.

We continue by considering the timing of the effect. Figure 7 shows effects separately for each of the three treatment groups that we use (see figure 1 in section 2.2). The first group (SDD) changed to duplex in the beginning of the second period (during week 6), whereas the two other groups (SSD and SMD) changed to duplex in the beginning of the third period (during week 10). The SMD group, further, is the one where the employees received the moral appeal message in the second period (before changing to duplex). As expected, there is a significant drop in the number of sheets per page for the SDD group during the second period, and for the SSD group during the third period. For SMD we expect no drop in the second period as the moral appeal treatment had no effect. In the third period, however, we expect the default effect to kick in, and it should be of similar size as for the two other groups. Notably, this is exactly what figure 7 shows. The fact that the default effect is equally large for SMD as for the other two groups shows that the employees who

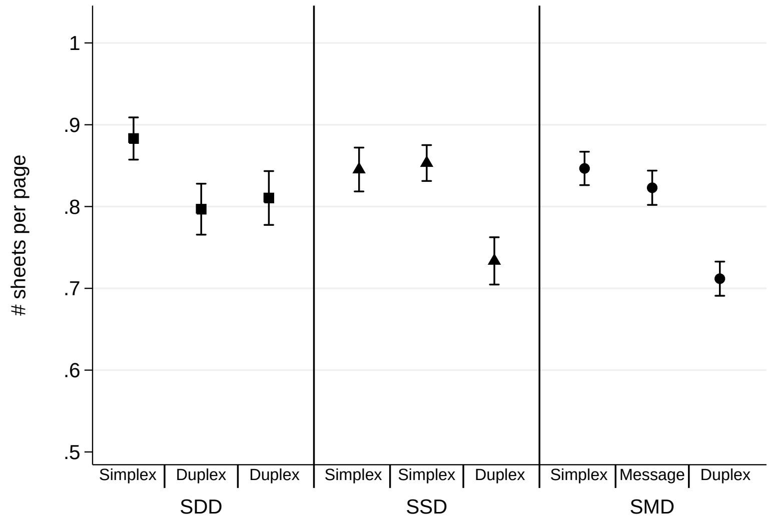
Figure 6: Default effect by department type



were subjected to the moral appeal message were capable of printing more duplex if they wanted. It also suggests that the default effect is not explained by inattention, or by implicit recommendation. From the figure we also draw the conclusion that there are no underlying time trends in printing behavior, as the outcome is stable for SDD (SSD) between period 2 and 3 (1 and 2).

We finally make use of two printers in the sample that report document-level data to assess whether the default effect varies by individual characteristics. We first divide the effect by gender. As shown by the dashed line in figure 8 (a), women are roughly 50 percent more likely to use duplex after the default intervention. What is more intriguing, however, is that the effect is substantially larger for men, as shown by the steeper slope of the solid line in the figure. It is striking to see men's strong tendency to stick to the current default regime: when simplex printing is the default, men use simplex printing; when duplex is the default, they use duplex. This tendency is much less pronounced for women. Using regression analysis, we have also confirmed that the difference in

Figure 7: Timing of the default effect



behavior between men and women is statistically significant.[7] The next graph in figure 8 looks at the importance of age in explaining the default effect. An employee is defined as "old" if his age is at or above the sample median, which is 46. The point estimates in figure 8 (b) suggest that the default effect is larger for employees above the median age, but the age difference is not statistically significant at conventional levels.

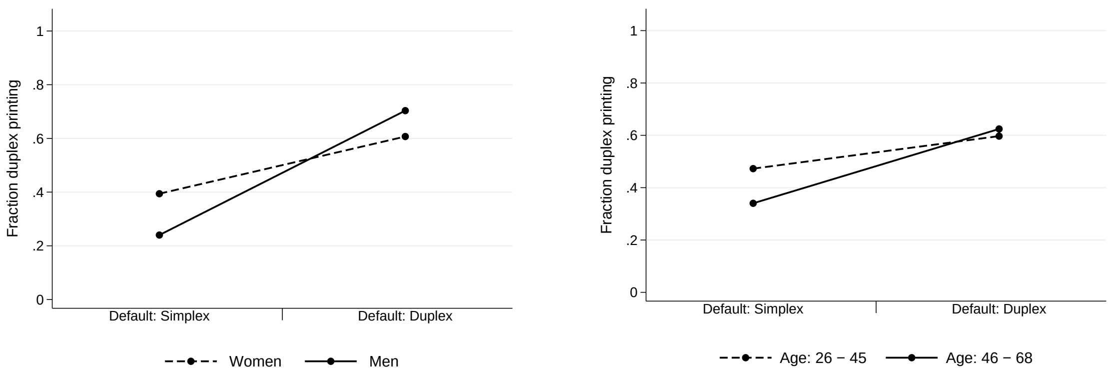
The above results suggest that, even though males (and the elderly) are not opposed to improvements in resource efficiency, they are less likely to *actively* take advantage of it. This could in turn reflect that male employees simply care less about the presentation of a document, or that they are less technologically savvy (and hence do not know how to opt out from the default option). To assess the different explanations we compare survey responses for men and women. Both men

[7]To confirm gender differences we estimate a linear probability model of the following form:

$$y_{i,t} = \alpha + \beta_1 After_t + \beta_2 Male_i + \beta_3 Male_i \times After_t$$

where $y_{i,t}$ is a dummy variable that equals one if individual i used duplex to print a specific document at time t, $After$ is a time dummy that equals one in the period after the default change, and $Male$ is a gender dummy that equals one if the employee is a man. The coefficient of interest, β_3, measures to what extent men are affected more than women by the default change (i.e., if they have a higher propensity than women to use duplex in the period after the change). When analyzing the importance of age we run the same model but with a dummy for age instead of gender. To account for serial correlation we clustered standard errors at the employee-level.

Figure 8: Default effect by gender and age



(a) (b)

and women know how to opt in and out of the default settings (men: 95%; women: 97%) and both groups perceive duplex printing as pro-environmental (men: 98%; women: 97%). Hence, there is no difference in technological skills or attitudes. When asked about their printing preferences (and their willingness to pay to get their preference implemented), 60 percent of the men are (classified as) indifferent between using simplex and duplex, against 50 percent of the women.[8] Men are also more likely to state that the default alternative affects them (Men: 55%; Women: 47%). Overall, the survey responses suggest that men tend to have weaker preferences than women, and that indifference seem to be a plausible explanation for the results in general.

3.3 Welfare

One motivation for implementing a default option in any setting is that it is Pareto improving: those who prefer the default alternative and thus save on time and effort are made better off, while those who nevertheless have to choose actively are unaffected. Choosing which alternative to use as the default is, however, nontrivial. We have already shown that changing the default option to duplex

[8]A person is classified as indifferent if he or she answered that it does not matter whether a document is printed simplex or duplex, or if they stated a preference but would not pay a positive amount to get the printing preference implemented in a situation where the opposite printing alternative was free of charge.

did, indeed, save resources. However, we should also take into account the fact that the intervention may have inconvenienced people. That is, even though it was good for the environment, it may have increased the time and effort employees spent on printing. To analyze this aspect we look at the opt-out behavior in the different default regimes. In figure 2 it is clear that the number of sheets per page drops from around 0.85 to 0.7 as we change the default from simplex to duplex. This suggests that one third opts out when simplex is the default $[(1 - 0.85)/0.45 = 1/3]$ and that one third opts out when duplex is the default $[(0.70 - 0.55)/0.45 = 1/3]$. It is therefore reasonable to assume that the fraction of users who prefer duplex is roughly similar to the fraction of users who prefer simplex. We have also looked closer at individual-level data for two of the printers in the sample to grasp whether the aggregate data grossly overestimate the share of people with duplex preferences. To the contrary, 88 percent of the employees in this sample either have preferences for duplex printing or are what we refer to as indifferent (i.e., they follow the default option). A second important observation is that people do not revert back to simplex printing in the longer run, as they eventually make more active decisions. Survey data supports these findings. 77 percent of the respondents state that they prefer, or strongly prefer, duplex printing, whereas only 13 percent state that they prefer to use simplex. Furthermore, 95 percent say they would use duplex as the default if they were in charge of the printers at the department. In sum, we conclude that the change in the default was, in fact, welfare enhancing.

4 Conclusion

We conduct a natural field experiment at a large public university to measure the causal effect of two paper saving programs. Changing printers' default option, from simplex to duplex printing, reduced paper consumption by as much as 15 percent. The default effect was immediate and remained intact more than six months after the intervention. An intriguing finding is the absence of behavioral responses pertaining to the demand for printing. This result indicates that defaults may offer an attractive benefit as a policy tool by avoiding unintended adverse effects. On the other hand, the fact that people did not respond actively to reduce (overall) consumption could imply that green defaults induce pro-environmental behavior without affecting people's environmental awareness. This, however, needs to be further explored. The second intervention relied on moral appeal to convince employees to cut back on paper use. In contrast to the substantial default effect, this policy had no effect at all, not even in the very short run. We do, of course, not claim that persuasive communication is useless in general, but the null result highlights that there are boundaries to what pro-environmental reminders and campaigns can achieve.

A compelling explanation for the default effect in our setting is that users have weak preferences over the two alternatives and therefore stick with the default option to avoid a tiny switching cost. Presumably, most people do not take the environment into account when making this decision inasmuch as their isolated action will have little (global) impact anyway. There are many other situations where the same logic applies (e.g., turning off the lights when leaving a room or leaving electronic devices on standby). In such situations, there is a potential in carefully choosing, and creating, "no-action" options. Policy makers should be particularly supportive of this type of intervention since a socially preferred allocation can be reached without spending money and without limiting individual choice (Thaler and Sunstein, 2003). One setting that our results can be applied to directly is so-called go-paperless-initiatives by for example banks, government agencies and telephone operators. Another possibility is to stimulate suppliers, through mandates or incentives, to preprogram home appliances with energy efficient default options.

References

Allcott, H. and S. Mullainathan (2010). Behavior and energy policy. *Science 327*(5970), 1204–1205.

Allcott, H. and T. Rogers (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review 104*(10), 3003–37.

Altmann, S., A. Falk, and A. Grunewald (2013). Incentives and information as driving forces of default effects. IZA Discussion Papers No. 7610.

Ayres, I., S. Raseman, and A. Shih (2009). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. NBER Working Paper 15386.

Bronchetti, E. T., T. S. Dee, D. B. Huffman, and E. Magenheim (2013). When a nudge isn't enough: Defaults and saving among low-income tax filers. *National Tax Journal 66*(3), 609–634.

Carlsson, F. and O. Johansson-Stenman (2012). Behavioral economics and environmental policy. *Annu. Rev. Resour. Econ. 4*(1), 75–99.

Catlin, J. R. and Y. Wang (2013). Recycling gone bad: When the option to recycle increases resource consumption. *Journal of Consumer Psychology 23*(1), 122–127.

Coppen, R., R. D. Friele, R. L. Marquet, and S. K. M. Gevers (2005). Opting-out systems: No guarantee for higher donation rates. *Transplant International 18*(11), 1275–1279.

Costanzo, M., D. Archer, E. Aronson, and T. Pettigrew (1986). Energy conservation behavior: The difficult path from information to action. *American Psychologist 41*(5), 521 – 528.

DellaVigna, S. and M. Gentzkow (2010). Persuasion: Empirical evidence. *Annual Review of Economics 2*(1), 643–669.

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2014). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in germany.

EU (2011). *A Resource-efficient Europe – Flagship Initiative Under the Europe 2020 Strategy.* European Union.

Fellner, G., R. Sausgruber, and C. Traxler (2013). Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association 11*(3), 634–660.

Goldstein, N. J., R. B. Cialdini, and V. Griskevicius (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research 35*, 472 – 482.

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic Literature 42*(4), 1009–1055.

Johnson, E. J. and D. Goldstein (2003). Do defaults save lives? *Science 302*, 1338–1339.

Kallbekken, S., H. Sælen, and E. A. T. Hermansen (2013). Bridging the energy efficiency gap: A field experiment on lifetime energy costs and household appliances. *Journal of Consumer Policy 36*(1), 1–16.

Khazzoom, J. D. (1980). Economic implications of mandated efficiency in standards for household appliances. *The Energy Journal 1*(4), 21–40.

Li, M., H. Colby, and G. Chapman (2014). Do defaults change what people eat? dietary defaults and their boundaries. mimeograph.

Löfgren, Å., P. Martinsson, M. Hennlock, and T. Sterner (2012). Are experienced people affected by a pre-set default option—results from a field experiment. *Journal of Environmental Economics and Management 63*, 66–72.

Madrian, B. C. and D. F. Shea (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics 116*, 1149–1187.

McKenzie, C. R., M. J. Liersch, and S. R. Finkelstein (2006). Recommendations implicit in policy defaults. *Psychological Science 17*(5), 414–420.

Pruckner, G. J. and R. Sausgruber (2013). Honesty on the streets: A field study on newspaper purchasing. *Journal of the European Economic Association 11*(3), 661–679.

Samuelson, W. and R. Zeckhauser (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty 1*(1), 7–59.

Schultz, P. W., J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science 18*(5), 429–434.

Shafir, E. (Ed.) (2012). *The Behavioral Foundations of Public Policy.* Princeton, NJ: Princeton University Press.

Stokes, L. C., M. Mildenberger, B. Savan, and B. Kolenda (2012). Analyzing barriers to energy conservation in residences and offices: The rewire program at the university of toronto. *Applied Environmental Education and Communication 11*(2), 88–98.

Sunstein, C. R. and L. A. Reisch (2014). Automatically green: Behavioral economics and environmental protection. *Harvard Environmental Law Review 38*, 128–158.

Sunstein, C. R. and R. H. Thaler (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review 70*(4), 1159–1202.

Thaler, R. H. and C. R. Sunstein (2003). Libertarian paternalism. *American Economic Review: Papers and Proceedings 93*(2), 175–179.

UN (1992). *Agenda 21.* United Nations.

A Appendix

Table A.1: Survey responses

QUESTION	All	Men	Women
1. E-mail			
Read e-mail? (% answered yes)	83.19%	80.95%	85.53%
E-mail will affect my own printing? (% answered yes)	35.35%	35.29%	35.39%
E-mail will affect other employees printing? (% answered yes)	74.74%	76.47%	73.85%
Perceived e-mail:			
Very positive/encouraging/needed	8.48%	7.14%	9.21%
Positive/encouraging/needed	49.15%	59.12%	43.42%
Neither positive or negative	37.29%	26.19%	43.42%
Negative/intrusive/irritating	5.09%	7.14%	3.95%
Very negative/intrusive/irritating	0.00%	0.00%	0.00%
2. Printing preferences (% of all respondents)			
Strongly prefer duplex	35.59%	30.95%	38.16%
Prefer duplex	40.68%	42.86%	39.47%
Does not matter	11.02%	14.29%	9.21%
Prefer simplex	10.17%	9.52%	10.53%
Strongly prefer simplex	2.54%	2.38%	2.63%
3. Willingness to pay: Duplex (% of those with duplex preferences)			
More than 1 kr	4.44%	6.45%	2.39%
Less than 1 kr but more than 50 öre	10.00%	16.13%	6.78%
Less than 50 öre but more than 1 öre	30.00%	22.58%	33.90%
Less than 1 öre but more than 0 kr	10.00%	6.45%	11.86%
Nothing, I would print simplex	45.56%	48.39%	44.07%
4. Willingness to pay: Simplex (% of those with simplex preferences)			
More than 1 kr	0.00%	0.00%	0.00%
Less than 1 kr but more than 50 öre	6.67%	0.00%	10.00%
Less than 50 öre but more than 1 öre	26.67%	20.00%	30.00%
Less than 1 öre but more than 0 kr	6.67%	0.00%	10.00%
Nothing, I would print double-sided	60.00%	80.00%	50.00%
5. Knowledge			
Know how to change from single-sided/double-sided? (% answered yes)	96.61%	95.24%	97.37%
Use to change some of the pre-set settings before printing? (% answered yes)	79.83%	85.00%	77.03%
Have you made changes to your local computer's print settings? (% answered yes)	58.48%	64.29%	55.26%
6. Default			
Default alternative will affect my own printing? (% answered yes)	50.00%	54.76%	47.37%
Default alternative will affect other employees printing? (% answered yes)	91.53%	92.86%	90.79%
Which default alternative would you choose? (% answered duplex)	95.02%	95.24%	94.74%
Do you perceive duplex printing as pro-environmental? (% answered yes)	97.46%	97.62%	97.37%
What is the default on your main printer? (% of respondents)			
Simplex	20.34%	11.90%	25.00%
Duplex	72.88%	83.33%	67.11%
Do not know	6.78%	4.76%	7.89%
Change printer if your main printer changed default setting? (% answered yes)	26.27%	19.05%	30.26%

A.1 Print screens

Figure 9: Simplex default

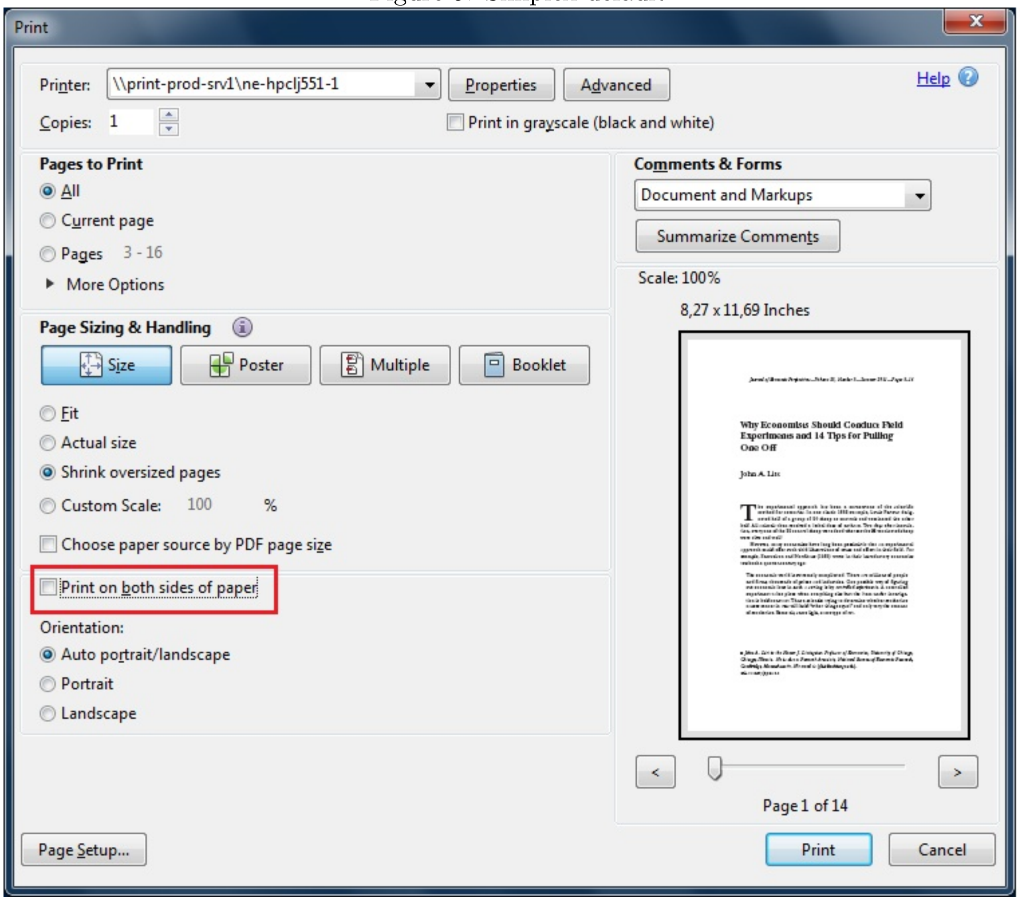
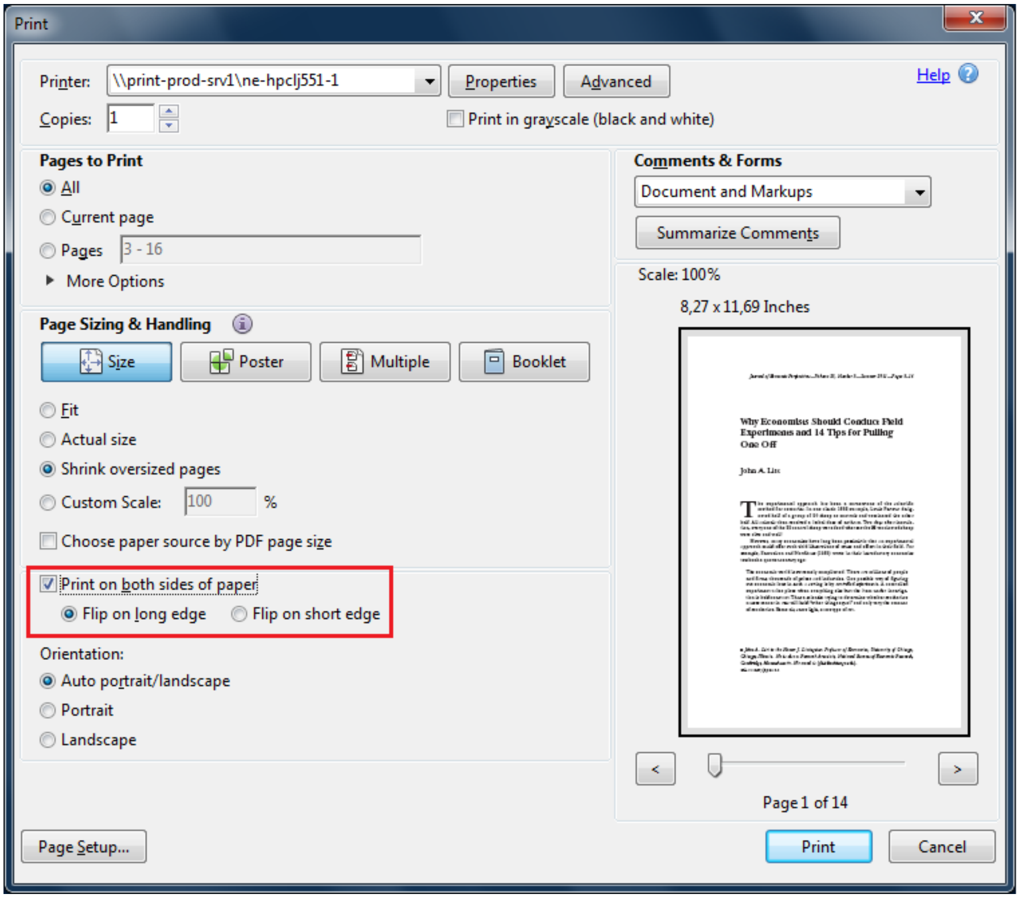


Figure 10: Duplex default



A.2 Default change notice

Dear Colleague,

In order to save the environment and resources, we have decided to change the default setting from simplex to duplex printing on printer X. The change enters into force tomorrow, x/x-2012. The possibility to print simplex remains, and if you absolutely want one-sided print jobs as a pre-selection, you can change your personal default to simplex. This is done under ”Printer Settings” on your particular computer. If you have any problems in the transition or if you want help with your own printer settings, you can contact the IT media helpdesk by phone: XX-XXXXXX.

Best regards

X X, Head of Department

A.3 Moral appeal message

Dear Colleague,

At X we strive to reduce our impact on the environment, and as an employee you can be part of this process. An easy way to save resources is to keep paper consumption low. You can contribute by choosing to print on both sides of a sheet. Duplex printing is available on most of the printers at X, and you can also make this printing mode the default option on your computer. If you need help with setting up your own printer settings, you are welcome to contact the IT media helpdesk by phone: XX-XXXXXX. Thank you for your cooperation!

Best regards

X X, Environmental coordinator, Faculty

X X, Environmental coordinator, The Environmental council

The Origins of Behavioral Contagion:

Evidence from a Field Experiment on Facebook *

Johan Egebark[†] Mathias Ekström[‡]

Abstract

We explore the micro-level foundations of behavioral contagion by running a natural field experiment on the networking site Facebook. Members of Facebook express positive support to content on the website by clicking a Like button. Making use of people's actual accounts, we study whether users are more prone to support content if someone else has done so before. We distinguish between three different treatment conditions: (1) one unknown user has Liked the content, (2) three unknown users have Liked the content, and (3) the most connected person in the network has Liked the content. Whereas the first condition had no effect, the latter two more than doubled the probability that people conform. The existence of threshold effects in our experiment shows that both group size and social proximity matters when opinions are shaped.

Key words: Social Influence; Contagion; Field Experiment; Online Content

JEL classification: A14; C93; D03; D83

---

*We want to thank all the Facebook users who made this experiment possible. We also want to express our gratitude to Pamela Campa, Stefano DellaVigna, James Fowler, Peter Fredriksson, Patricia Funk, Magnus Johannesson, Niklas Kaunitz, Erik Lindqvist, Martin Olsson and Robert Östling, as well as numerous conference and seminar participants, for helpful discussions and valuable comments. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. All errors are our own.

[†]Department of Economics, Stockholm University, and the Research Institute of Industrial Economics (IFN). E-mail: johan.egebark@ne.su.se

[‡]Department of Economics, Norwegian School of Economics, and Research Institute of Industrial Economics (IFN). Email: mathias.ekstrom@nhh.no

1 Introduction

Whenever a new trend arises, be it within fashion, on product markets or even in politics, it is relevant to ask if the popularity is explained by better quality, or if it simply reflects a desire people have to do what everyone else does. The latter supposition, if true, has wide implications since it could explain, among other things, the formation of asset bubbles and dramatic shifts in voting behavior. Unfortunately, identifying herding behavior is by its nature difficult and hence we still know little about the importance of this phenomenon (Manski, 1993, 2000). In this paper we explore the details of contagion dynamics. In particular, we examine how small changes in the size of the influencing group, and the introduction of social ties between the source and the target, affects the spread of behavior.

To study the origins of behavioral contagion, we use the world's leading social networking service, Facebook. Each Facebook user has a network of friends with whom he or she may easily interact through several different channels, e.g., by mailing, chatting or uploading photos or links. The most popular feature allows users to post short messages, so called status updates, for people in the network to read. Anyone of the friends in the network may react to these messages, either by writing their own comments or by pressing a "Like button". Pressing the button is a popular way of showing appreciation, i.e. to express positive support to content on the website.

We set up a natural field experiment to study whether users are more willing to Like an update if someone else has done so before. With access to real Facebook accounts, we post authentic status updates during a seven-month period. For every new update, we randomly assign subjects (i.e., the account holder's friends) into either a treatment or a control group. Both groups are exposed to *identical* status updates; however, while the treated individuals see the update together with previous opinions, this is not the case for individuals in the control group. To uncover the existence of threshold effects, we expose the subjects to three different treatment conditions: (1) one unknown individual has Liked the update, (2) three unknown individuals have Liked the update and (3) one user with a central position in the network has Liked the update. A comparison of conditions (1) and (2) determines the importance of the number of predecessors, whereas a comparison of (1) and

(3) tests whether social ties matter (holding mean group behavior constant). The results from this exercise are striking: while the first treatment condition left subjects totally unaffected, both the second and the third more than doubled the probability of Liking an update. The results are robust to different specifications, and apply to both men and women. Moreover, the effect is both content and sender independent, which suggests it applies more broadly. A closer look at the data shows that the effect is not driven by inattention; rather, it seems to be normative social influence that is the main mechanism behind the observed behavior.

The finding of low threshold effects in our experiment will hopefully contribute to a wider understanding of contagion dynamics. We show that single individuals are indeed influential, but only among those within a network. This implies that contagion can take off from a single node and evolve endogenously from peer to peer. We also demonstrate that once a sufficient number of peers have adopted the behavior, they will affect people outside the network. Hence, size and source may serve as complements in the proliferation of behavior—source increases penetration whereas size generates dispersion. The existence of a social multiplier in this setting has important direct implications. For a firm that seeks to use word-of-mouth as a way to increase visibility it is valuable to understand that once the number of recommendations reaches a certain threshold there will be a "snowball effect". More importantly, we show that the threshold is reached relatively quickly, even in the case when predecessors are complete strangers. The finding of multiplication effects within a social network gives credibility to the practice of offering discounts in return for Likes in order to multiply exposure.

An important contribution of this experiment is that subjects act in their natural environment, and are unaware of the fact that they are part of an experiment (Al-Ubaydli and List, 2012). This leads us to conclude that decades of social influence research from lab settings, including Asch's (1955) influential study and more recent contributions such as Goeree and Yariv (2010), cannot be dismissed as an artefact shaped by suspicious subjects, strange environments, or influential experimenters.

2 Background

2.1 Description of Facebook

Facebook is the leading social networking site and the second most visited website of all. When we ran the experiment in 2010, the average user had 130 friends, spent over one day per month on Facebook and created 90 pieces of content each month (links, blog posts, notes, photo albums etc.). Moreover, 50 percent of what Facebook defines as active users logged on to the website on any given day.[1]

Ultimately, Facebook is an arena for people who seek to interact with their network of friends. Other users are added to your network when they accept your friend requests. Once you have become friends you may visit each other's profiles and can easily interact through different channels, e.g., by mailing, chatting or uploading photos or links. The most popular feature allows users to inform their friends of their whereabouts in status updates. These updates are short messages made visible to the network on the so-called News Feed. Immediately after a status update has been posted, friends may react to it either by writing their own comments, or by pressing a Like button to show their appreciation. Both types of responses are shown together with the update and are thus clearly visible to the author of the update, and to the network of friends. Status updates are typically short and most often revealing what the author is doing at the moment, or where he or she is. Figure 5 in the appendix shows how status updates were displayed on the News Feed in the period of the experiment. As is seen, the first update has no feedback while the second has received one comment and the third one Like.

By using different icons on the News Feed, users could choose to see the most recent content (in reverse chronological order) or the content that Facebook defined as top news. According to the company's webpage, the top news algorithm was in 2010 based on "the number of comments, who posted the story, and what type of post it is (e.g., photo, video, status update, etc.)." We discuss this functionality further when addressing potential mechanisms in section 5.

[1]Figures reported here are from CheckFacebook.com which, although not affiliated with Facebook, claims to use data from its advertising tool. Since 2010, the website has grown substantially. According to the latest figures, the website has over 1 billion users every month, and over 600 million daily active users.

2.2 The Like button

In February 2009, Facebook introduced the Like button. The company's description of this feature, and how it works, is as follows:

> *We've just introduced an easy way to tell friends that you like what they're sharing on Facebook with one easy click. Wherever you can add a comment on your friends' content, you'll also have the option to click "Like" to tell your friends exactly that: "I like this."*

Today, hundreds of millions of people Like everything from their favorite artists, products or presidential candidates, to their local venues such as restaurants and hairdressers. Businesses are increasingly taking advantage of this digital word-of-mouth. For example, they use it as a customer acquisition channel, offering discounts in return for Likes. The idea is that this will increase brand visibility in the customer's social network and, ultimately, generate additional revenue. There is growing evidence that online word-of-mouth affects a variety of purchasing decisions (see, e.g., Tucker and Zhang 2011). The importance of Likes is further manifested in the fact that there is a market for them: numerous companies sell Likes to paying customers who want to signal popularity. There is also an emerging academic interest in this phenomenon: Kosinski et al. (2013) show that Likes can be used to predict highly sensitive personal attributes such as sexual orientation, ethnicity, religious and political views, and intelligence.

3 The Experiment

3.1 Recruiting Experimenters

To be able to execute the experiment we needed to collaborate with several of Facebook users. In principal, these users could have been recruited from a representative population. Our conjecture, however, was that very few would have accepted, which would lead to substantial non-random attrition. More importantly, since it would be difficult to retain control of information leakage to subjects, we would risk the benefits that come with conducting a natural field experiment (Harrison and List, 2004). Consequently, we used the strategy to approach only a small group of candidates, and selected candidates that we trusted would not reveal the experiment to anyone. All of the candidates that we approached accepted, and the fact that they gave us full access to their accounts highlights the degree of mutual trust. Confidentiality was our top priority and we repeatedly instructed the users to never reveal anything about our research. It should be noted that heretofore we have not received any indications that there was ever a breach of this confidentiality. We refer to five of these trusted users as experimenters, and to their more than 700 Facebook friends as the experimental subjects. While the experimenters may not represent the population of Facebook users in general, their Facebook friends should, to a larger degree, do so.

3.2 Design

Normally, posting a status update on Facebook means that all of your friends see it. However, each user has the ability to control who sees a specific update through privacy settings. Thus, if users wish, they can create a subset of friends, e.g., family members or close friends, and make the message visible exclusively to this group. We use this feature in our experiment since it allows us to simultaneously post identical status updates to different groups—in our case treatment and control groups. Importantly, the members of a group can only follow the communication within the specific group, and this communication is displayed as normal to the selected members. Hence, we have no concern that the subjects perceived the updates that we posted differently from the ordinary stream of information on the News Feed.

We posted 44 status updates in total during a seven month period, using the five Swedish experimenters' accounts. Table 1 briefly describes our six step procedure for posting an update. Every time we executed the process we used one experimenter's account; hence, the 44 updates are distributed over all the five accounts. In the first step, we asked one of the five experimenters to text his or her upcoming status update to us, and we explicitly instructed them not to think differently about the updates that we used within the experiment.[2] The updates in the experiment are thus

[2]In fact, they themselves stressed that it was important that the updates that we used expressed something they would have posted anyway, arguably because they did not want to gain a bad reputation by letting us post updates which they could not stand for.

Table 1: The experimental procedure

Step	Description
1	Ask for a status update from one of the five experimenters
2	Random draw of one of the three treatment conditions
3	Random assignment of subjects into treatment and control groups
4	Post identical status updates in treatment and control groups
5	Expose treated subjects to the condition drawn in step 2
6	Collect data on responses

authentic and will appear as a natural part of the ongoing communication on the website. The experimenters continuously posted updates on Facebook before, after, and during the experiment. From the list of examples given in table A.1 in the appendix, we note that the updates are trivial in the sense that they are short, fairly easy to interpret and do not say anything which could be perceived as sensitive, such as political opinions or religious views.[3]

In the second step, after we received the content for a status update, we randomly drew one of three types of treatment conditions: (1) one unknown user has Liked the update, (2) three unknown users have Liked the update, or (3) the most connected person in the network has Liked the update. We use these treatments to explicitly test the importance of group size and social ties as determinants of influence. Randomization in this step means that we can eliminate systematic differences in updates between treatments.

The third step implied random assignment of subjects into either a control or a treatment group. Since randomization occurred at the individual level, a subject's treatment status varies across updates, and within the type of treatment condition. This is convenient since we avoid a group cluster design, and can control for subject fixed effects. In the fourth step, we posted the same status update to all of the subjects (i.e., to everyone in the treatment and the control groups).

[3]We want to study behavior in the simplest possible setting. It would be interesting in further research to see if conformity depends on the type of update but this question is outside the scope of this study.

In the fifth step, we immediately exposed the subjects in the treatment group to the condition that was drawn in step two.[4] Finally, we collected data on responses. Since this six-step procedure is repeated for every new update, it is useful to think of the experimental design as 44 trials distributed over five subsamples.

Each one of the treatment conditions alters the initial condition that the subjects face, and the question that we are interested in is whether this affects the final number of Likes. To keep information close to the initial conditions, we used the strategy to partition groups into smaller entities. The strategy turned out to be successful: control group subjects were unexposed to Likes in more than 70 percent of all possible cases. (We also note that the advent of endogenous Likes will, if anything, introduce a downward bias.) The treatment effect is estimated as the difference in the outcome across treatment and control groups, and we study the importance of group size and social ties by comparing the treatment effect across the three treatments.

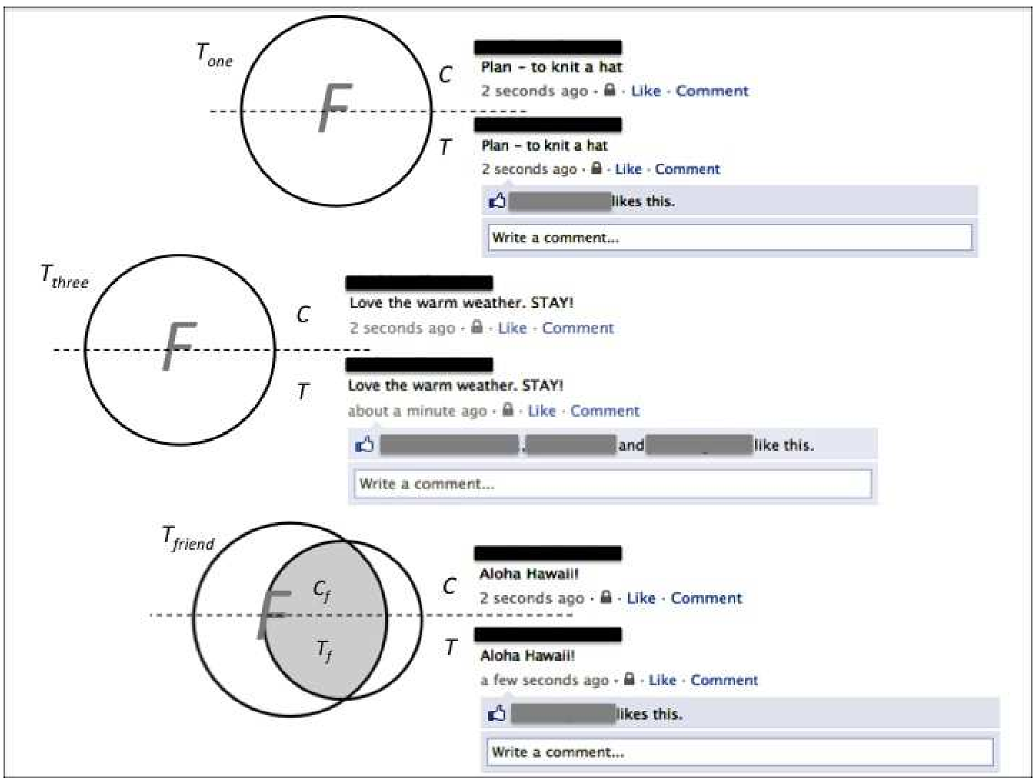
Figure 1 provides a graphical illustration of the three treatment conditions that we use. The baseline treatment condition T_{one}—one unknown user Liking the update—is illustrated at the top of the figure. Randomization divided our experimenter's set of friends, F, into two equally sized groups, the control group, C, and the treatment group, T. Both groups were exposed to identical status updates, and for the treated individuals we added one Like made by a user who is unknown to the subjects.[5] The figure shows how the update is displayed in each of the two groups on Facebook's News Feed. Since this treatment reveals only one person's opinion, and the person is unknown to every one of the subjects, we think of this as the lowest possible trigger.

It is reasonable to assume that social influence is stronger if predecessors are more unanimous (Asch, 1956; Latané, 1981). A natural extension, therefore, is to add more Likes to the updates. The middle section of figure 1 illustrates the next treatment condition, T_{three}, which increases the number of influencers to three. We still want the users who has Liked the update to be unknown to the subjects, since this implies that it is straightforward to compare the results from the first two

[4]It is we, the researchers, who press the Like button using the accounts that we have access to. The updates are thus Liked by one (or three) of the users' accounts available to us. Note that we have access to more than five accounts in total.

[5]The unknown users were added to our five experimenters' networks at the outset of the study, and they were chosen in such a way that we are certain they are unknown to all the subjects.

Figure 1: Illustration of treatment conditions



treatment conditions. If there turns out to be a difference in effects, we can learn something about whether the number of predecessors matters. Again, we randomly assigned friends into either a treatment or a control group and we exposed both groups to identical status updates. The decision to add exactly three Likes in this treatment condition has several reasons. First, increasing them one step at the time would have been too time-consuming. Further, we want to signal to some extent that the update in question is popular without making it stand out too much in the News Feed. Finally, the seminal studies by Solomon Asch on how subjects change private answers to simple questions when exposed to group opinions show that three confederates have the largest marginal influence on subjects' decision to conform (Asch, 1955, 1956).[6]

The last treatment condition, $T_{central}$, measures whether behavioral contagion depends on the strength of a relationship. From previous research we know that conformity is stronger among in-group members, and that "strong ties" exert more impact relative to "weak ties" (Abrams et al., 1990; Bond et al., 2012). It thus seems reasonable to expect that people are more likely to conform to the opinion of someone who is close than to the opinion of a stranger. The specific question that we ask is whether one close person is influential in isolation. Facebook is a good testing ground for this question since it is built around the concept of friendship. In particular, we are able to first objectively define what we denote as a central person, and then study the influence that this person has on subjects. Imagine again our experimenter's set of friends, F, illustrated by the bigger circle in the bottom of Figure 1. Each friend in this set has his or her own set of friends. For a majority of the friends in F the sets are overlapping, but the number of friends in common varies. The central person is the person in F who has the most friends in common with our experimenter (i.e., the largest overlapping area).[7] A subject belonging to the gray area in figure 1 is referred to as a common friend—he is a friend to both the experimenter and to the central person. In the last treatment condition, the central person is the one who has Liked the update. We notice that the group of common friends is always a subset of F (in our case the fraction is around 50 percent),

[6]Whether or not this result translates to our setting is an open question; nevertheless, we use this result as guidance.

[7]Several centrality conditions exist within social network theory. At this stage we found it appropriate to focus on the simplest possible condition, the so called degree centrality. Other measures, such as the Bonacich centrality or the intercentrality condition, take into account the centrality of the people you know and how important the indirect links are (Ballester et al., 2006).

Table 2: Experimenter characteristics

Background variables				Treatment conditions				Responses	
Exp.	Gender	Age	# Friends	N	T_{one}	T_{three}	$T_{central}$	Like	Comment
1	Female	29	120	960	3	5	-	19	8
2	Male	27	204	816	2	1	1	10	12
3	Male	27	152	608	1	1	2	6	4
4	Female	27	176	2464	4	5	5	42	18
5	Male	28	58	812	4	6	4	13	16
Total:			710	5660	14	18	12	90	48

and that random assignment of the friends in F automatically splits this group into a treatment and a control group. Importantly, for the subjects who are not common friends the third treatment condition is identical to the first condition described above.

4 Data

Table 2 describes the five experimenters from whose accounts we post status updates. Collaborating with real Facebook users was crucial since we wanted to use subjects that acted in their natural environment (Harrison and List, 2004). One drawback with this strategy is that we are restricted to a small and selected sample of experimenters. Table 2 shows that at least there is variation along important variables such as gender and number of friends. The unit of analysis is the subject, a unique friend-user combination, which means we have 960 observations for our first experimenter (120 friends times 8 updates), 816 observations for our second experimenter (204 friends times 4 updates) and so on. Consequently, we have in total 710 subjects and 5660 observations (which are evenly distributed across gender and the number of friends of the experimenter). Columns 6–8 show that the treatment condition T_{three} was drawn more frequently than the other two conditions. The difference is quite small, however, especially when considering that we could not use condition $T_{central}$ for one of the experimenters.[8] The last two columns give the distribution of responses. The respective outcomes are dummy variables that equal one if a subject responded to the status update. The two outcome variables are defined analogously: if the first response from the subject

[8]We were unable to approach experimenter 1's central person since this would risk the benefits of conducting a natural field experiment (see discussion in section 3.1).

was to press the Like button (give a comment) we define this as a Like (comment). We note that 90 Likes and 48 comments were made in total during the experiment. Finally, we have checked that the randomization process worked out well by verifying that the treatment and control groups balance in terms of background variables.

5 Results

5.1 Main findings

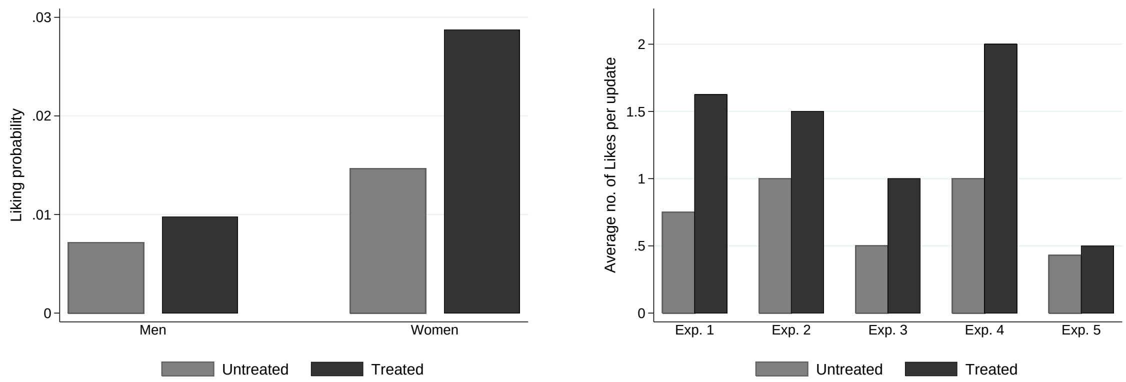
We begin this section by presenting some basic findings that serve as first evidence of a general treatment effect. First, we confirm that there is a striking difference in the total number of Likes across treatment and control groups: whereas control group subjects showed their appreciation (by hitting the Like button) 32 times during the experiment, treatment group subjects did so 58 times.[9] Second, as is shown in figure 2 (a), both men and women are affected by the treatment, even though women appear to respond somewhat stronger. (The fact that women are more prone to Like content is in line with Facebook reports saying women are behind 62 percent of all activity.) Third, as shown in figure 2 (b), subjects respond to treatment irrespectively of which of the five experimenters we consider. This means that the general treatment effect is not driven by one or a few of the experimenters.[10] The main message from these findings is that other people's opinions matter: when a subject is exposed to previous Liking he is about twice as likely to express support to the shared content on the website.

From figure 2 we cannot say whether the number of predecessors matters, or to what extent a central person is more influential. We therefore continue by studying each treatment condition separately in table 3. In all of the columns in the table, we simply regress the dummy variable Like on six different "assignment categories", i.e., three control groups and three treatment groups. The constant in the first row gives the probability to Like for the T_{one} control groups whereas rows 2–3

[9] As is described in section 3 above, subjects are assigned to treatment and control groups repeatedly; hence, one subject can be treated in one trial and untreated in another.

[10] This is also suggested by the fact that regression estimates are robust to the inclusion of user fixed effects in table 3.

Figure 2: Treatment effect by gender (a) and experimenter (b)



(a) (b)

show whether the probability to Like differ for the control groups of T_{three} and $T_{central}$. Rows 4–6, thus, give the effect for each of the three treatment conditions.

The point estimate in row 4 column 1 is small and statistically insignificant. This clearly shows that when one unknown person has Liked a status update, people do not care. For the other two treatment conditions, however, results are strikingly different. Row 5 shows that when three unknown persons have Liked a status update, people are about twice as likely to express their appreciation (with the effect being statistically significant). From this we conclude that increasing group size only slightly is enough to induce contagion in this setting. Row 6 of column 1 looks at what influence a central person has. The treatment effect is again substantial: when the central person has acted before, subjects are roughly twice as likely to express support. Finally, we test the robustness of these results by adding various fixed effects. Columns 2–3 include experimenter and subject fixed effects, respectively. In order to include update fixed effects in the last column we use the fact that the control group means are statistically indistinguishable—as is seen in rows 1–3 of columns 1–3—and pool all of the control groups into one. The results are robust to all of these different specifications.

When studying the central person's influence, in row 6 of table 3, we use the entire sample,

Table 3: Regression results using the full sample

	Dependent variable: Like			
	(1)	(2)	(3)	(4)
Constant	0.015***	0.015***	0.015***	0.011***
	(0.003)	(0.003)	(0.004)	(0.002)
T_{three} *control*	−0.004	−0.004	−0.004	–
	(0.004)	(0.004)	(0.005)	
$T_{central}$ *control*	−0.008	−0.007	−0.007	–
	(0.005)	(0.005)	(0.006)	
T_{one} *treatment*	−0.003	−0.003	−0.004	−0.003
	(0.006)	(0.006)	(0.005)	(0.006)
T_{three} *treatment*	0.016***	0.016***	0.014***	0.016***
	(0.005)	(0.005)	(0.005)	(0.005)
$T_{central}$ *treatment*	0.015*	0.015*	0.015*	0.015*
	(0.008)	(0.008)	(0.008)	(0.008)
Experimenter FE	NO	YES	NO	NO
Subject FE	NO	NO	YES	NO
Update FE	NO	NO	NO	YES
N	5,660	5,660	5,660	5,660
R^2	0.003	0.003	0.181	0.013

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. The outcome is a dummy variable that equals one if a subject responded by Liking the status update. The constant in columns 1–3 gives the probability to Like for the T_{one} control groups. The constant in column 4 gives the probability to Like for all the control groups combined. Standard errors clustered on the update level in parenthesis.

lumping together subjects belonging to the group of common friends and those who do not. (As described in section 3.2, a common friend has a relation to both the experimenter and to the central person.) Note, however, that we do not expect any reactions from subjects who have no relation to the central person; otherwise, our previous finding of no effect for the baseline condition would be questioned. Table 4, where we separate out the group of common friends, clearly shows that it is only those who have a social relation to the central person who respond to the $T_{central}$ treatment. Common friend subjects are more than three times as likely to show appreciation when they observe their friend's action, and the effect is statistically significant at the 5%-level.[11] Columns 1–2 further show that both subgroups are affected similarly by treatment conditions T_{one} and T_{three}.

The last two columns of Table 4 separate between women and men. The treatment response is more or less similar across gender for conditions T_{one} and T_{three}. However, judging from the point estimates, women appear to respond stronger to the $T_{central}$ condition (even though the insignificant estimate for men makes it difficult to draw decisive conclusions).

5.2 Influence or Attention?

In this section we consider the possibility that the effect that we observe is driven by limited attention, rather than social influence. In particular, we want to rule out that treatment increased the probability of actually observing and reading an update, since this would raise the number of responses in a mechanical fashion.[12]

We first note that status updates with at least one Like is more salient than those without because of their altered appearance (figure 1 shows that a blue rectangular area is added beneath the status update). The three treatment conditions affect the look of an update identically; thus, if saliency is driving the observed herding, we would expect to see effects across all three treatment conditions. From the above results we know that there is one condition without any effect, and we thus rule out this channel as an explanation.

[11] The point estimates in column 1 of Table 4 suggest that the effect is smaller for T_{three} than for $T_{central}$. However, a Wald test cannot reject equality of the two estimates (*p-value* $= 0.483$). In sharp contrast, we can reject that T_{one} has the same effect as T_{three} (*p-value* $= 0.025$) or $T_{central}$ (*p-value* $= 0.028$).

[12] For studies on limited attention see, e.g., Barber and Odean (2008); Ariely and Simonsohn (2008); DellaVigna and Pollet (2009).

Table 4: Regression results for different subgroups

	Dependent variable: Like			
	CF	Not CF	Women	Men
Constant	0.008***	0.014***	0.017***	0.006***
	(0.002)	(0.002)	(0.003)	(0.002)
T_{one} *treatment*	−0.006	−0.002	−0.004	−0.002
	(0.007)	(0.008)	(0.010)	(0.004)
T_{three} *treatment*	0.015***	0.015**	0.021**	0.008**
	(0.006)	(0.006)	(0.005)	(0.004)
$T_{central}$ *treatment*	0.025**	0.006	0.041**	0.011
	(0.012)	(0.008)	(0.019)	(0.013)
Status update FE	NO	NO	NO	YES
N	2,368	3,292	2,630	2,212
R^2	0.023	0.021	0.029	0.028

Notes: ***$p < 1\%$, **$p < 5\%$, *$p < 10\%$. The outcome is a dummy variable that equals one if a subject responded by Liking the status update. The first column (CF) gives the results for the common friend subjects, whereas the second column (Not CF) gives the results for those not in the group of common friends. The constant in columns 1–3 gives the probability to Like for the T_{one} control groups. The constant in column 4 gives the probability to Like for all the control groups combined. Standard errors clustered on the update level in parenthesis.

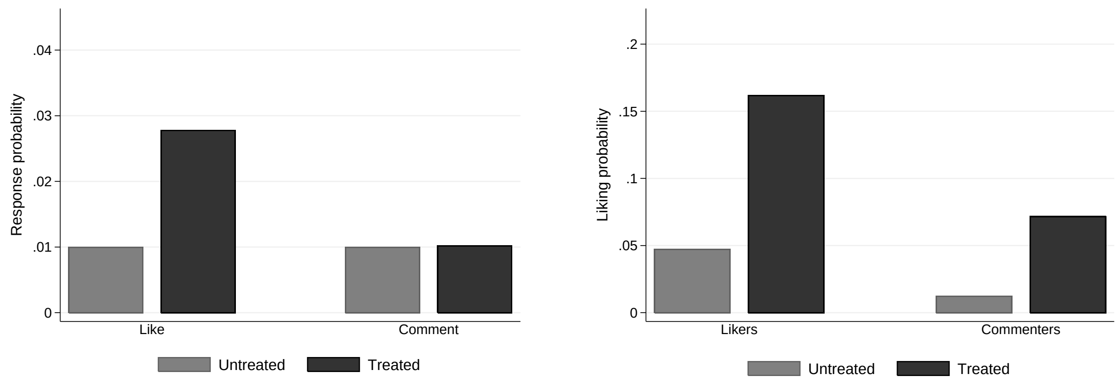
We next consider the possibility that subjects screen updates in another, more deliberate, sense. Users of Facebook may actively look for previous Likes (or any response for that matter) in order to quickly find the best updates, since this will save time and effort. If such a search rule exists, it is more likely that subjects in the treatment groups have noticed and read the status updates that we posted. In the end this could give rise to a difference in the number of responses seen across treatment and control groups. To address this possibility, we execute a simple placebo test where we focus on the type of response given to the updates. Deliberate searching should only affect which updates are read, not the mode of response. Hence, if this explanation is valid, we expect a treatment effect for comments as well. Figure 3 (a) shows that comments are completely unaffected ($t(2986) = 0.059, p = 0.953$), which is a finding that runs counter to the searching explanation. Comparing the gray bars, we also note that both response modes are equally popular in the absence of treatment. Consequently, it is not the case that Liking was the only suitable response to the updates that we used in the experiment.[13] The finding in figure 3 (a) also speaks against a more general top news-effect (see section 2.1). This is in line with tests that we conducted before launching the experiment: we posted several updates to different Facebook users, and varied the number of Likes (if any) that we attached to them. This procedure allowed us to establish that our treatment conditions did not affect how updates were displayed on the News Feed.

The above finding cast serious doubt on previous Liking as a general screening device. Nevertheless, it could be that users who prefer Liking (to commenting) screen updates based on Likes, while users who prefer to comment screen updates based on previous comments. This reasoning implies that the observed treatment effect is enhanced (or does only exist) among subjects who typically Like updates. We collected information on the subjects' response behavior before the experiment started, and therefore have a baseline measure of each subject's preferred response mode. Figure 3 (b) shows the probability to Like for "Likers" and "Commenters", respectively.[14] As is seen, response mode preferences are related to levels in the expected way, i.e., subjects who preferred Liking prior

[13]In this and in the remaining analysis we focus on the treatment conditions that proved to be meaningful, and treat them as one condition for visual ease. We can confirm that the T_{one} condition remained irrelevant, and that treating T_{three} and $T_{central}$ as separate conditions does not change the conclusions.

[14]A subject is defined as a Liker if she responded with a Like strictly more times than with a comment on ten status updates just before the experiment started (vice versa for Commenters). There are as many Likers as Commenters.

Figure 3: Treatment effects by outcome (a) and response preference (b)



(a) (b)

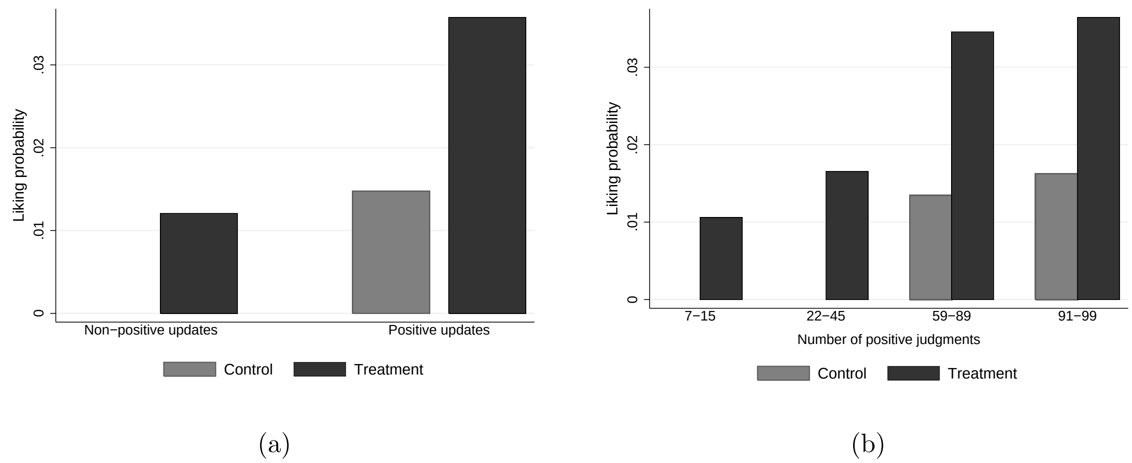
to the experiment continue to do so. However, the treatment effect is independent of being a Liker or a Commenter (a conclusion confirmed by regression analysis).

In summary, all of the mechanisms that we can think of related to limited attention—whether it be saliency or different types of searching—seem inconsistent with our findings. A plausible reason for the lack of support for this explanation is that status updates on Facebook are typically so short and easy to understand that the time and effort one can save by searching for prior responses is negligible. Presumably, a more popular screening method is to select updates based on who posted the update, especially since people will quickly learn who usually posts the best updates. Such a search mechanism will only affect the number of potential responders in our experiment, not the results.

5.3 Normative or Informational Social Influence?

We end the analysis by presenting some results that are suggestive of the type of social influence. Deutsch and Gerard (1955) refer to informational social influence as the "influence to accept information obtained from another as evidence about reality," and normative social influence as the "influence to conform to the expectations of another person or group." Cai et al. (2009) provide

Figure 4: Effect by update type



a good example of a setting where informational influence is at play. When diners in a Chinese restaurant chain are informed about the past weeks most popular dishes, the demand for these alternatives increases. Since deciding on what dish to order involves uncertainty, information about prior choices can serve as a quality signal that helps individuals make optimal choices. Contrary to the restaurant setting, choices in our experiment are made after subjects have experienced the "product" and have been able to evaluate it against comparable alternatives (figure 5 in the appendix shows that users can easily read and compare status updates on the News Feed). This indicates that informational social influence is less important in our setting. However, even though most updates must be seen as relatively easy to evaluate, it is not unthinkable that some subjects are unsure of how to interpret a specific update. For example, should we interpret the update 'I'm probably the only tourist who has visited Pisa but didn't see the tower...' as a good or a bad thing? If the user has spent their trip in the hospital after an accident, Liking the status update is inappropriate; if, instead, the user refers to a vacation in Pisa where too much fun was going on to bother with the leaning tower, a Like is a more proper response.

We argue that if informational social influence is driving behavior, the treatment effect should be pronounced for updates where the content is more ambiguous. Consequently, we let 40 persons judge whether they perceived each of the 44 updates as unambiguously positive, unambiguously negative, or difficult to interpret (i.e., ambiguous). This strategy allowed us to get a first indication of whether an update is perceived as ambiguous or not. We then recruited 61 persons and instructed them to guess, for each update, what the most popular answer was in the first group. These 61 persons were incentivized to make as good guesses as possible: they received 2 SEK (0.3 USD) per correct guess, and if someone had more than 34 correct answers we paid them 100 SEK (15 USD). The majority of the updates were labeled as unambiguously positive, and there was strong consensus across the two groups. We therefore define an update as positive if this was the most popular answer among the 101 persons asked, and non-positive if the majority labeled it as either unambiguously negative or as difficult to interpret. In figure 4 (a) we note that positive updates have a higher overall probability of being Liked. More importantly, the treatment effect is independent of whether the update is positive or not. In figure 4 (b), we sort the updates into quartiles depending on the number of positive judgments, allowing us to focus exclusively on the updates that a large majority defined as positive. Strikingly, and in contrast to the information hypothesis, even for the updates that more than 90 percent considered unambiguously positive there is a large treatment effect ($t(1220) = 2.209, p = 0.027$). In summary, we do not find any support for an explanation based on informational social influence. A tentative conclusion is therefore that normative social influence may offer a more compelling description of our findings.

6 Conclusion

This paper reports the results from a natural field experiment on the social networking service Facebook. We show that the decision to support content on the website is influenced by the existence of previous support. A key feature of the setting and the experimental design is the possibility to evaluate influence along two dimensions: group size and social proximity. In accordance with social impact theory developed by Latané (1981), influence is stronger when the group of predecessors is larger, or when influence comes from a person with a central position in the network. We find clear-cut evidence of threshold effects that hopefully will contribute to a wider understanding of contagion dynamics. Our results demonstrate that single individuals are influential within networks, whereas group size determines whether people outside networks are affected. The existence of a social multiplier in this setting has important direct implications. For a firm that seeks to use word-of-mouth as a way to increase visibility it is valuable to understand that once the number of recommendations reaches a certain threshold there will be a “snowball effect”. More importantly, we show that the threshold is reached relatively quickly, even in the case when predecessors are complete strangers. The finding of (even stronger) multiplication effects within a social network gives credibility to the practice of offering discounts in return for Likes in order to multiply exposure.

References

Abrams, D., M. Wetherell, S. Cochrane, M. A. Hogg, and J. C. Turner (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. *British Journal of Social Psychology 29*(2), 97–119.

Al-Ubaydli, O. and J. A. List (2012). On the generalizability of experimental results in economics. NBER Working Paper Series.

Ariely, D. and U. Simonsohn (2008). When rational sellers face nonrational buyers: Evidence from herding on eBay. *Management Science 54*(9), 1624–1637.

Asch, S. (1955). Opinions and social pressure. *Scientific American 193*(5), 31–35.

Asch, S. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs 70*(9).

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: The key player. *Econometrica 74*(5), 1403–1417.

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies 21*(2), 785–818.

Bond, R. M., C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler (2012). A 61-million-person experiment in social influence and political mobilization. *Nature 489*, 295–298.

Cai, H., Y. Chen, and H. Fang (2009). Observational learning: Evidence from a randomized natural field experiment. *American Economic Review 99*(3), 864–882.

DellaVigna, S. and J. M. Pollet (2009). Investor inattention and Friday earnings announcements. *Journal of Finance 64*, 709–749.

Deutsch, M. and H. B. Gerard (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology 51*(3), 629–636.

Goeree, J. K. and L. Yariv (2010). Conformity in the lab. Revise and resubmit Economic Journal.

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic Literature 42*(4), 1009–1055.

Kosinski, M., D. Stillwell, and T. Graepel (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences 110*(15), 5733–5734.

Latané, B. (1981). The psychology of social impact. *American Psychologist 36*, 343–356.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives 14*(3), 115–136.

Tucker, C. and J. Zhang (2011). How does popularity information affect choices? a field experiment. *Management Science 57*(5), 828–842.

A Appendix

Figure 5: Print screen, Facebook homepage

Table A.1: Examples of status updates from the experiment (translated from Swedish)

Treatment	Content
T_{one}	I'm probably the only tourist who has visited Pisa but didn't see the tower...
T_{one}	Party tonight. Prepare myself with intravenous drip and pain killers to be alive tomorrow...
T_{one}	I don't give a damn about your tax refund!
T_{one}	Plan - to knit a hat
T_{three}	Love the warm weather. STAY!
T_{three}	A warm welcome to you, dishwasher!
T_{three}	I'll be surprised if I don't get an A on today's exam
T_{three}	Towards the beach!
$T_{central}$	Rhubarb desert before the race. Hope the jogging tights still fits...
$T_{central}$	Aloha Hawaii!
$T_{central}$	Have the same posture as the Hunchback of Notre Dame. Lumbago please go away!
$T_{central}$	Premiere on the PGA tour ;)

ISBN: 978-91-7649-093-8
ISSN: 1404-3491

Stockholm University

Department of Economics