Leaving Poverty Behind? - The Effects of Generous Income Support Paired with Activation

By

Markussen, Simen and Knut Røed

The definitive publisher-authenticated and formatted version:

Markussen, Simen and Knut Røed, 2016, Leaving Poverty Behind? - The Effects of Generous Income Support Paired with Activation, American Economic Journal: Economic Policy, Vol 8 (1), 180-211.

is available at:

https://doi.org/

Forthcoming in American Economic Journal: Economic Policy

May 19, 2015

Leaving Poverty Behind?

– The Effects of Generous Income Support Paired with Activation

Simen Markussen and Knut Røed*

The Ragnar Frisch Centre for Economic Research, Oslo, Norway

Abstract

We evaluate a comprehensive activation program in Norway targeted at hard-to-employ social assistance claimants with reduced work capacity. The program offers a combination of tailored rehabilitation, training and job practice, and a generous, stable, and non-means-tested benefit. Its primary aims are to mitigate poverty and subsequently promote self-supporting employment. Our evaluation strategy exploits a geographically staggered program introduction, and the causal effects are identified on the basis of changes in employment prospects that coincide with local program implementation in a way that correlates with the predicted probability of becoming a participant. We find that the program raised employment prospects considerably.

Keywords: poverty, vocational rehabilitation, social insurance, treatment effects, program evaluation

JEL classification: C21, C26, H55, I30, J24

---

*This research is part of the project "Evaluation of a Norwegian Qualification Program," which was financed by the Norwegian Ministry of Labor and Social Affairs. It has also received support from the Norwegian Research council (grant # 236992). Thanks to Simen Gaure for the invaluable help with the estimations. Thanks as well to Kristian Heggebø, Ivar Lødemel, Angelika Schafft, three anonymous referees, the Editor Daniel Silverman, and seminar participants in Bergen, Oslo, Stavanger, Sydney and Uppsala for their comments and suggestions. The data made available by Statistics Norway have been essential for the research project. Correspondence to: Knut Røed, the Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, 0349 Oslo, Norway. E-mail: knut.roed@frisch.uio.no.

How should policymakers design social insurance institutions in order to fight persistent unemployment, marginalization and poverty? While economists often emphasize the moral hazard problems and the potential lock-in effects arising from generous social insurance programs, there is also an extensive literature focusing on the barriers associated with poverty itself, which are caused by, e.g. the type of myopic behaviors it promotes and the kind of unhelpful social networks it gives access to; see Dasgupta and Ray (1986; 1987), Dasgupta (1997), Calvó-Armengol and Jackson (2004) and Shah et al. (2012). The empirical evidence indicates that a severe scarcity reduces the ability to focus and concentrate on issues beyond the immediate needs, and that it causes sleep deprivation, erodes self-control and reduces work productivity; see Mullainathan and Shafir (2013) for a recent discussion of the literature. Some income support may therefore be required to break out of a poverty trap. But since income support is normally tested against earned income, generosity may discourage self-sufficiency and create a benefit trap instead. A potential solution to this dilemma is to couple generosity with activation, thereby effectively removing the leisure component from a life on income support, while also ensuring some “maintenance” of basic employment skills. Properly designed, activation requirements facilitate more ambitious social programs without aggravating moral hazard problems; see, for example, Moffitt (2007) for a review of empirical evidence in relation to the introduction of activation requirements in the cash-based welfare program for single mothers in the US, and Røed (2012) for a recent survey of the literature regarding activation strategies in unemployment and disability insurance programs.

In the present paper, we evaluate a “Qualification Program” (QP) that was launched by the Norwegian government in 2007 as its major tool to fight poverty. The program is both costly and ambitious, and designed to combine economic security and activation. It is targeted at persons with a severely reduced earnings capacity and no or very limited social insurance entitlements. The typical recruitment base is persons who have become, or are in danger of becoming, completely reliant on means-tested social assistance (welfare). QP participants may have a variety of problems in relation to a competitive labor market, such as poor language skills, disrupted schooling, little or no work experience, criminal records, and sometimes mental disorders and drugs problems. The aim of the QP is to help these often hard-to-employ persons into self-supporting employment through an individually tailored activation program, under which they also receive a standardized (and *not* means-tested) income support amounting to approximately one-third of the average full-time earnings level in Norway. The contents of the program vary a lot, but are normally made up of a combination of consultations, employment training, medical rehabilitation or therapy, social training and the upgrading of skills.

Based on the existing literature, it is not clear how we would expect the QP to affect the participants' future employment and earnings prospects. On the one hand, the program considerably raises overall benefit levels, and does so for a relatively long period of time. There is ample empirical evidence showing that more generous social insurance has negative effects on labor supply, ceteris paribus; see Krueger and Meyer (2002) for an overview of the literature, and Røed and Zhang (2003; 2005) and Fevang *et al.* (2013) for recent Norwegian evidence. On the other hand, the QP requires full-time participation in a tailored activation program. There exists no general consensus regarding the overall impacts of activation. For unemployed job seekers, the typical findings are that there are favorable "threat effects" prior to active labor market program participation, adverse "lock-in-effects" during participation and then (sometimes) favorable "post-program-effects" afterwards; see, *e.g.* Kluve *et al.* (2007) and Card *et al.* (2010) for recent reviews, and Raaum and Røed (2006) and Røed and Westlie (2012) for Norwegian evidence. There is also a literature focusing on programs specifically targeted at temporary disabled persons, though with a few notable exceptions – Frölich *et al*. (2004) and Aakvik *et al*. (2005), which have indicated effects close to zero for Sweden and Norway, respectively – this literature is more oriented toward comparing the effects of alternative rehabilitation strategies than toward evaluating the more general effects of applying an activation strategy as an alternative to pure income insurance.

To the best of our knowledge, the program we evaluate in the present paper is unique in its combination of offering economic security and (tailored) activation for a hard-to-employ target group with little (or no) social insurance entitlements. It represents a coherent – yet untested – strategy to fight persistent poverty. Fortunately, the program is also unique in that it was implemented in a way that facilitates scientific evaluation; *i.e*. it was gradually phased in over a three-year period, implying that potential participants got access to it at different points in time. In this paper, we combine the staggered program implementation with observed proxies for "participation propensity" to identify the causal impacts of QP participation on subsequent labor market outcomes.

Our main finding is that the program has been successful in terms of helping hard-to-employ persons into employment. Four years after program entry, we estimate that QP participation on average raises the employment rate by approximately 18 percentage points, ceteris paribus.

However, most of the extra employment comes in the form of poorly paid and/or very small jobs; hence, the dependency on transfers from the welfare state remains high. Although the program is designed to offer a considerable increase in income support to persons who would otherwise depend on social assistance, our estimates suggest that the increase in income support actually caused by the program is relatively modest; so from a purely pecuniary point of view, the main impact of the program is to enhance income security and predictability. We argue that the activation part of the program is the essential success factor, potentially together with its facilitation of a relatively secure economic environment.

I. Institutions and Data

The Qualification Program (QP) was launched in November 2007, and was then gradually rolled out over the next three years (2008-2010) – municipality by municipality – in tandem with an administrative reform that merged the local employment and social insurance offices into new joint administrations (NAV); see, e.g. Christensen et al. (2014). By January 2010, the QP had become a nationwide program. Although the empirical strategy we are going to use in this paper does *not* rely on a random-assignment-like order of local implementation, it is worth noting that the implementation sequence was primarily determined on the basis of administrative considerations in relation to the establishment of the new NAV offices, and not on the basis of, say, local employment opportunities.

The QP is designed to support persons who fall between the two stools of employment and social insurance, and thus potentially face serious poverty problems. Even though Norway is typically considered as a country with a relatively comprehensive welfare state, eligibility to social insurance is generally based on past contributions, with thresholds implying that workers with little and/or unstable labor market experience fail to qualify. There are exceptions from the requirement of past contributions in cases of serious disabilities/impairments that have been certified by a physician. Apart from the permanent disability pension, all social insurances are also time-limited (two years for unemployment insurance, four years for temporary disability insurance). Therefore, a considerable fraction of the population becomes reliant at some point on means-tested social assistance (welfare) administered by the municipalities. There are no national regulations regarding the amounts provided through social assistance, but (non-binding) guidelines recommend a support level for singles that corresponds to roughly 15% of the average earnings level in Norway.

The aim of the QP is to provide a stable and safe economic basis over a one-two-year period, while at the same time helping the participants onto a steady path toward self-sufficiency. The program offers an annual salary of NOK 170,000 (approximately $29,000) per year in 2013 (NOK 113,000 for participants below 25 years) – more than twice as much as what can be expected from social assistance – plus child allowances and housing benefits.[1] It is notable that the QP benefit is paid out by the local municipality's pay office rather than by the local welfare office, it is taxed like regular labor earnings and also entails the same holiday and leave privileges. The motivation for all this is to minimize the shame and stigma associated with participating in the program, and thus to reduce the risk of social isolation and withdrawal from networks that may be of importance in the process of breaking out of poverty; see, e.g. Walker et al. (2013). The QP benefit is normally granted for a period of up to two years, but additional extensions can be given on the basis of an individual assessment. In return, the participant is expected to fully take part in a tailored qualification and activation plan agreed upon by the claimant and the caseworker. A failure to do so – without any justifiable cause – normally results in a corresponding pay cut.

Because the program has been considered as a very important and visible element of the government's anti-poverty strategy, it has generally been given a high priority in the allocation of economic resources. Statistics reported by the Norwegian Labor and Welfare Administration (NAV) indicate that on average each QP caseworker has 18 clients. By comparison, caseworkers dealing with the temporary disability insurance program have 86 clients on average.

The legislation states that the Qualification Program applies for persons of working age with a substantially reduced work and income capability, and no or very limited social insurance entitlements. It further states that entitlement also requires that: a) the applicant's work ability has been assessed, b) that the program is viewed as both appropriate and necessary in order to increase the applicants' possibilities for labor market participation, and c) that the Labor and Welfare Administration is able to offer a suitable program. Given that these (somewhat vague) requirements are met, access to the QP is a legally protected entitlement. Formally, the QP benefit is granted for one year at a time, but during the period covered by our analysis, a two-year perspective was stated as the primary rule. Based on an individual assessment, the program can also be extended beyond two years, provided that the claimant has shown pro-

[1] All monetary amounts reported in this paper are inflated to 2013 value (based on the social insurance system inflator, which approximately corresponds to the consumer price index). To compute $ amounts, we have used the average exchange rate in 2013 of $1= NOK 5.88.

gress in his/her efforts to qualify for the labor market, and that a transition to employment appears probable within the near future. The program may be terminated at any time if the participant does not fulfill the obligations set out in the individual plan, or if he/she succeeds in finding a regular job.

For the QP target group, the *alternative* help offered by the welfare state will often be social assistance from the municipality, which is means-tested against household income and wealth. In cases of serious and lasting health problems, a temporary or permanent disability insurance benefit may become an alternative after some time, provided that appropriate rehabilitation attempts have been made first. For most of the participants, the program offers a considerable rise in personal income at the time of entry. Hence, from a pecuniary point of view, the program is typically viewed as attractive. The QP has therefore *not* been considered useful as a willingness-to-work test, whereby caseworkers threaten to terminate social assistance if participation is rejected. To the contrary, it has been emphasized that participation in the QP is voluntary, and should be considered a privilege rather than a duty. In that sense, the QP is more a “carrot” than a “stick.”

Whereas activation requirements are only used sporadically in relation to social assistance claims, participation in the QP entails a *full-time* activity based on an individually tailored plan. If the participants nonetheless have additional earnings during the participation period, the QP benefit is reduced in proportion to the number of work hours outside the program, such that, e.g. a half-time job results in a 50% reduction in the QP benefit.

Based on reports collected from the municipalities, the Norwegian Labor and Welfare Administration (NAV) has counted that 17,214 persons had participated in the QP by the end of 2010, out of which 4,968 had then completed the program according to the individual plan, and 1,414 had dropped out (Schafft and Spjelkavik, 2011). Among those who had completed the program, 31% were reported to have obtained regular employment afterwards, whereas 7% entered regular education. Most of the rest continued receiving some kind of income support, either in the form of temporary or permanent disability insurance or in the form of social assistance.

The data we use in the present analysis are collected from administrative registers and comprise the whole Norwegian population. Information on individual participation in the QP is based on a separate code for QP benefits used on the paycheck submitted for each participant by the municipalities to the tax authorities. This way, we can identify the year of program entry and the years of continued participation, but not the exact start and stop dates. By means of encrypted identification numbers, we merge these data to several other administrative registers containing a wealth of information on individual characteristics (such as gender, age, education, nationality, municipality of residence), as well as longitudinal information on past and future employment and income sources. From these data, we compute *person-year* observations on several *outcomes*, particularly related to employment, earnings and social insurance dependency.

Table 1: Descriptive statistics for participants in the QP

Number of participants	19 211
Entry year	
2008	2,919 (15.2%)
2009	5,857 (30.5%)
2010	6,060 (31.5%)
2011	4,375 (22.8%)
Still participating	
First year after entry year	82.9%
Second year after entry year	49.8%
Third year after entry year	23.1%
Mean age	33.7 years
Female	44.0%
Non-native	50.7%
High school as highest education	16.1%
College/University as highest education	7.0%

Note: Statistics are based on all QP entrants from 2008 through 2011. Participation rates one-three years after entry are based on entrants than can be followed for the period in question, with data ending in 2011. For example, participation three years after entry is based on 2008 entrants only.

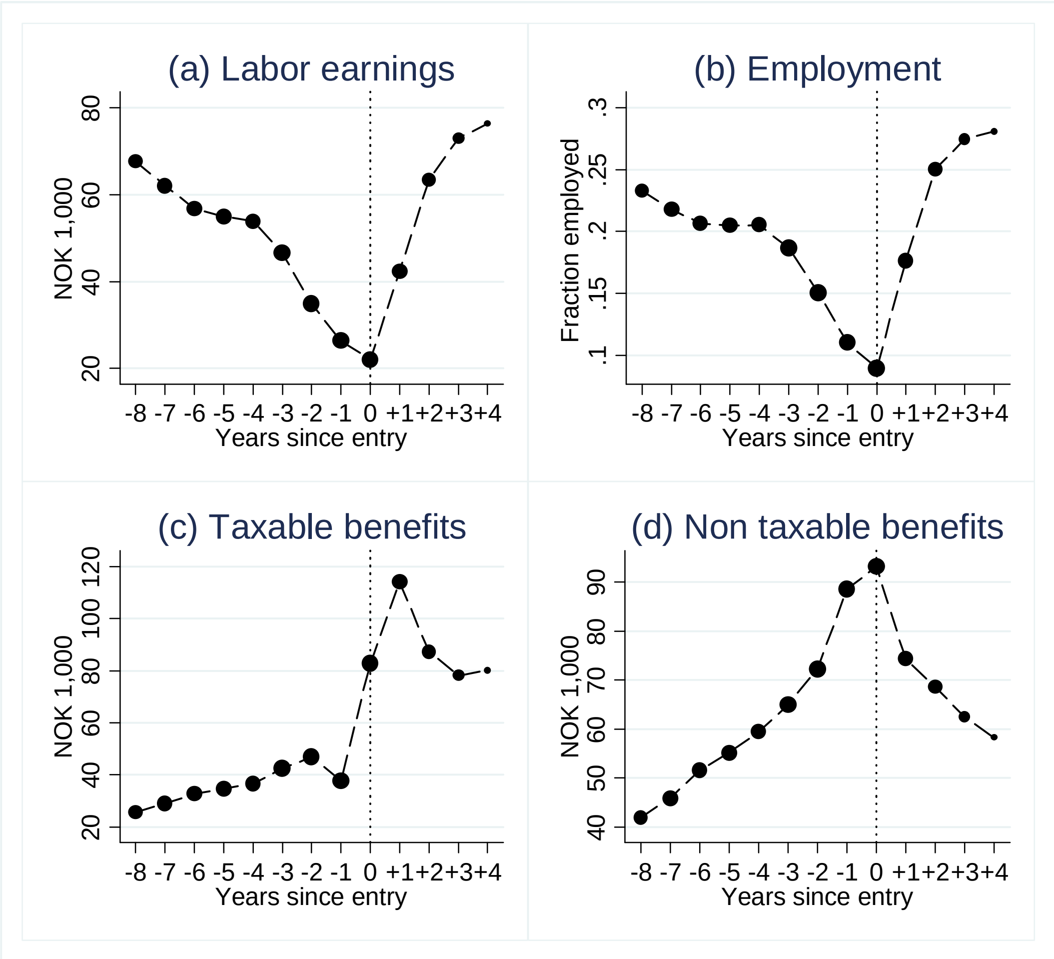

Figure 1: Annual employment, earnings and benefit claims among QP participants
Note: The graphs are based on all QP entrants 2008-2011 (19,211 persons); see also note to Table 1. The sizes of the dots are proportional to the number of observations behind each data point. Back-in-time-observations are censored at age 18, while forward-in-time-observations are censored in 2013. The employment outcome in panel (b) is derived directly from the earnings measure in panel (a), and a person is considered to be employed if annual labor earnings exceed NOK 85,000 (approximately 17% of average full-year-full-time earnings in Norway). Taxable benefits include all social insurance transfers, whereas non-taxable benefits include social assistance (welfare) and some child/housing allowances.

Based on the procedure described above, we identify 19,211 participants from 2008 through 2011, which is roughly in line with the numbers reported (manually) by the municipalities to the Labor and Welfare Administration. The number of entrants was largest in 2009 and 2010; see Table 1. This table also shows that many participants remain in the program for more than two years.[2] Figure 1 presents some key descriptive statistics for the participants identified in our data – in terms of their labor earnings, their employment propensity and their claims of (taxable) social insurance benefits and (non-taxable) means-tested social assistance (welfare)

[2] Since we do not have exact start and stop dates, we cannot compute accurate durations, but the numbers in Table 1 indicate that approximately 50% of the participants are in the program in at least three consecutive calendar years, and 23% are participating in at least four years.

for a period from eight years before until three years after entry into the QP. These statistics all indicate that entry into the QP coincided with a marked turning point in economic outcomes for the participants. Prior to the year of QP entry, the participant group members experienced a steady decline in average employment and labor earnings during the whole eight-year period covered by our data, and a corresponding rise in the dependency on means-tested social assistance. In the years after entry, these trends were significantly reversed. Moreover, from the year of entry, the level of taxable benefits also rose markedly, basically reflecting that the QP benefit itself falls into this category.

Figure 1 certainly conveys the impression that the QP program must have had large positive impacts on the participants' average employment and labor earnings trajectories. Yet, although the pre-program decline appears to have been a consistent feature of the participants' economic fortunes for many years, we cannot rule out that it mirrors the notorious "Ashenfelter Dip" (Ashenfelter, 1978), and, hence, that the apparent rebound reflects a regression toward the mean.

II. Methodology

The research question we seek to answer is how participation in the QP affects earnings, employment and benefit trajectories for up to four years after the year of program entry, i.e. we will attempt to find out how much – if anything – of the apparent rebound displayed in Figure 1 can be interpreted as causally related to the program. Given the relatively long durations of QP participation, we expect the causal impacts to change significantly with time since program entry. A probable dynamic effect pattern is that there are negative (lock-in) effects on employment and earnings during the first one-two years after entry, whereas the potentially favorable post-treatment effects build up gradually afterwards. Our main success criteria will be based on observed labor market performance in the fourth year after entry, at which point we can be relatively confident that the program participation period has ended.

Even conditional on observed covariates, we expect QP participation to be highly selective, so we cannot identify the causal effect of the QP by comparing participants and non-participants. Instead, our identification strategy relies on interacting individual (predicted) QP participation propensities with dummy variables representing the exact timing of local program implementation. In this way, we use the staggered rollout of the program as the source of random-assignment-like variation in actual QP participation. Somewhat simplified, our estimation strategy can be summarized as follows:

1. For each outcome observation, we compute the *hypothetical* probability that the person in question would have entered the QP in each of the last five years (including the current) provided that the QP had been implemented in the municipality and year in question. These probabilities are estimated on the basis of the subset of observations for which the program was actually available, but attributed to all observations.

2. To arrive at *actual* QP participation probabilities, we then interact the hypothetical probabilities with dummy variables that indicate whether or not the QP had been implemented in the municipality and year in question.

3. Finally, we use the *actual* (availability-interacted) probabilities as instruments for QP participation in regressions with economic outcomes as the dependent variables. In these equations, we control for the hypothetical probabilities and their interactions with time and municipality type (as defined by the timing of QP implementation), as well as for other potential confounders. In particular, we fully control for different time developments in different municipalities.

Somewhat simplified, the quasi-experiment we have in mind here consists of a treatment and a control group in which members – provided that the QP program becomes available – have exactly the same chances of participating, and otherwise face exactly the same economic developments. The treatment group members live in a municipality that introduced the program in, say, 2008, with the others in a municipality introducing the program in 2010. Looking at outcomes in, say, 2011, the former will then have had positive probabilities of entry over the whole 2008-2011 period, whereas the latter could only have entered in the last two years. These differences in probabilities are what we will interpret as being as good as randomly assigned from the prospective participants' point of view. And causal effects will materialize in the form differences in the correlation between outcomes and participation chances in the treatment versus the control group.

The statistical model we use portrays a person i, who in some *base-year* t, belongs to a risk group of potential QP entrants during the next four years ($t+1,\ldots,t+4$), provided that the QP becomes available in person *i's* municipality during this period.[3] To define the potential risk group, we rely on the eligibility criteria set out in the legislation, and include all persons aged 18-55, who in a base-year *t* received some kind of *temporary* income support (except sick pay) from the welfare state, and at the same time had low previous labor earnings and thus low (or no) social insurance entitlements in the coming years.[4] We use the term *potential* risk group to emphasize that they are *actually* at risk only if the program becomes available in the municipality during the next four years.

The base-years used in our analysis are 2000-2007, with outcomes measured in 2001-2012. This implies that all the base-years are strictly prior to the first local implementation of the program, thereby ensuring (by construction) that there is no QP participation in the base-years. It also implies that the base-year observations recorded in the first part of our data window (2000-2003) are never exposed to the risk of *actual* QP participation in the period we follow them (since we consider the entry risk to be negligible after four years and since the first municipalities introduced the program in 2008), whereas subsequent base-year cohorts are exposed to an increasing extent. This pattern is illustrated in Table 2. Note that the only group exposed to QP risk in the first year after the base-year (*i.e.*, in *t+1*) is the 2007 cohort in early reform municipalities. This is also the only group that can be followed for as much as four years after QP entry in our data.

Table 2: Entry possibilities in the four-year period after the base-year, by base-year and municipality reform year

	Reform year 2008 (55% of population)	Reform year 2009 (37% of population)	Reform year 2010 (8% of population)
Base-year:			
2000	No entry possibility	No entry possibility	No entry possibility
2001	No entry possibility	No entry possibility	No entry possibility
2002	No entry possibility	No entry possibility	No entry possibility
2003	No entry possibility	No entry possibility	No entry possibility
2004	2008	No entry possibility	No entry possibility
2005	2008, 2009	2009	No entry possibility
2006	2008, 2009, 2010	2009, 2010	2010
2007	2008, 2009, 2010, 2011	2009, 2010, 2011	2010, 2011

[3] We disregard the risk of QP entry more than four years after the base-year. In principle, we could have also modeled entry in a fifth year since the outcomes are modeled up to five years after the base-year. But, as becomes clear when we explain our statistical approach, this would have considerably complicated the model without adding anything of substance (the probability of entering the QP five years after the base-year is approximately 0.3%).

[4] The included benefits are unemployment insurance, temporary disability insurance (not including sick pay, which is payable for a maximum of one year), and social assistance. In the main specification of our model, the definition of low previous labor earnings is that max (last year's earnings, the average earnings over the last three years) does not exceed NOK 170,000 (measured in 2013 value). By comparison, the average full-time equivalent annual salary for all employees in Norway in 2013 was approximately 500,000).

Now, let $y_{i,t+r}$ be a labor market *outcome* in year $t+r$ for a base-year observation belonging to year t, and let $QP_{i,t+r-p}$ be indicator variables equal to one for persons who entered the QP p years before the outcome year in question (and zero otherwise). Furthermore, let $\mathbf{x_{it}}$ be a vector of observed individual characteristics measured no later than the base-year, incorporating detailed information on demographics, human capital, labor earnings, transfers and number of months with social assistance.[5] In the absence of unobserved sorting (i.e. if we were willing to assume conditional independence in the sense that participation is randomly assigned, conditional on observed characteristics), we could have regressed the various outcomes on a vector of program participation indicators, e.g.:

$$y_{i,t+r} = \mathbf{x}_{\mathbf{it}}^{'}\boldsymbol{\beta} + \alpha_{m\times t\times r} + \sum_{p=0}^{4}\lambda_p QP_{i,t+r-p} + u_{i,t+r}, \quad t = 2000,...,2007, \quad r = 1,...,5, \qquad (1)$$

where $\alpha_{m\times t\times r}$ is a fixed effect for all combinations of municipality (m), base-year (t) and years since base-year (r) (with municipality assigned in the base-year) and $u_{i,t+r}$ is a residual.[6] Here, λ_p represents the effect on the outcome of having entered the program p years ago, relative to not having entered the program at all.

The assumption of conditional independence is unconvincing in this case. Given the character of the program (in particular its explicit targeting of individuals under a high risk of becoming completely reliant on means-tested social assistance), we expect actual participants to be negatively selected, even conditional on observed characteristics. Redesigning (1) to include *individual fixed effects,* and thus exploit the before/after treatment dimensions illustrated in Figure 1 for actual QP participants, would also *not* solve the problem, since participants' earnings and employment profiles prior to entry have been anything but fixed, and since it is probable that QP participation is triggered by unobserved events and/or by changes in attitudes/motivation that in any case would have broken pre-program outcome trends.

---

[5] In our main specifications, the $\mathbf{x_{it}}$-vector includes age (44 categories), gender, years of education (eight categories), immigrant status (three categories), labor earnings (in the base-year and on average during the three years leading up to the base-year), social insurance benefits, social assistance, number of months with social assistance, number of months with UI claims and number of months with temporary disability insurance benefits.

[6] We apply this very flexible specification of municipality time effects throughout the paper, implying the use of 14,624 dummy variables. Note that this is more general than what we would obtain from more "standard" municipality-year dummies, because we allow the effects in a given municipality in a given calendar year to depend on time since base-year, i.e. the time since the condition of being dependent on temporary income support was imposed.

As explained above, we instead rely on an instrumental variables approach, whereby we use the rollout of the program during 2008-2010, *interacted with individual predicted participation propensities,* as instruments. A preparatory step in this empirical strategy is to construct the instruments for the four different entry alternatives that may become relevant (1, 2, 3 or 4 years after the base-year). We do this by estimating individual participation propensities as functions of observed explanatory variables on the basis of the set of baseline observations for which there is a genuine risk of entering the QP. As it turns out, the selection process into the QP appears to have varied quite a bit over time and between municipalities with different reform implementation years (Schafft and Spjelkavik, 2011). Therefore, we cannot simply estimate a single QP participation propensity. Instead, we estimate separate QP entry probabilities for each relevant combination of base-year, potential entry year and municipality-specific reform-year. Let $z_{i,t,q,r}$ be the predicted probability that person *i* observed in a base-year *t* living in a municipality implementing QP in year *q* entered the program *r* years after the base-year. We then compute the predictions

$$z_{i,t,q,r} = f(\tilde{\mathbf{x}}_{\mathbf{it}}^{'}\hat{\boldsymbol{\gamma}}_{\mathbf{t},q,\mathbf{r}}), \quad t = 2004,...,2007, \quad q = 2008,...,2010, \ r = 1,...,4, \qquad (2)$$

separately for all the (19) existing combinations of *t, q* and *r* in our data (see Table 2). In our main specification, we specify *f(.) as exp(.)/(1+exp(.)), i.e.* we use binomial logit models, but the results presented later in this paper would have been almost exactly the same had we used a linear specification instead. The vector $\tilde{\mathbf{x}}_{\mathbf{it}}$ in Equation (2) contains the same variables as $\mathbf{x}_{\mathbf{it}}$ in Equation (1), but with scalar variables instead of indicator variables.[7]

Now, although Equation (2) can be estimated on the subset of actual risk groups only, the 19 resultant predictions can be attributed to all persons and base-year observations in the dataset (since $\tilde{\mathbf{x}}_{\mathbf{it}}$ is always observed). They can then be interpreted as predicted annual *hypothetical* entry probabilities *had* the respective combinations of *t, q,* and *r* become relevant for *i*. For further use, we stack them in a (19×1) vector denoted $\mathbf{z}_{it}$.

[7] The reason for this is that we need a more restrictive specification in this case to avoid problems with no variation in QP entry for particular values of covariates. The following variables are included in $\tilde{\mathbf{x}}_{\mathbf{it}}$: age, gender, education level, immigrant status, earnings in the base-year, max of earnings in the base-year and in the last three years leading up to the base-year, taxable benefits in the base-year, non-taxable benefits in the base-year, number of months with social assistance in the base-year, number of months with UI benefits in the base-year and number of months with temporary disability benefits in the base-year.

Turning back to Equation (1), we note that what we need in order to instrument the endogenous variables $QP_{i,t+r-p}$ are predicted actual QP entry probabilities for the *0-4* years *prior to the outcome year in question*. For each outcome year *t+r,* we construct these predictions such that they represent the corresponding estimated probability $(z_{i,t,q,r})$ if the program was available for the (*t*,*q*,*r*) combination in question, and zero otherwise. In doing this, we end up with five instruments representing actual entry probabilities timed relative to the outcome year, which we denote $\mathbf{z}_{it}^{*} = \begin{bmatrix} z_{i,t+r}^{*} & z_{i,t+r-1}^{*} & z_{i,t+r-2}^{*} & z_{i,t+r-3}^{*} & z_{i,t+r-4}^{*} \end{bmatrix}$. The first- and second-stage equations of our two-stage least squares (2SLS) model thus become:

$$QP_{i,t+r-p} = \mathbf{x}_{\mathbf{it}}^{'}\boldsymbol{\beta}^{qp} + \alpha_{m\times t\times r}^{qp} + \mathbf{z}_{it}^{*}{}'\boldsymbol{\sigma}^{p} + \mathbf{d}_{\mathbf{t}\times r}^{'}\mathbf{z}_{\mathbf{it}}\boldsymbol{\tau}^{qp} + \mathbf{d}_{\mathbf{q}}^{'}\mathbf{z}_{\mathbf{it}}\boldsymbol{\rho}^{p} + \zeta_{i,t+r-p},\ p = 0,...,4, \tag{3}$$

$$y_{i,t+r} = \mathbf{x}_{\mathbf{it}}^{'}\boldsymbol{\beta} + \alpha_{m\times t\times r} + \sum_{p=0}^{4}\left(\lambda_{p}\widehat{QP_{i,t+r-p}}\right) + \mathbf{d}_{\mathbf{t}\times\mathbf{r}}^{'}\mathbf{z}_{\mathbf{it}}\boldsymbol{\tau} + \mathbf{d}_{\mathbf{q}}^{'}\mathbf{z}_{\mathbf{it}}\boldsymbol{\rho} + \varepsilon_{i,t+r}, \tag{4}$$

for *t=2000,…,2007 and r=1,…,5,* where $\mathbf{d}_{\mathbf{t}\times\mathbf{r}}$ is a vector of base-year × outcome-year dummy variables (one dummy for each possible combination of base-year and outcome-year), $\mathbf{d}_{\mathbf{q}}$ is a vector of reform-year dummy variables (time constant, but varying across municipalities with different reform implementation years) and $\widehat{QP_{i,t+r-p}}$ are the OLS predictions from (3).[8] The coefficients of interest are λ_p, *p=0,…4,* which represent the effects of having entered the QP in the same year (*p=0*), the last year (*p=1*) and so forth, in all cases relative to non-participation.

The rationale behind including the control variables $\left\{\mathbf{d}_{\mathbf{t}\times\mathbf{r}}^{'}\mathbf{z}_{\mathbf{it}}, \mathbf{d}_{\mathbf{q}}^{'}\mathbf{z}_{\mathbf{it}}\right\}$ in the statistical model, in addition to those already included in (1), is as follows:

[8] Note that we could have substituted these predictions directly for the QP participation indicators in Equation (1). However, we would then not end up with a correct 2SLS model, since there are a significant number of actual QP entries that are recorded in municipalities and/or years for which the program does not exist. The most likely explanation for this is that there are errors in municipality assignment or that persons have migrated to another municipality after the base-year. This represents a potential source of contamination bias (the non-participant group is contaminated with a number of participants), which unaccounted for will bias estimated effects toward zero.

- $\mathbf{d}'_{\mathbf{t\times r}}\mathbf{z}_{\mathbf{it}}$ (8 base years × 5 outcome years × 19 hypothetical entry probabilities = 760 variables) is included to control for any differences in the outcomes and their time-developments that correlate systematically with the QP participation propensities.[9]
- $\mathbf{d}'_{\mathbf{q}}\mathbf{z}_{\mathbf{it}}$ (3 QP implementation years × 19 hypothetical entry probabilities = 57 variables) is included to control for any differences in the correlation between QP propensities and outcomes between municipalities that implemented the reform at different times.

As a result, by including these controls, we narrow down the variation in participation propensities used to identify the causal effects to the desired quasi-experimental part of it.

Because the instruments used to identify the causal effects of program participation incorporate the phasing-in of the program itself, all actual participants must have been directly induced to participate by the instruments. In the terminology used by Angrist *et al.* (1996), all actual participants are "compliers," and there exists no "always-takers." Provided that the QP influences the clients' outcomes through actual participation in the program only, our statistical approach thus identifies the average treatment effects among the treated (ATET).

A final point to note is that, as we show in the next section, many of the individuals in our dataset qualify for being included in the risk-group in more than one base-year. Given the way we construct the analysis data, these persons will contribute with multiple – and sometimes overlapping – five-year outcome sequences. We have done this to ensure a completely symmetric risk group composition throughout the data window, while at the same time exploiting as much of the information content in the data as possible. However, we take the multiplicity of observations into account when we compute standard errors. This is done in two ways. For all the coefficients presented in this paper, we have computed standard errors based on clustering at the individual level. For the main results, we have also performed a complete non-parametric bootstrap with 600 replications, each based on a full resampling of base-year observations with replacement.[10] As we show in Section IV, these standard errors are very similar, although there is a slight tendency for the clustered standard deviations to exaggerate statistical uncertainty.

[9] When participation propensities $\mathbf{z}_{\mathbf{it}}$ are estimated by linear probability models, we impose one reference (zero-restriction) for each of the 19 entry routes to avoid perfect colinearity with $\mathbf{x}_{\mathbf{it}}$.

[10] As there are five outcome observations for each base-year observation in the actual data population, we also include all five outcome observations for each selected base-year observation in the bootstrap.

III. The Analysis Population

In this section, we give a brief description of the analysis population used in the statistical analysis. Given the rather vague eligibility criteria, it is not a trivial exercise to identify the population at risk of entering the QP over a forthcoming four-year period based on observed characteristics. The formal rules described in Section I target persons with a substantially reduced labor income capacity and no or very limited social insurance entitlements. In principle, this implies that everyone who have had low labor earnings over some time, and also received some kind of temporary income support from the welfare state, may become eligible. Since social insurance entitlements in Norway generally depend on labor earnings during the last calendar year and/or the average earnings over the last three years, we base our definition of "low labor earnings" on the maximum of these two amounts. Deciding on the location of earnings threshold involves a tradeoff, as over a forthcoming four-year period, persons may become eligible for QP almost regardless of previous earnings. By setting the threshold low, we "get rid of" the least likely QP participants, but at the cost of also throwing out a number of persons who nonetheless made it into the program. By setting the threshold high, we ensure the inclusion of more actual participants, but at the cost of including a larger number of persons for which the program never becomes relevant. In our main specification, we set the ceiling threshold to NOK 170,000 ($29,000). This is roughly one-third of the average full-time, full-year earnings level in Norway, and corresponds to the limit set for eligibility to full unemployment benefits. In a robustness analysis, we increase the threshold to 340,000. This raises the overall sample size by 65% and the number of actual participants covered by 11%.

Table 3: Descriptive statistics analysis population

	Participants	Non-participants
Number of base-year observations	21,082	1,386,310
Number of individuals	8,896	307,003
Mean age	32.5	36.7
Women %	61.0	46.8
Non-native %	36.2	15.3
High school as highest education %	17.8	35.8
College/University as highest education %	6.1	10.8
Mean labor earnings base-year (NOK 1,000, 2013)	19.5	28.8
Mean social assistance base-year (NOK 1,000, 2013)	103.4	54.4
Mean taxable benefits base-year (NOK 1,000, 2013)	50.8	135.4

Note: Averages and fractions are computed over base-year observations; NOK 1,000 = $170 (based on the average exchange rate in 2013).

Table 3 shows some descriptive statistics for the sample used in the main part of the analysis. There are 315,899 individuals included, out of which 8,896 (2.8%) actually participated in the QP. This reflects that our definition of the "risk group" is very wide (including everyone with temporary income support and low previous earnings), and that only a tiny minority in practice will end up in such a bad situation that the QP is a relevant alternative. The participants tend to be quite different from the non-participants, e.g. in the form of lower labor earnings and taxable social insurance benefits, higher levels of social assistance and lower levels of education. Women and immigrants are significantly overrepresented in the participant group.

On average, each individual contributes 4.5 baseline-year observations, and thus $5 \times 4.5 = 22.5$ (partly overlapping) outcome-year observations. A point to note here is that our statistical analysis only includes roughly half of the 19,211 QP participants described in Section I above. The primary reason for this is that to ensure a completely symmetric composition of analysis populations in municipalities with different QP implementation dates, our risk groups are defined on the basis of individual characteristics recorded strictly prior to program implementation, i.e. no later than 2007. As a result, we lose persons who entered the risk group for the first time during 2008-2011, and this alone accounts for 62% of the overall loss of actual participants in the analysis data. In addition, our definition of the risk population is imperfect, implying that some persons enter the QP even though they were not considered (by us) to be at risk, i.e. they had too high an income in the base-year, were too old or did not receive the types of temporary income support that we have used to define the risk population. The latter is particularly relevant for humanitarian immigrants, who sometimes enter the program directly without first receiving the temporary transfers. For this reason, the participant group included in our statistical analysis deviates somewhat from the group of all participants described in Table 1. In particular, we oversample female and under-sample immigrant participants.

IV. Results

We are interested in the effects of QP participation on a number of outcomes. Since the main aim of the program is to help persons into regular employment, we focus on regular labor earnings and employment as the main success criteria. We start out this section by summarizing the computation of the QP participation propensities, and show some illustrative graphical evidence. Thereafter, we turn to the estimation results from the statistical model.

A. Participation Propensities and Labor Earnings Profiles

Given that we estimate as many as 19 different participation models based on Equation (2) (depending on base-year, the number of years that have passed since the base year and the timing of the reform in the municipality of residence), we do not present these results in any detail. What all the regressions show is that observed individual characteristics have a considerable influence on the participation propensities. Those with the highest participation probabilities are young immigrants with almost zero labor earnings and little schooling, and who have virtually no access to (taxable) social insurance transfers and relatively large amounts of means-tested social assistance. At the other end, those with the lowest participation probabilities are older natives with some previous labor earnings and schooling, and significant social insurance transfers, but with virtually no means-tested social assistance.

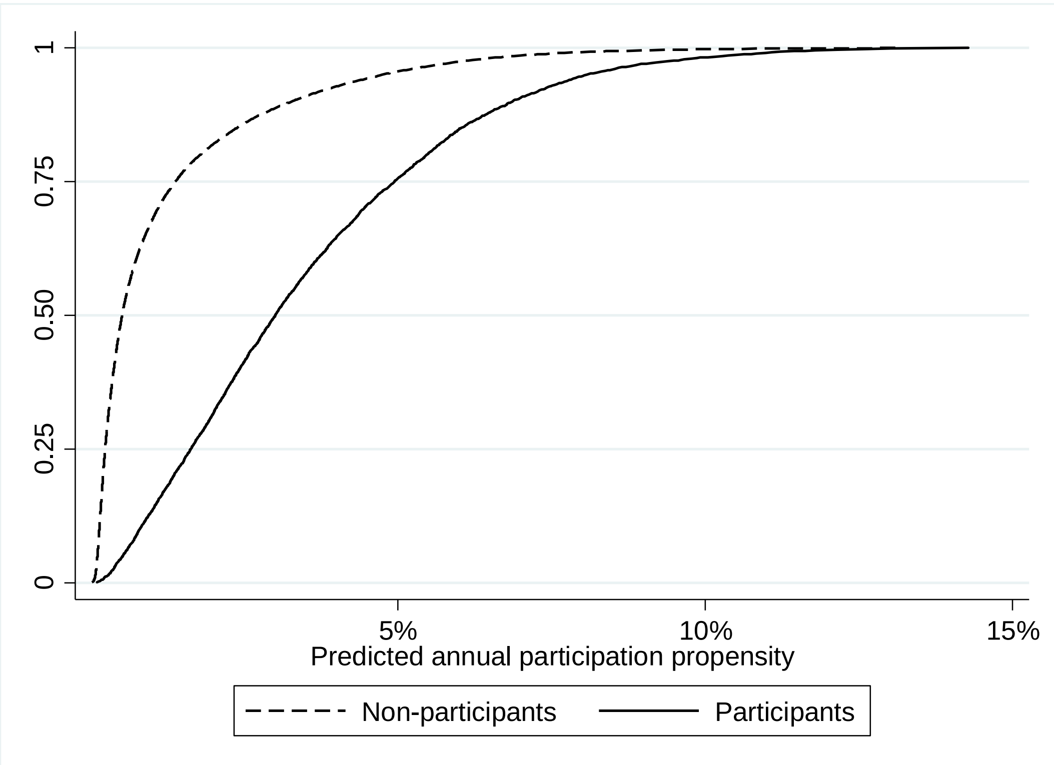

Figure 2: The cumulative distribution of predicted average annual QP participation probabilities

Note: The distribution functions are based on the averages of the 19 hypothetical QP propensities predicted from Equation (2).

To help illustrate the variation in predicted participation propensities among participants and non-participants, Figure 2 shows distribution functions for the average of the 19 probability predictions made for each individual base-year observation, for participants and non-participants, respectively. It is clear that there is a marked difference between the two groups, but that the average predicted annual QP entry probabilities (taken over the 19 possible entry routes) are low for virtually everyone: 3.5% for participants and 1.2% for non-participants. According to our model, provided independence in probabilities over entry years, this implies that over a four-year risk period, actual participants had a participation probability around 13.3 %, whereas non-participants had a participation probability of roughly 4.7%. The relatively low participation probabilities estimated, even for participants, reflects that it is difficult to identify a genuine high-risk group based on observed characteristics only, thereby highlighting the magnitude of the selection problem. It also illustrates why it would not be a good idea to base a causal evaluation of QP on a comparison of participants and non-participants, even with very good data (in terms of observed individual characteristics).

Rather, our empirical approach relies on a comparison of persons with high and low participation propensities before and after local implementation of the QP, controlled for differential time effects and geographical differences. To illustrate how this identification strategy plays out in practice, we present a graphical difference-in-differences analysis in Figure 3 based on a comparison of potential participants with different participation probabilities in the municipalities that implemented the QP first (in 2008) with those living in the municipalities that implemented it last (in 2010). The basis for this exercise is persons belonging to the risk population in 2007. The outcome period we look at in this case is the five-year period following this base-year (2008-2012). We would obviously expect causal QP effects to be larger in the early-implementing municipalities (as many of the risk-group members in 2007 would no longer be at risk in 2010), and any QP effects should also show up there with a two-year lead. Differences should also be larger the higher the predicted participation propensity. We look at four groups with increasingly higher participation propensities: the upper quintile (25%), the upper decile (10%), the upper vigintile (5%) and the upper percentile (1%), in all cases relative to the lower quintile. To eliminate noise arising from persistent geographical variations, we subtract the corresponding differences that prevailed in the five-year period prior to the reform (i.e. in 2003-2007, based on the population (hypothetically) at risk in 2002); hence, the numbers reported in Figure 3 are the triple differences along the dimensions of participation propensity, the municipality’s time of implementation and the calendar time period. For example, focusing on the outcomes four years after the QP launch (2012), the result indicated for the upper percentile in the QP propensity distribution that the most likely QP participants would have had approximately NOK 19,000 higher earnings in 2012 if they lived in a municipality with implementation in 2008, than if they lived in a municipality with implementation in 2010, after having subtracted the difference for less likely participants (the lowest quintile) in the same municipalities and years, in addition to the corresponding difference-in-difference five years earlier.

A relatively clear pattern emerges: There are indications of negative earnings effects in the year of QP launch as well as in the year after. The effects then tend toward the positive side, but at this stage it is important to bear in mind that the effects are measured against the presumed negative launch effects in the municipalities with a late implementation. Four years after the program launch, there appears to be clear positive effects. All the effects – both the initial lock-in effects and the subsequent positive effects – appear to be larger the higher up in the participation propensity distribution we go.

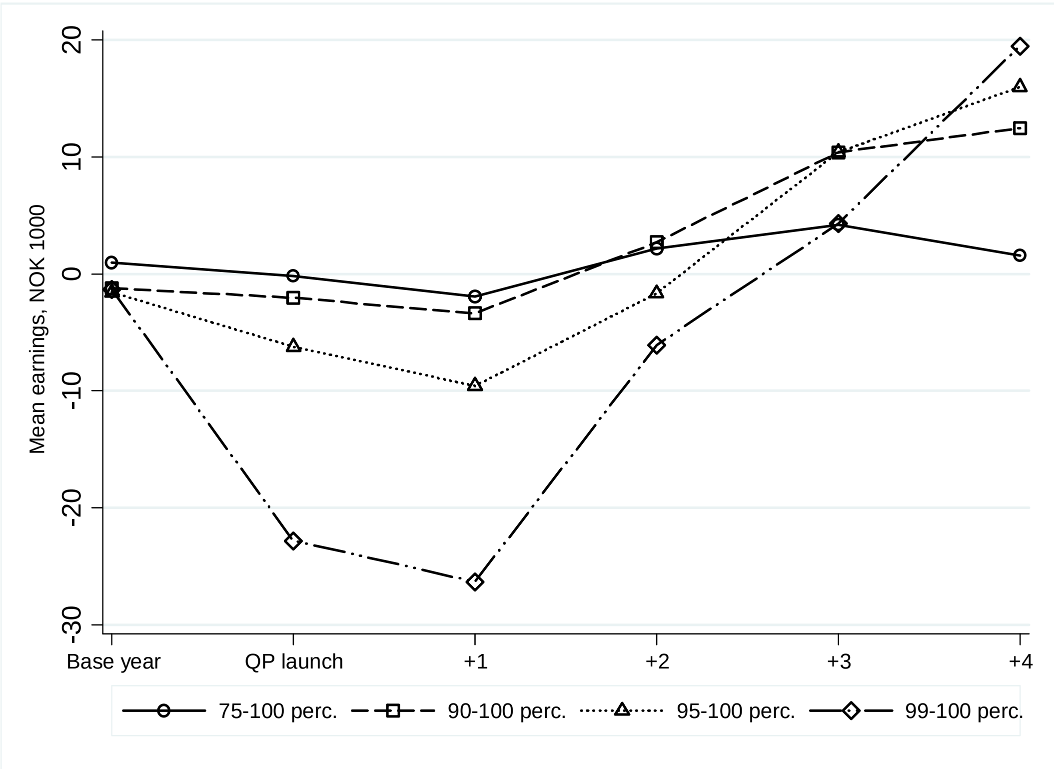

Figure 3: DiD estimates of the effect of QP program implementation (by position in participation propensity distribution relative to first quintile and by time since program launch)

Note: The reported numbers are the (triple) differences in annual earnings based on: i) relevant QP participation propensities (highest quintile/decile/vigintile versus lowest quintile), ii) time of local reform implementation (2008 versus 2010), and iii) base-year (2007 versus 2002).

B. Main Estimation Results

We now turn to the estimation results from the 2SLS model (Equations 3 and 4). In this section, we focus on the results of substantive interest, i.e. the second-step equations. The first stage results are presented in the Appendix, and they confirm that our instruments are strong, even in a multiple endogenous variable setting. In addition to using annual labor earnings as a continuous outcome variable, we derive discrete employment outcomes based on alternative annual earnings thresholds. Given the program's target group of hard-to-employ persons, we set these thresholds at relatively low levels. We also use outcomes measuring the level of various types of welfare state transfers. For these outcomes, we of course have considerable a priori knowledge about the true causal effects of the QP program, since *by design* the program offers a taxable full-year-equivalent transfer of NOK 170,000 ($29,000), which to some extent substitutes for non-taxable benefits (means-tested social assistance). This implies that we can use the models' estimated effects on these outcomes as a sort of plausibility test. In total, we specify six annual outcomes, each measured over a five-year period (the year of entry and the subsequent four years):

a) Employment, defined as having annual labor earnings above NOK 85,000 ($14,500);
b) Employment, defined as having annual labor earnings above NOK 170,000 ($29,000);
c) Annual labor earnings (measured in NOK 1,000);
d) Log annual labor earnings (log (earnings measured in NOK 1,000 plus 1));
e) Taxable benefits (measured in NOK 1,000);
f) Non-taxable benefits (measured in NOK 1,000).

Figure 4 summarizes the main results of this paper in the form of estimated causal effects of QP participation on our six outcome variables by year since QP entry. These estimates are based on our preferred instrumental variables (2SLS) specification (Equation 4), and we present point estimates together with 90% confidence intervals. Table 4 reports the estimation results with standard errors in more detail, also including a comparison with the naïve OLS estimates. Standard errors and confidence intervals for the main 2SLS results are based on a full nonparametric bootstrap (600 replications); see Section II. Given the computational cost associated with this bootstrap procedure, other results in this paper are based on standard errors clustered at the individual level.[11] We show in the next subsection that the cluster-based

[11] Note that any shocks common at the municipality level are absorbed by the municipality×base-year×outcome-year×fixed effects.

standard errors are very close to the bootstrapped ones, with the latter actually being a bit smaller.

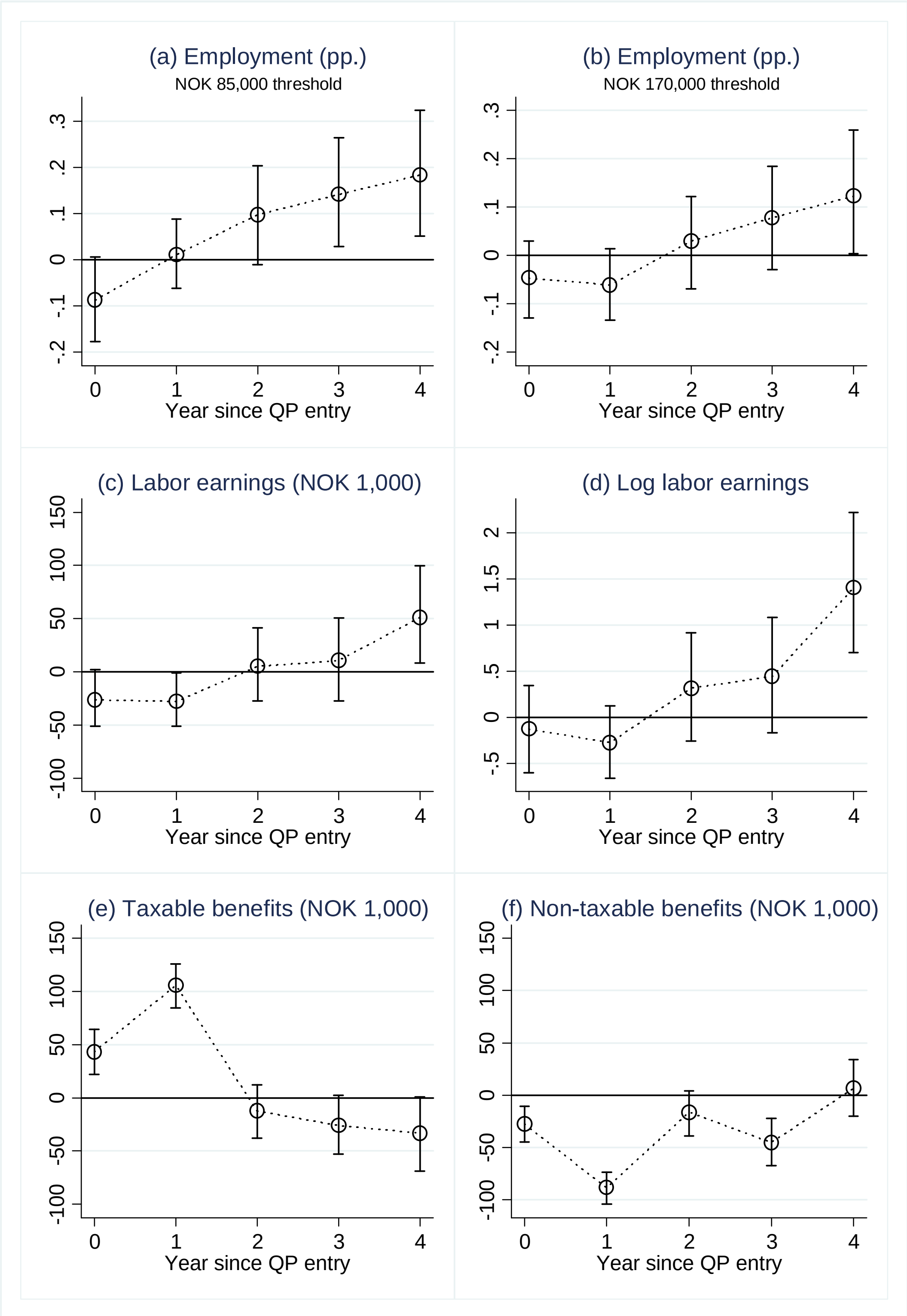

Figure 4: Main estimation results from the 2SLS model (Equation 4) with 90% confidence intervals.

Note: Confidence intervals are based on the nonparametric bootstrap with 600 replications (with replacement).

Based on the widest employment definition (annual earnings exceeding NOK 85,000), the 2SLS estimates indicate a 10 percentage point increase in the employment probability two years after QP entry, a 14 point increase after three years and an 18 point increase after four years (Figure 4, panel a).[12] The estimated impacts on the levels of annual earnings follow a similar time pattern (panel c); after four years, QP participation is predicted to raise annual earnings by approximately NOK 50,000 ($8,500). Given that the average full-time equivalent annual earnings level in Norway is roughly NOK 500,000, this is in quantitative terms not a huge effect. It is considerable, though, relative to the average labor earnings level of only NOK 19,500 among actual participants in the base-year. Because the QP effects appear to be concentrated in the extreme lower tail of the earnings distribution, the linear earnings equation probably does a poor job in representing them. Using a log-specification instead (panel d), we estimate a QP impact of 1.4 log-points after four years, thereby indicating that those who actually participate in the program (and typically have very poor earnings prospects) experience large *relative* increases as a result of QP participation.

The estimated impacts on employment and earnings may also be put in perspective by noting from the descriptive statistics presented in Figure 1 that *observed* average earnings among participants four years after entry was approximately NOK 77,000, whereas the employment rate (based on the lowest earnings threshold) was 28%. Our estimates suggest that without program participation, annual earnings would have been as low as NOK 27,000, and that the employment rate would have remained at about 10%, i.e. very close to the levels actually observed around the time of entry.

Although the statistical uncertainty associated with all the employment and earnings estimates is relatively large when evaluated one-by-one, it is notable that the overall pattern of estimated coefficients conveys a coherent and plausible story: QP participation reduces employment and labor earnings slightly during the first one-two years of participation and raises them afterwards. At the same time, it sharply (and statistically significantly) increases the level of taxable benefits (panel e) and reduces the level of social assistance (panel f). The sizes of the latter effect-estimates correspond closely to what we would expect on the basis of prior knowledge. For example, the estimated rise in taxable benefits of NOK 106,000 in the first year after QP entry accords well with the fact that annual full-year benefits are equal to NOK

[12] Recall that the effects reported for different number of years after entry are based on different parts of the sample; hence, if the effects are heterogeneous, the dynamic effect pattern may also reflect changes in participant composition. In particular, the effects four years after entry are identified solely on the basis of 2007 base-year observations in municipalities implementing the reform in 2008.

170,000, as some QP entrants would have been eligible for small amounts of taxable benefits even without the program and some entrants do not participate for the whole year. Still, it may appear somewhat surprising that the overall income support level (the sum of taxable and non-taxable benefits) increases only slightly as a result of QP entry.[13] This finding suggests that many of the QP participants would have obtained a considerable increase in income support even in the absence of QP, and that relative to realistic alternatives, QP is not that "generous" after all. From a pecuniary point of view, the main difference between the QP benefit and alternative income support sources (particularly social assistance) may be that QP offers economic stability and security, and perhaps some sense of dignity.

Given that we summarize the main results with as much as 30 coefficient estimates, one may question whether some of them display statistical significance by coincidence. We emphasize, however, that although our statistical inference in Figure 4 and Table 4 is based on separate tests, it is the overall pattern of the results – e.g. in the form of negative employment and earnings effects just after entry, and then monotonously increasing positive effects afterwards – that we interpret as convincing evidence that QP actually had the intended impacts. Our primary interest, though, lies in identifying the post-program effects, as one may expect that these have some bearing on longer term employment and earnings outcomes. The closest we get to that in our analysis are the impacts identified for the fourth year after entry, since we expect that almost all participants have completed the program at this point. For this specific period, we therefore compute a joint test based on the 600 bootstrap estimations, reflecting the hypothesis that the QP failed to achieve improved employment prospects according to at least one of the employment/earnings measures. A test of this hypothesis gives rejection with a p-value of 0.05, i.e. in 95% of our 600 bootstrap samples, we obtain positive employment/earnings effect in the fourth year regardless of which of the outcomes we focus on. If we *also* require a negative effect on total transfers (the sum of taxable and non-taxable benefits), the results are less certain: In 83% of the samples, we obtain both positive earnings and employment effects *and* negative effects on total transfers in the fourth year after entry.

[13] It is not possible for us to identify the tax consequences of a given increase in "taxable benefits," but the earnings levels typical for QP participants entail very low tax rates.

Table 4: Main estimation results from the 2SLS (Eq. 4) and OLS (Eq. 1) models (clustered (OLS) or bootstrapped (2SLS) standard errors in parentheses)

	Employment (NOK 85,000 threshold)		Employment (NOK 170,000 threshold)		Labor earnings (NOK 1,000, 2013 value)		Taxable benefits (NOK 1,000, 2013 value)		Non-taxable benefits (NOK 1,000, 2013 value)		Log labor earnings (log(earnings+1)) (NOK 1,000, 2013 value)	
	I Eq. 1 (OLS)	II Eq. 4 (2SLS)	III Eq. 1 (OLS)	IV Eq. 4 (2SLS)	V Eq. 1 (OLS)	VI Eq. 4 (2SLS)	VII Eq. 1 (OLS)	VIII Eq. 4 (2SLS)	IX Eq. 1 (OLS)	X Eq. 4 (2SLS)	XI Eq. 1 (OLS)	XII Eq. 4 (2SLS)
Effects of QP												
Same year (p=0)	-0.201*** (0.003)	-0.086 (0.057)	-0.177*** (0.002)	-0.047 (0.049)	-60.02*** (0.66)	-26.33 (16.35)	12.70*** (0.68)	43.50*** (12.49)	45.18*** (0.68)	-27.00** (10.77)	-0.829*** (0.019)	-0.133 (0.302)
First year after entry (p=1)	-0.145*** (0.004)	0.010 (0.047)	-0.128*** (0.003)	-0.061 (0.045)	-47.00*** (0.88)	-28.18 (15.14)	55.41*** (0.83)	106.50*** (12.68)	21.57*** (0.72)	-88.12*** (9.65)	-0.735*** (0.021)	-0.267 (0.233)
Second year after entry (p=2)	-0.075*** (0.005)	0.096 (0.064)	-0.071*** (0.004)	0.028 (0.058)	-28.35*** (1.20)	4.25 (20.03)	30.01*** (1.05)	-10.92 (15.17)	11.59*** (0.81)	-16.01 (12.62)	-0.448*** (0.026)	0.304 (0.348)
Third year after entry (p=3)	-0.042*** (0.006)	0.143** (0.072)	-0.034*** (0.005)	0.081 (0.064)	-15.34*** (1.64)	12.18 (22.76)	19.86*** (1.31)	-25.23 (17.27)	0.54 (0.94)	-45.84*** (13.56)	-0.261*** (0.033)	0.465 (0.383)
Fourth year after entry (p=4)	-0.012 (-0.010)	0.182** (0.086)	-0.008 (0.009)	0.121 (0.077)	-5.80** (2.72)	50.54* (27.86)	23.53*** (2.12)	-32.08 (21.88)	-8.23*** (1.58)	6.73 (16.65)	-0.096* (0.053)	1.395*** (0.466)

Note: For the 2SLS models, the reported standard errors are based on a nonparametric bootstrap with 600 replications (sampling with replacement). For the OLS models, standard errors are clustered at individuals.

*(**)(***) indicate significance at the 10(5)(1) % levels.

Number of observations in all models: 7,036,980. NOK 1,000 = $170 (based on average exchange rate in 2013).

Control variables included in all models: Municipality×base-year×years-since-base-year-fixed effects (18,280 dummy variables), age (44 dummy variables), gender, education (eight dummy variables), immigrant status (three dummy variables), labor earnings (in the base-year and as average over the three years leading up to the base-year), taxable benefits in the base-year, non-taxable benefits in the base-year, number of months with social assistance in the base-year, number of months with UI benefits in base-year and number of months with temporary disability benefits in base-year.

Additional control variables in IV models: Estimated QP participation propensities interacted with base-year×outcome-year (760 variables), and estimated QP participation propensities interacted with reform-year in the municipality (57 variables).

As pointed out in Section II, provided that QP affects the performance of actual participants only, our 2SLS effect estimates can be interpreted as the estimated average treatment effect among the treated (ATET). If there are indirect effects on non-participants, e.g. in the form of "threat effects" or in the form of changes in the local treatment culture that spill over to other social assistance recipients, our 2SLS estimates are still valid as measures of the overall effects *relative to the number of treated*. But since they attribute all (reduced form) effects to the actual participants, they will in this case underestimate the number of affected clients and thus overestimate the causal effect for each of them.

For comparison, Table 4 also presents OLS estimates, based on the conditional independence assumption (Equation 1). It is of some interest to note that they consistently – and with overwhelming statistical significance – indicate large adverse effects of QP. Hence, despite our fairly impressive vector of observed covariates, including a wide range of individual characteristics and dummy variables for all combinations of base-year, outcome-year and municipality, OLS estimates appear to be seriously biased. This illustrates the danger of relying on conditional independence assumptions in order to evaluate the causal effects of programs with selective participation. In particular, it confirms our suspicion of a large negative (remaining) selection into QP even conditional on everything we are able to observe; persons are allowed into this program only when they are in serious trouble, and our observed covariates are far from accounting properly for this selection.

Since the OLS results generally go in the opposite direction of the 2SLS estimates, it may also be noted that if we had a problem with weak instruments – which we arguably do not; see the Appendix – this would imply that the main results reported in this section are biased toward zero, and hence represent conservative estimates of the true effects of QP.

C. Robustness

How robust are these results? The statistical model outlined in Section II and estimated in the previous subsection is arguably not particularly transparent; it is highly complicated, and it contains an extraordinarily large number of explanatory variables. This may raise legitimate concerns that the results are not robust, and perhaps that they have come about after a data-mining process involving a number of alternative (not reported) specifications. Note, however, that a much simpler – yet still convincing – difference-in-difference approach is not feasible on the basis of our data. The reason for this is that to ensure the use of reliable identification sources only, we essentially have to extract and exploit very weak signals of causality from data characterized by a lot of disturbing "noise." Recall that even among the most likely QP participants (according to our predictions), only a tiny minority actually enrolls each year, and our model not only builds on a comparison of participants and non-participants, it needs to facilitate inferences about the likelihood of several alternative entry years, in addition to easing out not one estimated effect, but a dynamic effect curve covering five years. And all this must be identified on the basis of a mere three-year phasing-in period. As we have shown above, this is entirely possible given a sufficient amount of data, but to distill the weak signals of interest we actually need a model that effectively controls for all disturbing factors. That is why we have modeled the control functions along the dimensions of municipality (reform-year), base-year, outcome-year and predicted participation propensities in the most flexible manner possible. Along these dimensions, the model is in some sense "saturated" (it cannot be made more flexible without taking away the foundation for identification); hence there is actually little room for trial and error (data mining). We could of course have imposed restrictions on these interaction variables, e.g. by assuming that calendar time effects were the same regardless of QP participation propensity. But such restrictions would clearly be invalid, and thus undermine our identification strategy.

The ways we have selected the analysis population and modeled the influences of individual characteristics are completely different stories. Here, we have had to make choices, both with respect the determination of data-inclusion thresholds and with respect to which explanatory variables to use and how to use them (functional forms). This should not matter for the model's ability to sort out the true causal effects, however, only for the degree of precision. Recall that our identification strategy does *not* rely on the (questionable) assumption that all unobserved heterogeneity is accounted for (conditional independence). Although we cannot evaluate the model's robustness with respect to individual characteristics that we do not observe, we can easily *make unobserved* some of the variables that we do observe. We can also use alternative criteria for inclusion into the analysis population. If our identification strategy is appropriate, including too many (or too few) observations or leaving out important groups of individual explanatory variables should make the model poorer (less precise), but it should not cause bias in the estimated effects of interest. Therefore, in this section we offer robustness exercises with respect to these choices. Moreover, we examine the model's robustness with respect to assumptions regarding the exogeneity of residential address.

In all these exercises, we focus on the main outcomes of interest, i.e. employment (according to two alternative definitions) and log earnings from unsubsidized jobs. We start out assessing the model's robustness with respect to the set of included observed covariates. Table 5 compares the estimates from our main model (with all explanatory variables included) with those obtained when we throw out groups of explanatory variables from the whole estimation exercise (i.e. "pretend" that these variables are unobserved). The reported standard errors in this (and the next) table are based on single estimations only, with clustering at the individual level. Columns I-III repeat some of the main estimates from the previous subsection, and for comparison we also equip these estimates with the cluster-based standard errors. It may be noted that these standard error estimates are somewhat larger than their bootstrapped counterparts used in the previous subsection, thus suggesting that the clustering strategy used here actually entails a slight exaggeration of statistical uncertainty and hence a somewhat "conservative" inference. Columns IV-VI report results from models in which we drop all the demography and human capital variables, whereas columns VII-IX report results from models where we instead drop all the labor earnings and social insurance transfer variables.[14] The results turn out to be highly robust. The point estimates keep telling almost exactly the same story regardless of which covariates we remove. The standard errors become larger, though, suggesting that we actually need to do as much as we can out of the covariates we have in order to facilitate precise estimates and informative statistical inference.

[14] These two regressions are thus based on disjoint sets of individual covariates, except for the variables describing social assistance claims, which appear to be essential for identification. To help illustrate the importance of the excluded variables in terms of identifying QP participation propensities, we ran an auxiliary linear regression using the average of the 19 predicted participation probabilities as the left-hand side variable and the variables originally used to form each of these predictions as right-hand side variables. When all variables are included, we obtain an R-square of 0.78. Leaving out demographics and human capital variables (as in columns IV-VI) reduces the R-square to 0.68. Leaving out labor earnings and social insurance transfers instead (as in columns VII-IX) yields an R-square of 0.70.

Table 5: Estimation results for models with reduced sets of individual covariates; Instrumental Variables estimates (Eq. 4) (clustered standard errors in parentheses)

	Main results based on all individual covariates (repeated from Table 4)			Results without demographic and human capital variables			Results without labor earnings and social insurance transfer variables		
	I	II	III	IV	V	VI	VII	VIII	IX
	Employment (NOK 85,000 threshold)	Employment (NOK 170,000 threshold)	Log labor earnings (log(earnings+1)) (NOK 1,000, 2013 value)	Employment (NOK 85,000 threshold)	Employment (NOK 170,000 threshold)	Log labor earnings (log(earnings+1)) (NOK 1,000, 2013 value)	Employment (NOK 85,000 threshold)	Employment (NOK 170,000 threshold)	Log labor earnings (log(earnings+1)) (NOK 1,000, 2013 value)
Effects of QP									
Same year (p=0)	-0.086 (0.074)	-0.047 (0.065)	-0.133 (0.396)	-0.111 (0.086)	-0.072 (0.063)	-0.287 (0.397)	-0.106 (0.096)	-0.056 (0.084)	0.014 (0.515)
First year after entry (p=1)	0.010 (0.064)	-0.061 (0.058)	-0.267 (0.330)	-0.019 (0.092)	-0.075 (0.057)	-0.467 (0.326)	0.022 (0.078)	-0.077 (0.072)	-0.206 (0.400)
Second year after entry (p=2)	0.096 (0.079)	0.028 (0.071)	0.304 (0.431)	0.091 (0.108)	0.025 (0.070)	0.139 (0.435)	0.117 (0.101)	0.063 (0.090)	0.711 (0.548)
Third year after entry (p=3)	0.143* (0.082)	0.081 (0.073)	0.465 (0.434)	0.155 (0.126)	0.088 (0.071)	0.475 (0.427)	0.110 (0.101)	0.037 (0.091)	0.411 (0.536)
Fourth year after entry (p=4)	0.182** (0.089)	0.121 (0.079)	1.395*** (0.477)	0.174 (0.125)	0.129* (0.076)	1.186** (0.476)	0.207* (0.113)	0.144 (0.101)	1.891*** (0.613)

Note: Standard errors are clustered at individuals. *(**)(***) indicate significance at the 10(5)(1) % levels.
Number of observations in all models: 7,036,980. NOK 1,000 = $170 (based on average exchange rate in 2013).
Control variables included in all models: Municipality×base-year×years-since-base-year-fixed effects (18,280 dummy variables), estimated QP participation propensities interacted with base-year×outcome-year (760 variables), and estimated QP participation propensities interacted with reform-year in the municipality (57 variables).

Table 6: Estimation results for sample with high earnings threshold; Instrumental Variables estimates (Eq. 4) (clustered standard errors in parentheses)

	Whole population			Natives only			Immigrants only		
	I	II	III	IV	V	VI	VII	VIII	IX
	Employment (NOK 85,000 threshold)	Employment (NOK 170,000 threshold)	Log labor earnings (log(earnings+1)) (NOK 1,000, 2013 value)	Employment (NOK 85,000 threshold)	Employment (NOK 170,000 threshold)	Log labor earnings (log(earnings+1)) (NOK 1,000, 2013 value)	Employment (NOK 85,000 threshold)	Employment (NOK 170,000 threshold)	Log labor earnings (log(earnings+1)) (NOK 1,000, 2013 value)
Effects of QP									
Same year (p=0)	-0.043 (0.069)	0.017 (0.062)	0.055 (0.372)	-0.054 (0.079)	-0.023 (0.069)	0.210 (0.471)	0.017 (0.118)	0.080 (0.110)	0.193 (0.624)
First year after entry (p=1)	0.052 (0.060)	-0.008 (0.056)	-0.115 (0.309)	0.028 (0.070)	-0.007 (0.064)	-0.307 (0.407)	0.018 (0.104)	-0.073 (0.100)	-0.107 (0.520)
Second year after entry (p=2)	0.156** (0.074)	0.138** (0.067)	0.668* (0.402)	0.139* (0.083)	0.125* (0.074)	0.808 (0.499)	0.238 (0.146)	0.211 (0.138)	1.124 (0.757)
Third year after entry (p=3)	0.198** (0.077)	0.131* (0.071)	0.629 (0.408)	0.201** (0.085)	0.105 (0.077)	0.718 (0.512)	0.001 (0.181)	-0.094 (0.175)	-0.424 (0.893)
Fourth year after entry (p=4)	0.273*** (0.084)	0.212*** (0.076)	1.848*** (0.449)	0.219** (0.091)	0.132* (0.080)	2.046*** (0.576)	0.404 (0.356)	0.473* (0.254)	2.747* (1.267)

Note: Standard errors are clustered at individuals. *(**)(***) indicate significance at the 10(5)(1) % levels. Number of observations: Column I-III: 11,567,275; Column IV-VI: 9,953,115; Column VII-IX: 1,451,965. NOK 1,000 = $170 (based on average exchange rate in 2013).
Control variables included in all models; see note to Table 4.

Table 6 examines the impact of making alterations in the composition of the analysis population. In particular, and as explained in Section III, we look at how the key estimates change when we double the upper base-year earnings threshold used as a data-inclusion criterion to NOK 340,000 ($58,000), thereby increasing the sample size by 65%. In addition, given the large fraction of immigrants participating in the program, we use this expanded dataset to estimate separate models for natives and immigrants, with the results shown in Table 6. Columns I-III first present estimates based on the whole expanded dataset, which are similar to the estimates reported in Table 4. The main difference is that the estimated employment and earnings effects two-four years after the QP entry become a bit larger and also more statistically significant than in our baseline model. Separate results for natives and immigrants are provided in columns IV-VI and VII-IX, respectively. The results for natives are closely aligned with the results reported for the entire population. The results for immigrants also display a similar pattern, although they are more unstable and subject to larger statistical uncertainty. Hence, there is not enough information in the data to draw conclusions regarding immigrant/native effect differentials.

Lastly, our empirical strategy hinges on the assumption that the unobserved composition of the base-year risk populations was not affected by whether the reform was implemented in 2008, 2009 or 2010. This assumption could have been violated if, for example, would-be QP participants migrated to the municipalities with an early introduction in order to take advantage of the program. Since the QP program was legislated in June 2007, there was a short window of opportunity during the autumn of 2007, whereby would-be QP participants could self-select into municipalities with an early (2008) QP implementation. On average, approximately 11% the risk-group population move across municipalities each year, but we see no pattern of increased migration during the period in question. Nevertheless, to check the sensitivity of our estimates with respect to endogenous migration, we re-estimate our baseline models with each base-year observation tied to the last year's (*t-1*) municipality (i.e. we disregard migrations that occurred between *t-1* and *t*). In this way, we eliminate the possibility of an endogeneity problem, but at the cost of inducing more measurement error into the model. The results of this exercise are presented in Table 7. As it turns out, the estimates change very little, which suggests that endogenous migration is not a big issue in the present context.

Table 7: Estimation results with last year's address used to identify the municipality (dataset with a low earnings threshold); Instrumental Variables estimates (Eq. 4) (clustered standard errors in parentheses)

	I	II	III
	Employment (> NOK 85,000)	Employment (> NOK 170,000)	Log labor earnings
Effects of QP			
Same year (p=0)	-0.050 (0.080)	-0.025 (0.070)	0.102 (0.432)
First year after entry (p=1)	0.024 (0.068)	-0.066 (0.063)	-0.252 (0.354)
Second year after entry (p=2)	0.139 (0.087)	0.065 (0.078)	0.604 (0.473)
Third year after entry (p=3)	0.158* (0.088)	0.086 (0.080)	0.341 (0.471)
Fourth year after entry (p=4)	0.194* (0.099)	0.100 (0.088)	1.348** (0.533)

Note: Standard errors are clustered at individuals. *(**)(***) indicate significance at the 10(5)(1) % levels. Number of observations: 6,903,650. NOK 1,000 = $170 (based on average exchange rate in 2013). *Control variables included in all models;* see note to Table 4.

D. Benefits versus Costs

Our analysis indicates that to some extent the QP has accomplished its aims of helping hard-to-employ persons back to (or into) work. However, the program has also been costly. Have the benefits outweighed the costs? To answer this question properly, we would obviously need a lot more information than what comes out of our estimation exercise, both with respect to impacts beyond the fourth year after QP entry and with respect to impacts on non-pecuniary outcomes such as health, life quality and criminal behavior. While a full-blown cost-benefit analysis is well beyond the scope of this paper, we can shed some light on the costs and benefits involved by accumulating the annual effects that we do identify and comparing them to administrative costs. To help accomplish this, we first add up the extra labor earnings estimated to have been generated by the program. We then subtract the marginal cost of public funds associated with any net increases in transfers.[15] Finally, we subtract the additional administrative costs associated with implementing the program, which based on admittedly highly imperfect information, can be estimated to approximately NOK 60,000 per participant (including the marginal cost of public funds).

[15] We assume that the marginal cost of public funds is 20%. To compute the net effect of changes in taxable benefits, we also assume an average income tax on these benefits of 20%.

To perform this cost-benefit analysis, we build on the 600 bootstrap estimations. This makes it possible to assess the (large) statistical uncertainty involved, and the results are presented in Figure 5. In this exercise, we have allocated the presumed administrative costs over the entry year and the two subsequent years. For this reason, it is no surprise that the accumulated costs exceed the accumulated benefits during the first few years after entry. According to our point estimates, the total benefits do not fully balance the costs during the four-year estimation period covered in our analysis. Hence, for this particular cost-benefit assessment to come out with a positive number, we need to assume that the favorable earnings effects to some extent persist after the fourth year. In Figure 5, we have extrapolated the estimated fourth-year effects for two additional years, and the estimates then indicate positive effects from the fifth year. However, the statistical uncertainty is large, as illustrated by confidence intervals, particularly when we extrapolate out of our four-year outcome period.

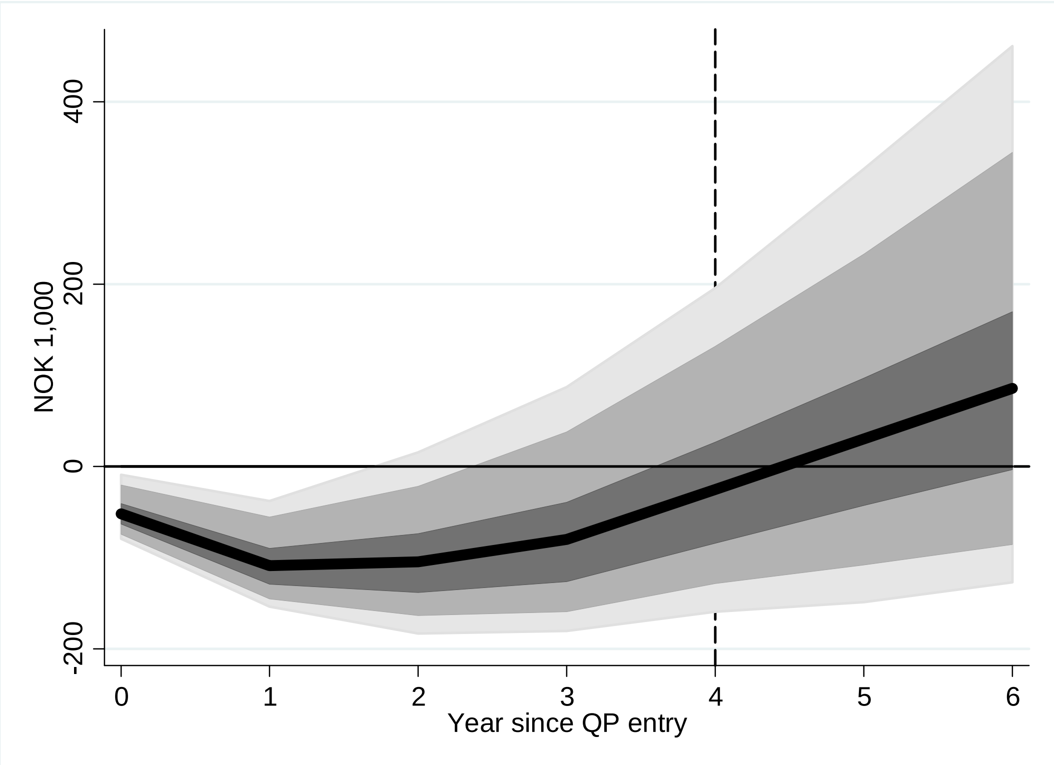

Figure 5: Estimated accumulated earnings effects minus program costs, with 50% (dark gray), 80% and 90% (light gray) confidence intervals.

Note: Impacts for the fifth and sixth year are based on the extrapolation of the effects estimated for the fourth year. Costs include NOK 60,000 in administrative expenditures and 20% of estimated impacts on public transfers (the cost of public funds). NOK 1,000 = $170 (based on average exchange rate in 2013).

In addition to the costs and benefits discussed here, there will of course also be costs and benefits associated with the particular activities that the QP participants take part in, which – as described above – range from medical rehabilitation to full-time work. Since we do not have any specific information about individual QP activities, nor about the activities that would have prevailed without QP, we are not able to provide a precise assessment of these cost-benefit components. However, statistics reported by Statistics Norway, based on information collected from the municipalities, indicate that work is the dominant activity. For example, among those who entered the QP in 2009, as much as 78% participated in some form of employment in 2010 as part of their individual QP plan.[16] This presumably implies that value is generated, though probably at some costs associated with individual support and workplace adaptations.

Additional favorable effects of QP may come from peer influences on persons outside our analysis population or at later points in time. Recent empirical studies have indicated that welfare dependency is contagious within social networks; see, e.g. Rege et al. (2012) and Markussen and Røed (forthcoming). Hence, if the QP succeeds in moving at least some persons from welfare to work, we may expect to see subsequent knock-on effects among their peers.

An important point to bear in mind is that a main aim of the QP was to reduce poverty. Our estimation results indicate that the overall level of take-home earnings was raised for all years after entry, with a possible exception for the third year, in which the increased transfer level has been tapered off while the favorable earnings effects are still small. Because of this, poverty problems appear to have been alleviated by the program.

A final and potentially important positive effect of the QP could be a reduction in criminal behavior. Economic theory suggests that any improvement in legitimate labor market opportunities reduces the incentives for committing crimes, and several empirical studies have confirmed that this mechanism is important in practice; see, e.g. Machin and Meghir (2004), Lin (2008) and Altindag (2012). Since many QP participants have criminal records prior to entry, it is probable that such effects are particularly relevant in the context of this program.

V. Conclusion

The main conclusion of our analysis is that the combination of activation requirements and economic security appear to have had the intended effect of helping a hard-to-employ group of persons back to (or into) work in Norway. Two years after entry into the Qualification Pro-

[16] These numbers are available from «Statistikkbanken» (https://www.ssb.no/statistikkbanken).

gram (QP), we identify a positive effect on the probability of having obtained at least some paid employment, and the effect increases in subsequent years. Four years after entry, our estimates indicate that the employment probability is raised by as much as 18 percentage points, ceteris paribus. Effect estimates of this order appear robust across different samples and subgroups. Still, the statistical uncertainty associated with each single estimate is large (standard errors of about seven-nine percentage points), so the point estimates should be interpreted with care.

Although most of the additional jobs are small, at least to start with, they may represent important stepping-stones into regular employment for a group of persons who in the absence of the program most likely would have had labor force participation rates close to zero. In combination with the structured and active everyday life implied by the program itself, this may also have built up a work habit among the participants with possible knock-on effects on the participants' peers. Whether the gains identified in this paper are sufficient to compensate for the costs depends on the extent to which the moderately increased labor earnings can be expected to last or perhaps even become a first step toward full employment, which is too early to say based on our data.

Given the comprehensiveness of the program evaluated in this paper, one may wonder which particular program features are most vital for its apparent success, i.e. what are the "active ingredients" behind our estimation results? Is it the poverty alleviation, the activation strategy or the combination of the two? While these ingredients cannot be empirically disentangled on the basis of our data – since all participants were subjected to the complete package – it is worth emphasizing that the basic idea of the program was to *combine* poverty alleviation and activation. A number of empirical studies on unemployed and temporary disabled individuals (referred to in the introduction to this paper) have shown that while income support *alone* reduces poverty problems, at least in the short run, it does so at the cost of also reducing the speed by which claimants find (or return to) employment. Thus, the finding in this paper that an increase in the level of income support was actually followed by a significant increase in employment propensity suggests that the activation part must have been of critical importance. This interpretation is corroborated by the finding that QP entry actually triggered a much lower increase in the income support level than suggested by a comparison of the QP benefit level with both the participants' own income history and with the national social assistance guidelines. Yet, although the realized income support did not increase very much, it is probable that the income security and predictability offered by the program may have played an important role, in that it gave the participants room to focus on longer-term issues.

Our analysis does not provide answers to questions such as: Was the high caseworker intensity really necessary? Could the same effects have been obtained at a lower cost? Would the results have been even better with a bit lower (or higher) income support? To empirically address such issues, we would need random-assignment-like variations not only in program participation, but also in program design. However, what we can conclude is that there exist combinations of quite generous income support and activation strategies that actually have the capacity to help even some of the most hard-to-employ persons into gainful employment.

Appendix

Table A1: First stage estimates (Eq. 3), based on sample with low earnings threshold; effects of excluded instruments on entry probabilities in year *t+r-p* for *p=0,1,2,3,4* (standard errors in parentheses)

	$QP_{i,t+r-0}$	$QP_{i,t+r-1}$	$QP_{i,t+r-2}$	$QP_{i,t+r-3}$	$QP_{i,t+r-4}$
$z^*_{i,t+r-0}$	0.740***	-0.015	-0.002	-0.004	0.002
	(0.025)	(0.012)	(0.006)	(0.003)	(0.001)
$z^*_{i,t+r-1}$	0.011	0.740***	-0.014	0.002	-0.006**
	(0.030)	(0.026)	(0.013)	(0.007)	(0.003)
$z^*_{i,t+r-2}$	-0.056**	0.009	0.741***	-0.013	0.010*
	(0.027)	(0.036)	(0.028)	(0.014)	(0.006)
$z^*_{i,t+r-3}$	0.020	-0.074**	0.014	0.752***	-0.023**
	(0.023)	(0.038)	(0.040)	(0.030)	(0.011)
$z^*_{i,t+r-4}$	-0.031*	0.023	-0.075	0.037	0.796***
	(0.018)	(0.038)	(0.046)	(0.047)	(0.035)
F-statistic excluded instruments (partial)	201.9	201.8	193.6	185.5	124.2
F-statistics excluded instruments (conditional)	989.2	1017.9	844.7	940.8	649.8

Note: Standard errors are clustered at individuals. *(**)(***) indicate significance at the 10(5)(1) % levels. Number of observations: 7,036,980.
Control variables: Municipality×base-year×years-since-base-year fixed effects (18,280 dummy variables), age (44 dummy variables), gender, education (eight dummy variables), immigrant status (three dummy variables), labor earnings (in the base-year and as an average over the three years leading up to the base-year), taxable benefits in the base-year, non-taxable benefits in the base-year, number of months with social assistance in the base-year, number of months with UI benefits in the base-year, number of months with temporary disability benefits in the base-year, estimated QP participation propensities interacted with base-year×outcome-year (760 variables) and estimated QP participation propensities interacted with reform-year in the municipality (57 variables).

Since the instruments used in the first stage (Equation (3)) partly mirror the availability of the program and – by construction – represent predicted entry probabilities given availability, there is in this case no doubt that they have causal effects on entry into the program. Table A1 presents the first stage estimation results. The coefficients along the diagonal, which correspond to the respective entries are all large (around 0.75) and highly significant (t-values above 22). However, note that the diagonal elements of Table A1 are not equal to one, nor are the off-diagonal elements always equal to zero. These patterns primarily reflect that participation sometimes occur in municipalities/years that were not considered to be at risk when the instrument was constructed. Moreover, while the instruments are constructed as a non-linear function of individual characteristics, they enter linearly in the 2SLS model.

We present two different F-statistics for the power of the instruments. The *partial F-statistic* gives the conventional test for the joint impact of the excluded instruments separately for each of the five endogenous variables. Given that we have five instruments, these F-statistics have been adjusted for five degrees of freedom. Since we also have five endogenous variables, there is no superfluous instrument in our model. And the partial F-statistics are unable to detect cases in which interdependencies between the instruments imply that our model is under-identified. We therefore also provide the *conditional F-statistics* suggested by Sanderson and Windmeijer (forthcoming) to test for weak instruments in a multiple endogenous variable setting. These F-statistics are conditional on the other endogenous variables, and take into account that there is only one instrument per endogenous variable. In our case, all the F-statistics are well above conventional threshold levels for weak instruments.

References

Aakvik, Arild, James J. Heckman, and Edward J. Vytlacil. 2005. «Estimating Treatment Effects for Discrete Outcomes when Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs» *Journal of Econometrics* 125: 15-51.

Altindag, Duha T. 2012. «Crime and unemployment: Evidence from Europe» *International Review of Law and Economics* 32: 145– 157.

Angrist, Joshua, Guido W. Imbens, and Donald B. Rubin. 1996. «Identification of Causal Effects Using Instrumental Variables» *Journal of the American Statistical Association* 91: 444-455.

Ashenfelter, Orley. 1978. «Estimating the Effect of Training Programs on Earnings» *Review of Economics and Statistics* 60: 47-50.

Calvó-Armengol, Antoni and Matthew O. Jackson. 2004. «The Effects of Social Networks on Employment and Inequality» *American Economic Review* 94 (3): 426-454.

Card, David, Jochen Kluve, and Andrea Weber. 2010. «Active Labour Market Policy Evaluations: A Meta-analysis» *The Economic Journal* 120: F452–F477.

Christensen, Tom, Anne Lise Fimreite, and Per Lægreid. 2014. «Joined-Up Government for

Welfare Administration Reform in Norway» *Public Organization Review* 14: 439-456.

Dasgupta, Partha. 1997. «Nutritional Status, the Capacity for Work, and Poverty Traps» *Journal of Econometrics* 77: 5-37.

Dasgupta, Partha and Debraj Ray. 1986. «Inequality as a Determinant of Malnutrition and Unemployment: Theory» *Economic Journal* 96: 1011-1034.

Dasgupta, Partha and Debraj Ray. 1987. «Inequality as a Determinant of Malnutrition and Unemployment: Policy» *Economic Journal* 97: 177-188.

Fevang, Elisabeth, Inés Hardoy, and Knut Røed. 2013. «Getting Disabled Workers Back to Work: How Important Are Economic Incentives?» IZA Discussion Paper No. 7137.

Frölich, Markus, Almas Hesmati, and Michael Lechner. 2004. «A Microeconomic Evaluation of Rehabilitation of Long-term Sickness in Sweden» *Journal of Applied Econometrics* 19: 375-396.

Kluve, Jochen, David Card, Michael Fertig, Marek Góra, Lena Jacobi, Peter Jensen, Reelika Leetmaa, Leonhard Nima, Eleonora Patacchini, Sandra Schaffner, Christoph M. Schmidt, Bas van der Klaauw, and Andrea Weber. 2007. «*Active Labor Market Policies in Europe - Performance and Perspectives*» Springer-Verlag Berlin and Heidelberg GmbH & Co.

Krueger, Alan B. and Bruce D. Meyer. 2002. «Labor Supply Effects of Social Insurance» In A. J. Auerbach and M. Feldstein, M. (Eds.), Handbook of Public Economics, Vol. 4. Elsevier Science, North-Holland, 2002; 2327-92.

Lin, Ming-Jen. 2008. «Does unemployment increase crime? Evidence from U.S. data 1974–2000» *Journal of Human Resources* 43 (2): 413–436.

Machin, Stephen and Costas Meghir. 2004. «Crime and economic incentives» *The Journal of Human Resources* 39 (4): 958–979.

Markussen, Simen and Knut Røed. Forthcoming. «Social Insurance Networks» *Journal of Human Resources.*

Moffitt, Robert. 2007. «Welfare Reform: The US Experience» *Swedish Economic Policy Review* 14 (2): 11-48.

Mullainathan, Sendhil and Eldar Shafir. 2013. «*Scarcity: Why Having Too Little Means so Much*» Time Books, Henry Holt and Company, New York.

Rege, Mari, Kjetil Telle, and Mark Votruba. 2012. «Social Interaction Effects in Disability Pension Participation – Evidence from Plant Downsizing. *Scandinavian Journal of*

Economics Vol. 114, No. 4, 1208–1239.

Røed, Knut. 2012. «Active Social Insurance» *IZA Journal of Labor Policy* 1:8 (doi:10.1186/2193-9004-1-8).

Røed, Knut and Oddbjørn Raaum. 2006. «Do Labour Market Programmes Speed up the Return to Work?» *Oxford Bulletin of Economics & Statistics* 68 (5): 541-68.

Røed, Knut and Lars Westlie. 2012. «Unemployment Insurance in Welfare States: The Impacts of Soft Duration Constraints» *Journal of the European Economic Association* 10 (3): 518-554.

Røed, Knut and Tao Zhang. 2003. «Does Unemployment Compensation Affect Unemployment Duration?» *Economic Journal* 113: 190-206.

Røed, Knut and Tao Zhang. 2005. «Unemployment duration and economic incentives - A quasi random-assignment approach» *European Economic Review* 49: 1799-1825.

Sanderson, Eleanor and Frank Windmeijer. Forthcoming. «A Weak Instrument F-test in Linear IV Models with Multiple Endogenous Variables» *Journal of Econometrics.*

Schafft, Angelika and Øystein Spjelkavik. 2011. «Evaluering av Kvalifiseringsprogrammet» (Evaluation of the Qualification Program). AFI-rapport 4/2012.

Shah, Anuj K., Sendhil Mullainathan, and Eldar Sharif. 2012. «Some Consequences of Having Too Little» *Science* 338: 682-685.

Walker, Robert, Grace B. Kyomuhendo, Elaine Chase, Sohail Choudhry, Erika K. Gubrium, Jo Yongmie Nicola, Ivar Lødemel, Leemamol Mathew, Amon Mwiine, Sony Pellissery, and Yan Ming. 2013. «Poverty in Global Perspective: Is Shame a Common Denominator?» *Journal of Social Policy* 42 (2): 215-233.