Empirical analysis on the economic effects of disability benefits

Alonso Quijano

April 2022

Disability benefits are government programs that exist with the goal of protecting individuals against income loss due to disability. They have for long time, however, concerned economists and policymakers as they may highly encourage nonemployment, causing a large imbalance between the private benefits for beneficiaries and society cost to taxpayers.

Just like many other social programs, it is hard if not impossible to implement a randomized experiment to estimate the effects of disability benefits, denying some people access to benefits while still granting access to others. Simply comparing beneficiaries with nonbeneficiaries is less of a viable option. Beneficiaries may be better informed, have more catastrophic illnesses, or be of an older age, where disability is more common and one may have enough earning history to be eligible for benefits. With enough data, a researcher would be able to control for many of the observable confounders, but would hardly be able to account of those unobservable variables that would influence one's decision to stop working and opt for a disability pension, such as one's own taste for work.

The goal of this essay is to analyze and comment on the empirical evidence on the economic effects of disability benefits. I will look at two famous papers on the topic, Gruber (2000) and Autor et al. (2019), both witch use different identification strategies, difference in differences and instrumental variables, respectively. I will end by summarizing the results of a study of my own that attempts to measure the effects of disability benefits in Ecuador, my homeland, and commenting on its limitations and potential biases.

1 Gruber's (2000) DID study on the Canadian pension system

Canada has two different disability insurance programs, one for the region of Quebec and another for the rest of the country. While both programs operate similarly, the Quebec Pension Plan (QPP) had always been more generous than the Canada Pension Plan (CPP). In 1987, however, the Canadian government increased the benefit levels of the CPP by 36 percent to equalize the generosity of the QPP. Gruber (2000) studied this policy change to measure the elasticity of labor force nonparticipation with respect to disability insurance benefits, which he estimated to be 0.36.

Gruber (2000) applies a logistic regression model to compare nonemployment before and after the policy using Quebec as the control region and the rest of Canada as the treatment region. The difference in differences (DID) model is as follows:

$$NE_i = f\left(\alpha + \beta_1\text{CPP} + \beta_2\text{AFTER} + \beta_3\text{CPP} \times \text{AFTER} + \beta_4 X_i + \epsilon_i\right)$$

Where NE represents a dummy for whether person i is nonemployed, CPP is a dummy for whether person i lives in a CPP province, and AFTER is a dummy for whether the year is after the policy change. X_i is a set of covariates (age, marital status, education, and number of children).

β_3 is the parameter of interest, which represents the effects of the policy change on nonemployment. From the logistic coefficient, Gruber (2000) calculated the implied probability effect as 2.3 percent, indicating that the 36 percent rise in benefits led to a rise in nonemployment of 11.5 percent (or an arc elasticity of 0.36).

In order for the DID model to provide an unbiased estimate of the policy effect, one key assumption must be met, the parallel trend assumption. This assumption requires that in absence of the policy change, the difference in nonemployment between both regions should remain constant over time. To prove this assumption, Gruber (2000) conducts a falsification test. This test consists of running the same DID model but on the years before the policy change. If the coefficient is significant, it suggests there was a preexisting trend in nonemployment and, thus, the parallel trend assumption would be violated. The results from this falsification test show a small and insignificant coefficient, implying that there was no preexisting trend in nonemployment prior to the policy change and proving the parallel trend assumption.

Like any other study, this one has its own limitations. For example, in the model he omits earnings history, which is unavailable in the cross sectional data he uses. Earnings history is an important variable as it determines whether someone is eligible for disability insurance. Not only that, but for those who are eligible it also determines the replacement rate (the percentage of one's employment income that is replaced by the disability allowance). Individuals with a higher replacement rate would probably choose not to work after becoming disabled, while those with a lower replacement rate would probably opt to keep working. Adding this variable to the model would most likely change the estimate. This omitted variable bias effect may be smaller for men, who may have enough years of contribution, but larger for women.

Overall, Gruber's (2000) study does a good job designing an identification strategy and proving its assumptions. Additional to the DID model, he estimates "synthetic earning histories" for groups of workers in order to calculate their potential replacement rate. He then adds this "synthetic replacement rate" to a parametrized model. The results from this model are similar to those obtained in the DID model.

2 Autor et al.'s (2019) IV model on the Norwegian disability insurance program

The Norwegian government provides disability allowance to people unable to work due to physical or mental impairment. Of those eligible to apply, approximately a quarter are rejected. Denied applicants are usually those claiming difficult to verify impairments, especially back pain. About 25 percent of denied applicants appeal in court, 15 percent of those being ultimately allowed benefits.

Autor et al. (2019) apply an instrumental variable (IV) model to measure the household impacts and fiscal costs of disability benefits. They use judges' leniency as an instrumental variable for disability allowance. They find that receiving disability allowance on appeal reduces annual labor earnings by approximately 6,800 USD in the first year after appeal, approximately 40 percent of the annual transfer benefit received. The IV model is as follows:

$$A_i = \gamma Z_{j(i)} + X_i'\delta + \epsilon_i,$$
$$Y_{it} = \beta_t A_i + X_i'\theta_t + \eta_{it}$$

The first equation above represents the first stage model. A_i is a dummy for whether the person i is allowed insurance allowance at the appeal. $Z_{(j)i}$ is the leniency measure for judge j to which appellant i assigned. X_i represents a set of relevant control variables. The second equation corresponds to the second stage equation, where Y_i represents labor earnings three years after the decision.

There are several assumptions an IV model needs to meet in order to provide a reliable estimate of a causal effect. First, the instrument (in this case the judges' leniency) should cause variation in the treatment variable (in this case being allowed or not insurance allowance). Second, judges' leniency should be correlated with the outcome (in this case labor earnings) only through the decision of allowing or not insurance allowance. In other words, the instrumental variable should not directly affect the outcome. This is known as the exclusion restriction. Third, the instrument should be assigned randomly or quasi randomly.

The exclusion restriction is generally hard to prove, and it usually relies on the researcher's own understanding of and conjectures about the instrument validity. For example, a violation of the exclusion restriction could happen if the judges were more lenient towards those with specific types of disability or sociodemographic background, or those who they consider would "make better use" of the disability insurance. Autor et al. (2019) argue that in Norway, appeals are presented in writing and judges do not have any personal contact with appellants. They further prove this assumption by running a regression of appellants predetermined characteristics (e.g. age, sex, education, type of disability, assets, and earnings) on judges' leniency and whether the case was allowed or not. These results do not only support the exclusion restriction but also the random assignment of the instrument as they show that cases were assigned to judges in a random fashion. Actually, as they then clarify, cases are assigned on a rotating basis depending on the date they are received and the alphabetical ordering of judges' last name.

In their study, Autor et al. (2019) manage to find a reliable and valid instrument to estimate the economic effects of disability insurance allowance. However, their estimate only represents the treatment effect for a subset of the sample (in this case the appellants) if and only if they were assigned the treatment, otherwise known as the Local Average Treatment Effect (LATE). As mentioned, those who appeal in court are usually those with impairments (e.g., back pain or mental disorders) that are particularly difficult to verify.

In their study, Autor et al. (2019) manage to find a reliable and valid instrument to estimate the economic effects of disability insurance allowance. However, their estimate only represents the treatment effect for a subset of the sample (in this case the appellants) if and only if they were assigned the treatment, otherwise known as the Local Average Treatment Effect (LATE). As mentioned, those who appeal in court are usually those with impairments (e.g., back pain or mental disorders) that are particularly difficult to verify. This means that the estimated treatment effects cannot be extrapolated to other populations, restricting the external validity of the results.

In addition to their primary findings, Autor et al. (2019) run separate models to test the effects of disability insurance allowance on married and unmarried appellants. However, one important analysis they do not consider is the heterogeneous treatment effects between men and women. Effects on labor earnings and household consumption may differ between men and women, as they could have different tastes for work, consumption preferences, and bargaining power. In development economics literature, measuring the treatment effects between distinct allowance recipients is key to understanding not only the welfare implications of a policy but also its effects on well-being. For example, Duflo (2003) found that pensions received by elderly female recipients in South Africa had a large impact on the nutrition of their young granddaughters. While this is not a development economics study, analyzing the gender effects of disability insurance allowances could still provide insights on the elasticity of labor earnings and household consumption. In the study in describe in the next section, I further discuss the gender effects of disability benefits.

3 The labor and health effect of disability benefits in Ecuador

Since 2003, the government of Ecuador issues disability IDs to people with disabilities. Disability ID holders can enjoy several benefits such as tax exemption on the importation of goods, income tax benefits and VAT reimbursements, and discounted fares for many public services. In 2009, the government of Ecuador launched a social program called Misión Manuela Espejo (MME) as part of a social research for registering and studying people with disabilities. Visited subjects were given disability IDs and caregivers of those with severe physical and mental disabilities became eligible to receive a $240 cash transfer.

The goal of this study is to explore the effects of disability benefits on health status and demand for healthcare services in Ecuador. We use instrumental variables to estimate the causal effect of having a disability ID—an ID card issued by the government that provides its holder the official status of disabled, discounted prices on public services, tax benefits, and the eligibility to apply for cash transfers. The instruments are the visits made during the Misión Manuela Espejo program. Findings suggest that male ID holders are less likely to be sick and more likely to report feeling healthy. Female ID holders, in contrast, are more likely to receive preventive care. We also estimate the effects on labor supply and, similar to previous research, find that disability ID holders are more likely to be non-employed.

The IV linear probability model is as follows:

$$ID_i = \alpha Z_i + X_i'\gamma + \epsilon_i$$
$$Y_i = \beta ID_i + X_i'\delta + \eta_i$$

Where ID_i is a dummy for whether person i has a disability ID. Z_i is a dummy for whether person i was visited by MME. X_i' is a vector of control variables (age, education, number of children, marital status, type of disability, degree of disability). Y_i represents an outcome of interest (employment status, feeling sick in the past 30 days, having received preventive care in the past 30 days, and self-assessment of health).

To test the instrument's relevance, we use the first stage equation. As expected, being visited by MME largely increases the probability of having an ID holder. The instrument coefficient is large and significant (F-statistic 51.023 for male and 75.174 for female).

Testing the instrument validity is not as straightforward. There are several circumstances that could threaten the validity of the instrument. First, MME visits could cause having a disability ID, but the opposite could also be possible as MME benefits were not restricted only to those who did not have a disability ID previous to the program. We cannot observe in the data whether a person had a disability ID prior to the MME visit. However, we know that the motivation for the MME program was the very small registry of people with disabilities in Ecuador. Thus, we believe that those visited by the MME who already had a disability ID represent a small and insignificant part of the sample. Another threat to the validity of the instrument is the possibility that during the MME, beneficiaries received more than just an ID. For example, in some cases, during the MME visits poor household received assistive devices likes wheelchairs and house items like mattresses. This could inflate the effect of having a disability ID on health status.

To assess the random assignment of the instrument, we examine the relationship between the MME visits individuals' home infrastructure. Because of the nature of the cross-sectional data, we cannot estimate the probability of

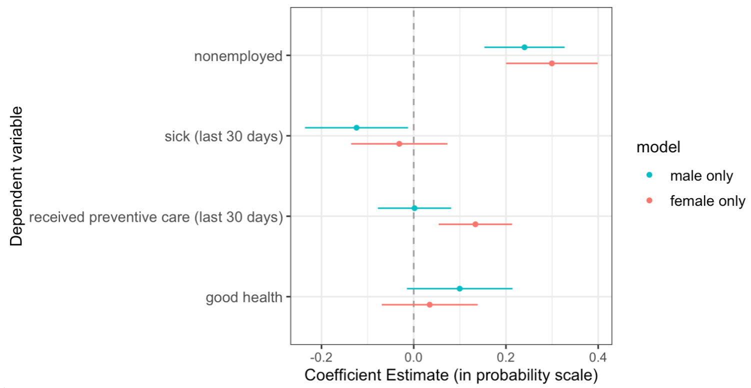

Figure 1: Effects of having a disability ID on nonemployment and health status

having been visited by the MME as a function of the subject's socioeconomic characteristics at the time of the visit. To address this issue, we looked at the subject's home infrastructure, which is highly predictive of economic status and less likely to change over time. We argue that, while the MME program was inherently targeted to poor people, at the neighborhood level the subject's characteristics are not correlated with the instrument assignment. We applied a linear probability model to test whether subjects' home infrastructure increase the probability of being visited by the MME. When applying no district fixed effects, home infrastructure is predictive of visits but coefficients are very small (less than five percent). When adding district fixed effects, coefficients become insignificant.

One important analysis in this study is the treatment comparison between men and women, which are not included in the studies by Gruber (2000) and Autor et al. (2019). We estimated a large effect of disability benefits on nonemployment for both men and women. We observe that health outcome, however, are different between groups. Men are less likely to be sick and feel healthier, but we find no effect on women. This could be because even after quitting paid labor, women may still choose to do domestic labor, which may hinder their recovery. Surprisingly, we find that women are more likely to receive preventive care, but we find no effect on men. Economic literature on disability has traditionally focused on the causal relationship between disability benefits and the labor supply of the disabled. There is less evidence, however, on their effects on health. These results are interesting as they provide insights on the policy benefits of disability programs on health, which may be crucial for coping with the long-term challenges of living with an impairment.

References

Autor, D., Kostøl, A., Mogstad, M., & Setzler, B. (2019, July).

Disability benefits, consumption insurance, and household labor supply. *American Economic Review*, *109*(7), 2613-54. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/aer.20151231` doi: 10.1257/aer.20151231

Duflo, E. (2003). Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in south africa. *The World Bank Economic Review*, *17*(1), 1–25.

Gruber, J. (2000). Disability insurance benefits and labor supply. *Journal of political economy*, *108*(6), 1162–1183.