ZEW
Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Universität
Konstanz

Rechts-, Wirtschafts- und
Verwaltungswissenschaftliche
Sektion
Fachbereich
Wirtschaftswissenschaften

Diskussionspapiere der DFG-
Forschergruppe (Nr.: 3468269275):

Heterogene Arbeit: Positive und Normative
Aspekte der Qualifikationsstruktur der Arbeit

Stefan Hupfeld

Rich and Healthy – better than Poor and Sick?

An Empirical Analysis of Income, Health, and the Duration of the Pension Benefit Spell

May 2007

Diskussionspapier Nr. 07/07
http://www.wiwi.uni-konstanz.de/forschergruppewiwi/

Diskussionspapier der Forschergruppe (Nr.: 3468269275) "Heterogene Arbeit: Positive und Normative Aspekte der Qualifikationsstruktur der Arbeit"

Nr. 07/07, May 2007

Rich and Healthy – better than Poor and Sick? An Empirical Analysis of Income, Health, and the Duration of the Pension Benefit Spell

Stefan Hupfeld

Universität Konstanz
Fachbereich Wirtschaftswissenschaften
Fach D136
78457 Konstanz
Germany

mail: Stefan.Hupfeld@uni-konstanz.de
phone: +49-7531-88-2534
fax +49-7531-88-4135

Zusammenfassung:

Abstract: We analyze the relationship between duration of the pension benefit spell and pension benefit claims from the German public pension system, with a special emphasis on differential results with respect to health. This relationship is crucial and causal for a potential structural pattern of redistribution between different income and health groups, induced by the public pension system. Evidence for such redistribution from poor to rich is present for most of the specifications in our analysis. The specification we believe to be correct is partially-linear, does therefore not impose any parametrical restrictions between duration and benefit claims, and allows for potential endogeneity. The relationship we extract is remarkably close to positively linear. Additionally we find that the income gradient is steeper for pensioners in bad health, meaning that redistribution from the least able to the most able individuals is more pronounced the worse the health status is.

JEL Klassifikation : H55, J26, (C14)
Schlüsselwörter : Public Pension System, Duration of the Benefit Spell, Income Gradient
Download/Reference : http://www.wiwi.uni-konstanz.de/forschergruppewiwi/

Rich and Healthy – better than Poor and Sick? An Empirical Analysis of Income, Health, and the Duration of the Pension Benefit Spell

Stefan Hupfeld*

May 25, 2007

Abstract

We analyze the relationship between duration of the pension benefit spell and pension benefit claims from the German public pension system, with a special emphasis on differential results with respect to health. This relationship is crucial and causal for a potential structural pattern of redistribution between different income and health groups, induced by the public pension system. Evidence for such redistribution from poor to rich is present for most of the specifications in our analysis. The specification we believe to be correct is partially-linear, does therefore not impose any parametrical restrictions between duration and benefit claims, and allows for potential endogeneity. The relationship we extract is remarkably close to positively linear. Additionally we find that the income gradient is steeper for pensioners in bad health, meaning that redistribution from the least able to the most able individuals is more pronounced the worse the health status is.

Keywords: Public Pension System, Duration of the Benefit Spell, Income Gradient

JEL-Classification: H55 J26 (C14)

*Department of Economics, University of Konstanz, Box D136, 78457 Konstanz, Germany, Stefan.Hupfeld@uni-konstanz.de. I gratefully acknowledge support by the German Research Foundation (DFG) under grant no. 454, "Heterogeneous Labor: Positive and Normative Aspects of the Skill Structure of Labor". A previous version of this paper has been presented at several conferences under the title *Disentangling the Relation between Realized Duration of the Pension Benefit Spell and Benefit Claims in Germany.* I thank Friedrich Breyer, Normann Lorenz, Joaquim Oliveira Martins, the participants of the FNA Doctoral Colloquium, and the RGS Doctoral Conference for helpful comments and discussions. Obviously, all remaining errors are my own.

1 Introduction

The length of the period between retirement and death (the duration under the benefit spell) is the crucial determinant for the rate of return from the German public pension system. Contributions are conditioned on income, whereas monthly benefits are paid according to the amount of benefit claims and the current value of these claims. The benefit claims are measured in points, where one point corresponds to the claims earned by the contributions based on one year of the average income. Since this point value is equal for all pensioners in each year, the individual monthly pension is proportional to the total amount of individual contributions. Yet, people tend to receive their pension for a different period, such that the total amount of benefits is not necessarily proportional to the contributions. Differences occur for two reasons, namely due to differences in the retirement age and differences in life expectancy.

The key question is, whether differences in duration (and hence, in the internal rate of return from the pension system) are systematic. If they are not, the public pension simply fulfils its task as an insurance against longevity, and individual deviation from the average rate of return is random. If, however, the duration under the benefit spell varies systematically with income, e. g., then a potential for unintentional redistribution among different income groups arises.

Especially life expectancy is usually perceived to be increasing in socio-demographic factors such as income. Recent empirical findings at least partially corroborate this conjecture, among them von Gaudecker and Scholz (2006).[1] Other things – especially retirement age – equal, this would be an indication for redistribution from the bottom to the top of the income distribution. Additionally, retirement age might be a function of income as well. Evidence in this case is controversial, see Berkel and Börsch-Supan (2004) versus Schils (2005) with opposing results with respect to the impact

[1] However, in their seminal paper, Adams, Hurd, MacFadden, Merrill, and Ribeiro (2004) find no causal link of socio-economic status on mortality. Instrumenting income or wealth in order to account for endogeneity, there exist two opposing results with respect to the impact of income on health, which is closely related to mortality. Meer, Miller, and Rosen (2003) find no or only a weak impact of income on health, whereas Lindahl (2005) does find a significantly positive impact. Even in the study by von Gaudecker and Scholz (2006), it is debateable whether the relationship between income and mortality is really monotonously increasing over the whole support of the income distribution.

of income or wealth on retirement age. The impact of socio-economic factors on the aggregate of both, namely on the duration, has been examined by Lauterbach, Lüngen, Stollenwerk, Gerber, and Klever-Deichert (2006). They find that duration is indeed higher for individuals with higher income. Still, the death cohort they observe is rather small, and benefit payments before the age of 65 are neglected. In contrast to this analysis, we propose to include any kind of pension benefits, especially those for early retirees and disability pensions. Another empirical contribution with regard to heterogeneous social security outcomes by different subgroups of the population in the United States has been made by Liu and Rettenmaier (2003), who find that the shorter life expectancy of blacks yields a lower return from social security even despite progressive elements in the U.S. social security, which are not in place in the German retirement benefit formula.[2]

Theoretical literature on duration as a function of income being the driving force behind heterogenous implicit taxation in pension systems is mainly based on models of optimal taxation. Diamond (2003), pp. 87, derives an optimal pension system (or age-dependent tax system) that takes exogenous differences in life expectancy into account, and concludes that different implicit tax rates (or, in other words, different rates of return[3]) that are actually second best optimal. Breyer and Kifmann (2002) analyze the effect of the adjustment of discounts for early retirees on the implicit tax rate – however, heterogeneous life expectancy does not matter in their analysis. Its separate effect on the implicit tax of a pension system is analyzed in Breyer and Kifmann (2004). A factual and an optimal life-cycle profile of the implicit tax imposed by a social security system is derived by Fenge, Uebelmesser, and Werding (2006), who find that actual as well as optimal implicit taxation differs even despite absent differences in duration under the benefit spell. An exact measure of the size of the implicit tax exerted on prolonged work is proposed by Butrica, Johnson, Smith, and Steuerle (2006), who find that the implicit tax in the United States is indeed smaller for individuals with a life expectancy above average, yet, their measure is

[2]Yet, other groups with – on average – slightly lower life expectancy like the low educated offset their disadvantage and can directly benefit from the progressive elements in the benefit formula.

[3]Both, internal rate of return and implicit tax measure the same phenomenon, the implicit tax being the difference between the rate of return on the capital market and the internal rate of return of the pension system.

not directly transferable to Germany, as the general income tax and health insurance issues are also taken into account.

Standard theory on the effect of income on longevity is mainly contained in Grossman (1999), who summarizes his own earlier work on health capital as well as contributions and extensions by subsequent authors. Though life expectancy is in general not a direct choice variable in these models, death is usually assumed to occur once health capital depreciates below a certain threshold and is therefore (positively) affected by income or initial wealth. There exists a comparably vast strand of theoretical literature with regard of the retirement decision, effectively beginning with Feldstein (1974), Diamond and Mirrlees (1978), and Sheshinski (1978), who estimate and analyze the effect of a pension system on savings and, even more important, the retirement decision. A recent theoretical contribution by Bloom, Canning, and Moore (2005) indicates that initial wages are negatively related to retirement age, but the growth rate of income has a positive impact. The first is explained by the income effect of higher wages on labor supply in general, which is also present for the retirement decision. Although longevity is a main driving variable their theoretical model, an empirical test of this model in Bloom, Canning, Moore, and Song (2006) shows that subjective probabilities of survival have no impact on the retirement decision in the United States.

In this paper, we find that the answer to our key conjecture of income-induced differences in duration is sensitive to the sub-group under consideration. Although we propose to restrict the relevant group to individuals with a certain minimum number of years of contribution, results for other sub-groups are also reported. Applying non-parametric and semi-parametric methods, we find evidence for duration to be increasing in benefit claims. The shape and level of the analyzed relationship depends additionally on the health status of the individual, which is measured by the time spent in ill-health or rehabilitation or by the existence of a private health insurance. In general, people in ill-health benefit from a longer benefit spell, but the positive impact of income is stronger for people in worse health. The specification we believe to be correct accounts for endogeneity and shows a positive impact of benefit claims on duration, which is remarkably close to positively linear. So, on average there is evidence for redistribution from poor to rich in the pension system. A comparison of these results to standard LS and 2SLS regressions indicates that the income gradient is always significant, although its predictive power is low.

The remaining paper is organized as follows. In section 2, the data set used in this analysis is introduced. Section 3 gives a short overview of the descriptive statistics, whereas section 4 introduces the econometric method, which is mainly based on locally linear regression. Section 5 discusses the results, and in section 6 we conclude.

2 The Variables

The variables used in this analysis stem from a data set with pension discontinuations from 1993 to 2003, published by the Federation of German Pension Insurance Institutes (VDR, now: Deutsche Rentenversicherung Bund). It contains a 10% sample of all discontinued public pensions due to the death of the beneficiary, which amounts to roughly one million observations. However, each observation corresponds to a pension, and not to an individual retiree, who can benefit from more than one pension. Correcting this, we are left with 98,000 observations. The most important variables are the sum of pension benefit claims, the length of the work life, and the duration under the benefit spell. All used variables are described in detail below:

- PSEGPT: The sum of pension benefit claims, measured in points. One point corresponds to one year of contributions based on the average income of those who contribute to the social insurance schemes.

- BYVL: The number of years in which own contributions have been paid.

- Duration 1: Constructed by subtracting the year of entry into the first pension from the year of death. The first pension may already be the old age pension, but it also encompasses all kinds of disability pensions. For redistributional issues, the total length of the benefit spell (and therefore Duration 1) is the relevant measure.

- Duration 2: Constructed by subtracting the year of entry into the actual pension from the year of death. The actual pension is the pension that was paid at the time the beneficiary died. If he was older than 65, it corresponds to the old age pension, as all pensions are transformed until this age. Excluding disability pensions paid due to a reduction in earnings capacity[4], all other pensions are already transformed into old-age pensions at the age of 61.

- GEBC: Gender, though the major part of this paper is only concerned with male pensioners.
- BLAND: The place of residence, aggregated to West Germany, East Germany and abroad for this analysis.
- AJAZ: Time spent in unemployment, in months.
- AT: Identification of private or public health insurance. The premium for a private health insurance is based on risk, whereas the premium for public health insurance is based on income. So, individuals who are eligible for private health insurance will choose it if they are a good risk and therefore benefit from a low premium. Only high-income individuals or their spouses/children can choose to leave the public health insurance.
- AUAZ: Time spent in ill health (as long as this time is relevant for the pension system), in months. This also covers time spent in rehabilitation treatments.

All other variables which are contained in the data set have to be taken with care. They are only reliable when they have been used for the calculation of benefits, otherwise they are unreliable or missing, which is the case for marital status, e. g. Other variables are not reliable for a similar reason: The number of children is actually important for the benefit calculation, as if the individual has interrupted his working life, some benefit claims can be earned by the upbringing of children as a counterbalance to the lost years of work. However, only one of the parents can make use of this feature, such that this variable does not measure the "physical" number of children.

It remains to be determined which observations should be included in the analysis. Though the results are reported for the whole sample as well, this is not necessarily the relevant one. Since income per year or working life (or the earnings capacity) cannot be observed directly, it has to be ensured

[4] "Reduction in earnings capacity" is (in terms of the pension system) the least severe form of a disability.

that the benefit claims are a good substitute. In the simplest case, namely when an individual has worked his whole career and contributed to the public pension system, benefit claims are a linear transformation of income.[5] This still holds even if the individual under observations had to suffer from periods of unemployment and (longer lasting) illness or if he or she raised children. The measure is then slightly diluted, but still a good proxy for income. The close relationship between total income and benefit claims however is not guaranteed once the individual has been self-employed or has worked as a public servant for some time in his career. During these times, usually no contributions are paid, as membership in the public pension system is not mandatory (or even possible) anymore.

From a methodological point of view it has to be stressed that the basis of our analysis are death cohorts, and not birth cohorts, because the set of the latter is not self-contained. If we had based our estimations on birth cohorts, we would suffer from a severe selection bias, as there are only very few birth cohorts of which all members have died and of which total duration under the benefit spell has been realized. Although there is a measure for remaining life expectancy (and therefore for expected duration) conditional on age, this measure only captures the *average* remaining life expectancy, disregarding the differences among income groups, which we are interested in.

3 Descriptive Statistics

Standard descriptive statistics for the variables of interest in each data set can be found in table 1. Histograms for the respective main independent variable are presented in figure 1. As one can see there, the distribution of our claims measure is not as skewed to the right as we would expect it from an income distribution. Those who have a very high income and contribute to the pension system are not identified with their actual income, but with a maximum sum of points, because the amount of yearly contributions has an upper limit, and so have the benefit claims. Furthermore, outliers very far to the right are truncated due to institutional restrictions, as only those

[5]Up to a certain income, beyond which contributions (and therefore claims) are capped. The maximum contributions are based (in 2006) on a monthly gross income of EUR 5250 and are adjusted on a yearly basis.

individuals are observed who work in a job where contributions to social security are mandatory.

Though the data is representative for those who take part in the pension system, the mean of the points per year of contribution is not exactly 1. This (statistically significant) slight deviation is not only due to the exclusion of some observations, namely the ones with missing values in the main variables of interest, but also with the official report.[6]

The realized duration however differs substantially from what is reported, which requires an explanation. Our measure yields (for male observations who died between 1993 and 2003) 7.38 years for total duration and 4.72 years for duration under the last benefit spell, whereas the officially reported figure for this time was larger. One difference between the official data and ours is the treatment and interpretation of missing values. We deliberately exclude especially these observations whose total amount of claims, whose amount of years of contributions, and whose duration from the entry into first and into actual pension are not available, since we want to base our estimations on exactly these measures. However, their number is not small and their impact on the level of duration is substantial. Secondly, the nature of sample imposes a small selection bias. Since we observe a death cohort, shorter life expectancies are over-sampled: Life expectancy has been rising almost linearly, but those born later (with therefore higher life expectancy) are only observed when they have died young. Both effects only capture a potential bias in the level of duration, but not in the respective income gradient, such that all relative conclusions we draw remain valid.

In line with expectations is that the average sum of points increases if the sample is restricted to a minimum of 25 years of contribution. More interesting, the average points-per-year measure is only slightly (though significantly in both data sets) affected, which indicates that the major fraction of the increase in total claims can be attributed to the sheer number of additional work years alone. The fraction of male pensioners increases once the sample is restricted, which is also in line with the observation that female careers are distinctively different from male careers in the observed time span.

[6] See Deutsche Rentenversicherung Bund (2005), p. 107.

4 Econometric Method

4.1 Locally Linear Regression

In this analysis, no restriction on the shape, hence no parametric form of the relationship is imposed. Instead, a fully non-parametric function is estimated:

$$y_i = f(x_i) + \epsilon_i \tag{1}$$

The method used here to fit y on x is locally linear estimation[7]. At every pair of observations (x_i, y_i), a linear relation is fitted around this pair, using the neighboring observations with a weight decreasing in the distance to x_i. This method is proven to be efficient and is especially not biased at the left and right boundaries, where only few observations can be found. The weights applied here come from the Gaussian kernel, such that the whole range of x and y is used for each local regression, however, with differing weights. For our application, the slight efficiency loss of the Gaussian kernel as compared to the Epanechnikov kernel[8] is more than outweighed by the computational advantages, which cover the exclusion of too many sparse matrices.

For the application of the kernel it remains to be determined how fast the weights decrease.[9] The bandwidth h is chosen according to the asymptotically optimal plug-in method proposed by Loader (2004):

$$h = \left(\frac{\sigma^2 (b-a)^2 \int K(v)^2 dv}{n \left(\int v^2 K(v) dv \right)^2 \int m''(x)^2 dx} \right)^{1/5}, \tag{2}$$

where σ^2 is the error variance, $m''(x)$ is the second derivative of the estimated function, and a and b are the lower and upper bounds of x. Using a first stage or pilot estimate, the error variance can be estimated by

[7] In principle, this method goes back to Cleveland (1974). On implementation and selection of the smoothing parameter, refer to the more recent Loader (2004).

[8] See Mittelhammer, Judge, and Miller (2000), p. 606.

[9] For a finite kernel, e. g. the triangular or rectangular kernel, this choice corresponds to the choice of the distance around (x_i, y_i) that determines the included observations for each local regression.

$$\widehat{\sigma^2} = \frac{1}{n - 2\nu_1 + \nu_2} \sum_{i}^{n} (y_i - m(x_i))^2, \tag{3}$$

with ν_1 and ν_2 as adjustment for the degrees of freedom (see again Loader (2004) for the computation). The second derivative $m''(x)$ of the estimate is obtained by a local *quadratic* regression to the data, which serves as the pilot estimate. It remains open to pick a bandwidth for the pilot estimate; in this analysis, $h_{\text{pilot}} = 5n^{-1/5}$ is used, which is of the optimal order of magnitude.[10]

4.2 Bivariate Locally Linear Regression

In principle, the regressor x is not restricted to be a scalar. Any regressor matrix of the dimension $n \times k$ can be implemented. In the case of two regressors, the result is a surface of fitted $m(x_{i1}, x_{i2})$ values above the (x_1, x_2) plane, which allows to extract as partial results all conditional moments, namely the functions $m(x_{i1}|x_{i2})$ and vice versa. Though this can (in theory) easily be extended to $k > 2$, two restrictions arise: First, a diagrammatic illustration of $m(x_{i1}, x_{i2}, \cdots, x_{ik})$ is no longer feasible. Secondly, the so-called curse of dimensionality arises, which states that the number of observations has to be increased more than proportionally with each additional regressor if the same degree of precision[11] is desired. Generally, none of the $x_j, x \in (1, \ldots, k)$ should be a constant, as the level in $m(x_{ij})$ is determined via the *local* regressions.

Though the data sets used in this analysis are comparatively large, we

[10] The `GAUSS 7.0` code of this procedure we have written and applied is available upon request. In order to speed up computation, the final estimation of the conditional moment vector $E(y|x)$ is performed on an equally spaced grid of 50 points on the total range of x, whereas for each *local* regression the complete x and y vectors are used. The computation of the plug-in bandwidth, however, requires that the complete set of observations is used in a locally quadratic regression. The standard approach in this case would be cross validation, which is more burdensome in terms of computer resources. In addition, the determination of the smoothing parameter by plug-in methods is more stable (see Fan (2000) for a short overview), which is corroborated by experiments with smaller subsamples of the data set on hand, where the cross-validated bandwidth varies by the factor 5. Our bandwidth-selection procedure requires for a set of variables x, y with approx. 50,000 observations roughly 1:05 hours of CPU time on a `P4HT` with 2.8GHz for each conditional moment vector.

[11] See Yatchew (2003), pp. 17. "Degree of precision" is inversely defined by the approximation error, which has the order of magnitude of $1/n^{1/k}$.

restrict our regressor matrix to be of dimension $k = 2$ in order to allow for a full graphical inspection of the results.[12]

4.3 Endogeneity

Using total benefit claims (hence, PSEGPT) as explanatory variable, endogeneity might arise: Duration is directly shortened by delayed retirement, whereas benefit claims rise, actually indicating a negative relationship between the two variables. The problem does not appear if claims are used on a yearly basis, which are not directly affected by the retirement age.[13] A different approach is to account for endogeneity by the use of an instrument, e. g. benefit claims normalized at a certain age. This can be constructed by subtracting (adding) the claims which (would) have been earned between retirement age and the normalization age. Although the method used here is fundamentally non-parametric, an approach similar to 2-stage least squares can be used. As a first step, the relationship between the instrument and the original explanatory variable is estimated by least squares:

Suppose that x_i is endogenous with respect to y_i; however, there exists a variable $\widetilde{x_i}$ which does satisfy the restriction of conditional orthogonality and which is associated with the original x_i by

$$x_i = \widetilde{x_i}\theta + u_i. \tag{4}$$

Under the further assumption of linearity of $E(\epsilon_i|x_i, u_i) = u_i'\rho$, the relationship among the residuals can be expressed by

$$\epsilon_i = u_i'\rho + \nu_i, \tag{5}$$

where ϵ_i is the residual of the model in equation (1). The endogeneity-

[12] Also in the bivariate case, the final estimation is performed on a grid of 25 equally spaced data points on *each* of both x vectors. The local regression on each grid point, however, uses again the whole sample.

[13] Average benefit claims per year of contribution may still be affected if income (and therefore contributions) are rising during the career. Delayed retirement would then lead to a larger weight on the high-income period at the end of a career. This would not vanish once a hypothetical measure of "benefit claims at a certain age" is used directly, since these are not directly observable and have to be constructed, using the average claims again.

adjusted model can then be written[14] as

$$y_i = f(x_i) + u_i'\rho + \nu_i \tag{6}$$

and finally, in the second step, be estimated semi-parametrically, with u_i entering the model as additional parametric regressor. As u_i is not directly observable, it has to be replaced by an estimate, namely by the residual of equation (4) estimated by least squares. The function f can be identified as follows: First, write the model in conditional expectations:

$$E(y_i|x_i) \quad = \quad f(x_i) + E(u_i|x_i)'\beta \tag{7}$$

The conditional expectations are now estimated non-parametrically, by fitting a local polynomial. Denote the estimates by

$$\begin{aligned} \widehat{E(y_i|x_i)} &=: m_y(x_i) \\ \widehat{E(u_i|x_i)} &=: m_u(x_i). \end{aligned}$$

The partially linear model in terms of conditional expectations of equation (7) can then be rewritten as

$$y_i - m_y(x_i) = [u_i - m_u(x_i)]'\beta + \epsilon_i, \tag{8}$$

and β can be estimated by least squares. Denoting the estimate by $\widehat{\beta}$, an estimate for $f(x_i)$ is finally obtained by

$$\widehat{f}(x_i) = m_y(x_i) - m_u(x_i)'\widehat{\beta}. \tag{9}$$

The significance of IV-residual in the final partially linear estimation will falsify the hypothesis of exogeneity of the respective regressor. As the

[14]See Yatchew (2003), pp. 87 for a textbook treatment, Speckman (1988) for the introduction of partially linear models with smoothed conditional moments as statistical method, and Blundell and Duncan (1998) for a seminal application, including the problem of endogenous regressors.

instrument $\widetilde{x}$, we use the sum of total claims, normalized to the claims hypothetically earned until the age of 65:

$$\widetilde{x} = x + \frac{x}{\mathrm{BYVL}}(65 - \mathrm{RET1}), \tag{10}$$

where x is the sum of claims and RET1 is the age at retirement into the first pension.[15]

4.4 Validity of Instruments

In the least squares framework, we can apply formal tests for the validity and strength of the instruments[16] we use. First, the instruments used in the regressions are strong. A standard F-test applied to the first stage regressions in a 2SLS setup (which also applies for the first stage regression in the partially linear framework) never falls below 33609.13 in all specifications. The requirement of no correlation between the instruments and the residual is also fulfilled; the maximum correlation coefficient between one of the instruments and the residual of each IV-regression is (rounded) 0.0000 and therefore *very* small, and more important, not significantly different from zero.

4.5 Least Squares Regression

In order to quantify the impact of the independent variables we apply a standard least squares approach as well. The main independent variable "total claims" or "claims per year" x enter as a polynomial of degree 3, the remaining set of covariates (including a constant) form the matrix z, such that the regression equation is

$$y = \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + z\gamma + \epsilon. \tag{11}$$

[15] See Angrist and Lavy (1999) for another example and a justification of the validity of the prediction of a variable as its instrument.

[16] In the partially linear regressions, we use only one instrument, namely the predicted benefit claims for benefit claims. In the least squares regression framework, however, we fit a polynomial of degree 3 in the explanatory variable, so we need (at least) 3 instruments. We use squared and cubic terms of the predicted benefit claims as second and third instrument.

In the case of total claims as independent variable, we use two versions of the matrix z, first the one with original regressors, and second, one with fitted regressors $\hat{z}$ from a first stage regression of z on the instrument for duration, $\widetilde{x}$. As we need at least as many instruments as endogenous variables, we compute the squared and cubic values of our instrument as well.

5 Results and Political Relevance

5.1 General and Differential Results

The most striking single result can be found in figure 2. Using the restricted sample (which we find is the relevant one) and instrumenting total benefit claims, the relationship between benefit claims and duration from the first benefit payment is remarkably close to linear for male pensioners, although by application of non-parametric estimation of the conditional expectations, no shape (especially no linear relation) has been imposed ex ante. Using total benefit claims PSEGPT as explanatory variable, the results differ substantially once PSEGPT is instrumented and the shape of the relationship is estimated by partially linear methods. See figures 9 through 11 for the differential results using the whole sample, and figures 2 through 4 using the sample restricted to observations with at least 25 years of work. Our variable of interest, the duration from the first benefit payment[17], depends positively on total claims, both for female and male pensioners in the complete sample, whereas the sign of the relationship for female pensioners is reversed once the sample is restricted to the minimum number of years of work. This reversed shape is the main reason to exclude female observations from the following differential results, which we derive by residence and by months of unemployment for male pensioners.

Whether the pensioner lived in the West or in the East part of Germany does not matter for the general positive relationship between income and duration. The relationship for those living abroad is unspecific and subject to a relatively large variation in all our specifications. We observe this pattern both in the complete as well as in the restricted sample. The impact of total claims stratified along the time spent in unemployment yields that duration in all subgroups (with no time in unemployment, between 1 and 6 months in

[17]The results with Duration 1 as dependent variable can always be found in the left panels, and the results with respect to Duration 2 in the right panels, respectively.

unemployment, and more than 6 months in unemployment) reacts positively on total claims, however, the gradient becomes flatter the more time the pensioner has spent in unemployment, meaning that the impact of income on duration is less important. Still, the average duration is higher for those being only a few months or even never unemployed.

Once we use average claims PY as independent variable (which does not have to be instrumented, see figures 14 and 15 for the complete sample and figures 7 and 8 for the restricted sample), the results are qualitatively similar. Namely, we do find evidence for a positive impact of income on duration, especially for male pensioners. Yet, the impact stratified along residence or unemployment does not reveal the same pattern as with PSEGPT as explanatory variable.

The least squares regressions with polynomials in benefit claims also yield similar results. Applying this method, the significant signs of the linear, quadratic, and cubic elements indicate an inverse U-shape relationship in all specifications (meaning that there is an increasing relationship over a significant range of the income distribution), including the IV-regressions. All specifications yield only a low R-squared, despite the high significance of the regressors, so a major fraction of variation in duration is not explained by income.

5.2 Differential Results by Measures of Health Status

From a redistributional point of view, we derive the most interesting results with a stratification along the individual health status. Health is measured twofold, firstly, by a dummy variable indicating whether the individual bought private health insurance or whether he was insured in the public health insurance (see figures 12 and 5 for the case of PSEGPT being the explanatory variable). Secondly, we distinguish three groups by the time they spent in ill-health during their career, starting from no time at all, between 1 and 6 months, or more than 6 months (see figures 13 and 6 for the case of PSEGPT being the explanatory variable). In general, the impact of income on duration is stronger the worse the health status is. Assuming that ceteris paribus the choice of public health insurance is an indicator for worse health, observations with public health insurance reveal a positive relationship between income and duration, no matter if PSEGPT or PY (see again figures 14 and 7) is used as explanatory variable and no matter if the data set is restricted or not. Individuals with private health insurance exert a rather unspecific pattern instead.

If health is measured by the months spent in ill-health, the general finding is that the income gradient is positive for all subgroups. However, using PY as explanatory variable (figures 14 and 15 for the total sample and figures 7 and 8 for the restricted sample), the univariate as well as the bivariate estimations show that the gradient is steeper the more time has been spent in ill-health.[18]

The effect of time in ill-health estimated by least squares or 2SLS is complex as well: In both samples (total and restricted), the duration under the benefit spell is positively related to time spent in ill health. This is driven by institutional factors such as the possibility of a disability pension, as the comparison with duration 2 as dependent variable shows. Here, duration under the benefit spell currently paid at the time of death is higher for those who are poor, but healthy; having spent time in ill health however, leaves the poor worse off, such that being relatively rich in terms of benefit claims even overcompensates being healthy. Secondly, the impact of both health measures included in a straight forward regression equation yields the following results: In all specifications, time spent in ill health significantly increases duration, which (unsurprisingly) corroborates out non-parametric results (see table 3), and the choice of public health insurance is significantly associated with lower duration.

5.3 Policy Remarks

The question of heterogeneous rates of return, based on different income groups cannot be answered easily, for several reasons. The most important one is the fact that the pattern is not consistent across different subgroups, and it is beyond the focus of this essay to name the subgroup which is relevant, although we can give some guidance for the discussion. Furthermore, though in some cases monotonous relations can be extracted in the locally

[18] The three curves from the respective univariate estimations can be understood as conditional expectations (or marginal distributions) from the bivariate estimations. Graphically, the direct univariate estimates and the marginals from the bivariate estimates are not exactly congruent, as the bandwidths differ. Note that in this case, the bivariate bandwidths are not chosen optimally, but by subjective comparison of the smoothness/roughness of the resulting plots. For this estimation, our choice is $h = 1.2$. The general shape is not affected by the bandwidth choice; the higher the bandwidth is in this case, the closer is the estimation to a plane surface.

linear estimations, the results are not always robust with respect to the choice of the sub-sample. This phenomenon is also expressed in the low R^2 of all least squares regressions, even despite the high significance of benefit claims measures. A major part of the variation in duration has to be explained by variables other than benefit claims.

However, the sub-group which indeed suffers from a positive income gradient are male pensioners with at least 25 years of work, who have public health insurance and who have spent some time in ill-health. This subgroup is not marginal. Still: other major subgroups, especially female pensioners, exert an either unspecific or even *negative* pattern.

From the preceding analysis we know that people in ill-health slightly profited from the pension system, as their duration was ceteris paribus higher. This can easily be explained by the existence of pension due to occupational disability[19], which, however, was restricted in the year 2001 to those born before 1961. Yet, the sub-sample of these individuals born afterwards is small and nobody out of this cohort has reached the usual retirement age so far. The eligibility for those born later is subject to stricter conditions of occupational disability: the retirement age for the severely disabled has been increased from 60 to 63.[20] The effect of both policy measures is not fully observable yet, though the conjecture is obvious that people in ill-health will most likely suffer from it in terms of a reduced duration. Analogously, the increased minimum age for a pension after unemployment will most likely reduce the duration and therefore the return from the pension system for the unemployed. At the same time, the effect of income on duration and hence redistribution from poor to rich becomes more pronounced for individuals in a poor health status.

Although we cannot answer whether this form of redistribution *should* be neutralized within the public pension system[21], we can provide guidelines

[19]Which is not the same as "general" disability.

[20]See e. g. Deutsche Rentenversicherung Bund (2005), pp. 263 for a short chronological overview of changes in the pension system.

[21]One could argue that the issue of redistribution of any kind has to be addressed by income taxation; the argument made here would then be in favor of a (ceteris paribus) more progressive income taxation. However, Joaquim Oliveira Martins kindly discussed these findings on the 5th Workshop on Pension and Saving and brilliantly argued that a higher rate of return from the pension system for high-income individuals could actually be optimal, once education is taken into account. We know that education both increases income and life-expectancy, such that a higher rate of return due to higher life-expectancy can be understood as a fraction of the education premium, which therefore provides further

what mechanisms *could* reduce the implicit transfers from poor to rich. If we believe in the partially linear specification which accounts for endogeneity, we have to address the question of a feasible policy instrument which is capable of softening the positive ability-duration relationship. One instrument which is already in use are discounts for early retirement. In general, each month the retirement age falls short of the age of 65 (for males) comes with the loss of 0.3% of the monthly pension benefits. We can use this device and adjust the monthly discounts in a way to approximate a rate of return which is independent of the total benefit claims. With an additional degree of freedom, conditioning these discounts on other parameters would allow to come even closer to the ideal of an independent rate of return. Without addressing legal impediments (or even the political process), an obvious dimension of discrimination would be the sex of the beneficiary. Furthermore, discounts may not be constant, but varying with each additional month of early retirement. The computation of such discounts (either constant or being a function of income) is subject of our current research. Further possibilities of adjusting the benefit formula include a progressive element, meaning that benefit claims increase less than proportionally with income. Such an instrument is in place, e.g. in the United States. If the benefit formula should be left unaltered, incentives and possibilities for early retirement (either into the old-age pension or into any kind of disability pension) for poor individuals and individuals in poor health add to an income-neutral rate of return.

6 Conclusion

We addressed the question whether the annuity payments in the German public pension system lead to redistribution from poor to rich. Although the major fraction of variation in duration cannot be explained by income, the specification which would in general allow for non-linear results and which accounts for possible endogeneity yields a positive and almost linear relation. The comparison with standard techniques shows that this impact is also significant. We disentangle the influence of income on duration for different sub-groups, especially stratified along the health dimension. In our sample, individuals in ill-health could still benefit from from a policy

incentives to educate. See Gorski, Krieger, and Lange (2007) and Lau and Poutvaara (2006) for a theoretical treatment of this phenomenon.

of disability pensions, which has been tightened in the meantime. However, conditional on a certain health status, the positive impact of income on duration is more pronounced, and hence is redistribution.

Beyond the scope of this paper are other major forces behind differential duration. Differences in retirement age and differences in realized life expectancy both drive duration and have to be analyzed separately, especially with respect to their sensitivity to measures of socio-economic status and health. Furthermore, the revealed relation gives rise to policy interventions to weaken the redistribution. A feasible instrument has to found, and we propose to adjust the discounts for early retirement, which is subject of our current research agenda.

A Tables

Descriptive Statistics				
Data set:	93-03 full		93-03 restr.	
n:	98,399		63,020	
Variable	Mean	St. Dev.	Mean	St. Dev.
PSEGPT	31.14	17.42	40.21	13.54
P/Y	1.06	0.43	1.07	0.312
BYVL	29.00	12.63	37.26	6.19
DUR1	7.38	6.18	7.23	5.69
DUR2	4.72	2.99	4.62	2.93
WHORT west	0.91	—	0.96	—
GEBC male	0.67	—	0.81	—
AT priv.	0.11	—	0.10	—
AUAZ	2.34	5.76	2.70	6.07
AUAZ (Dummy)	0.36	—	0.42	—
AJAZ	8.21	19.83	8.50	19.39
AJAZ (Dummy)	0.35	—	0.39	—

Restricted data sets include only observations with BYVL larger than 25.

Table 1: Descriptive Statistics

Plug-In Bandwidths

Data set:	full data set				restricted data set			
Indep. Var.	PSEGPT		P/Y		PSEGPT		P/Y	
Dep. Var.	DUR1	DUR2	DUR1	DUR2	DUR1	DUR2	DUR1	DUR2
GEBC male	29.152	23.953	32.221	24.540	21.794	19.172	25.895	19.650
GEBC female	22.230	17.890	20.928	19.572	13.271	11.753	13.498	11.777
BLAND west	22.399	20.293	22.031	19.602	21.238	18.519	20.490	18.644
BLAND east	18.338	14.633	5.856	5.012	3.990	3.819	3.762	3.894
BLAND abroad	9.117	8.900	9.082	8.878	5.750	5.572	5.873	5.739
AT public	26.490	21.500	21.055	19.476	19.653	18.172	19.198	18.513
AT private	9.690	9.051	8.606	8.493	8.324	7.120	8.258	7.018
AUAZ $= 0$	24.434	20.765	18.799	18.072	16.149	16.104	16.149	16.104
AUAZ ≤ 6	14.899	12.329	16.206	13.529	14.325	12.329	16.001	13.714
AUAZ > 6	11.886	9.658	14.849	10.775	11.509	9.658	10.688	8.415
AJAZ $= 0$	26.021	21.990	20.350	18.101	18.243	16.119	18.195	16.106
AJAZ ≤ 6	10.296	9.965	11.830	9.564	10.163	9.965	10.259	9.124
AJAZ > 6	13.791	12.628	13.972	12.784	11.446	12.628	12.704	11.631
IV-Residual	33.305	33.305	—	—	32.047	32.047	—	—

Restricted data sets include only observations with BYVL larger than 25.

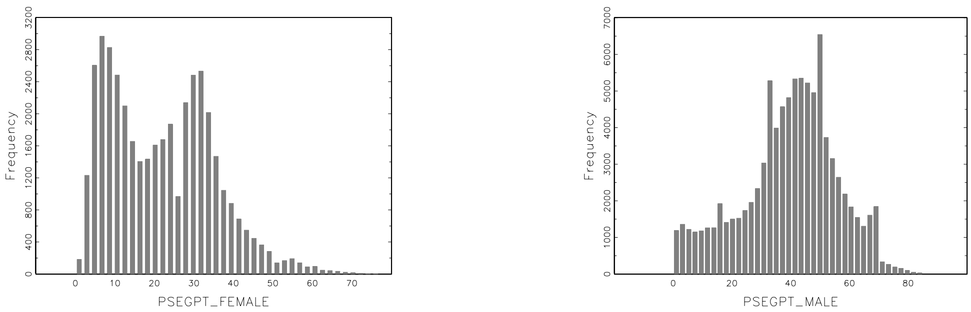
Table 2: Plug-In Bandwidths

Least Squares Estimation — Dependent Variable: ln (DUR1)						
Data Set	full		full		full, IV	
x	β	S.E.	β	S.E.	β	S.E.
constant	1.62***	.017	1.48***	.0252	.562***	.075
PSEGPT	.011***	.002	—	—	.178***	.012
$PSEGPT^2/100$	−.045***	.006	—	—	−.543***	.035
$PSEGPT^3/1000$	.004***	.001	—	—	.0473***	.003
P/Y	—	—	.329***	.070	—	—
P/Y^2	—	—	−.0196	.056	—	—
P/Y^3	—	—	−.040***	.013	—	—
WHORT west	.006	.014	−.106***	.013	−.329***	.028
AT priv	−.354***	.010	−.323***	.009	−.375***	.011
AUAZ	.026***	.001	.0249***	.001	.024***	.001
AJAZ	−.003***	.000	−.002***	.000	−.004***	.000
R^2	0.048		0.051			
Data Set	restr.		restr.		restr., IV	
x	β	S.E.	β	S.E.	β	S.E.
constant	1.41***	.033	1.48***	.034	−2.19***	.203
PSEGPT	.0206***	.003	—	—	.471***	.024
$PSEGPT^2/100$	−.054***	.009	—	—	−1.24***	.062
$PSEGPT^3/1000$	.005***	.001	—	—	.098***	.005
P/Y	—	—	.091	.115	—	—
P/Y^2	—	—	.059	.111	—	—
P/Y^3	—	—	−.011	.034	—	—
WHORT west	−.023	.021	−.007	.022	−1.39***	.080
AT priv	−.439***	.012	−.450***	.011	−.492***	.014
AUAZ	.025***	.001	.024***	.001	.022***	.001
AJAZ	−.002***	.000	−.002***	.000	−.004***	.000
R^2	0.054		0.057			

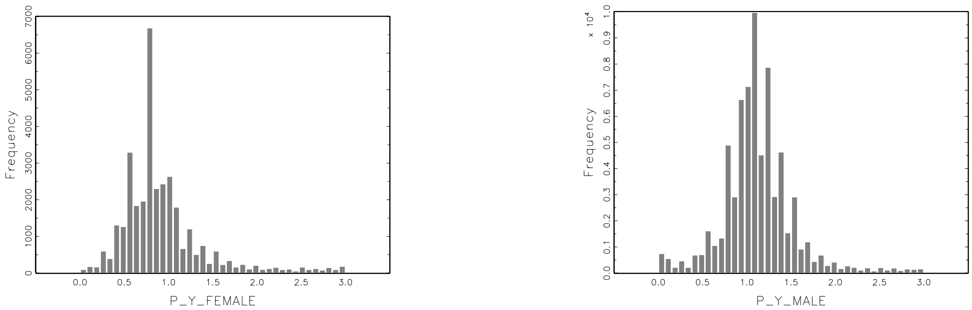
Restricted data set include only observations with BYVL larger than 25. Robust Standard Errors, * denotes significance on 90% level, ** denotes significance on 95% level, and *** denotes significance on 99% level. Only male observations included.

Table 3: Least Squares Estimation, Data Set 1993 - 2003

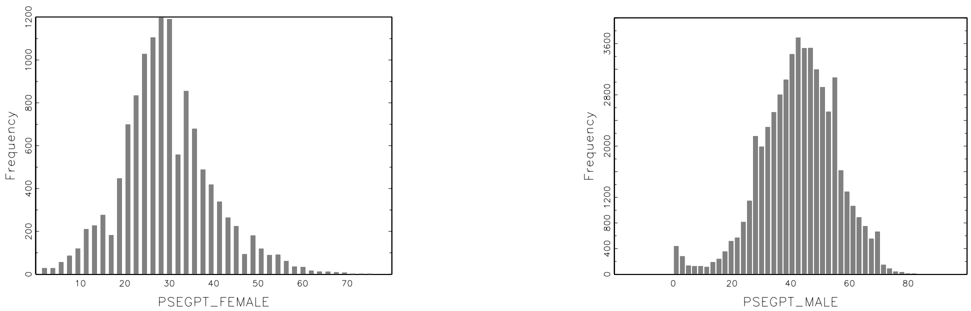
B Histograms



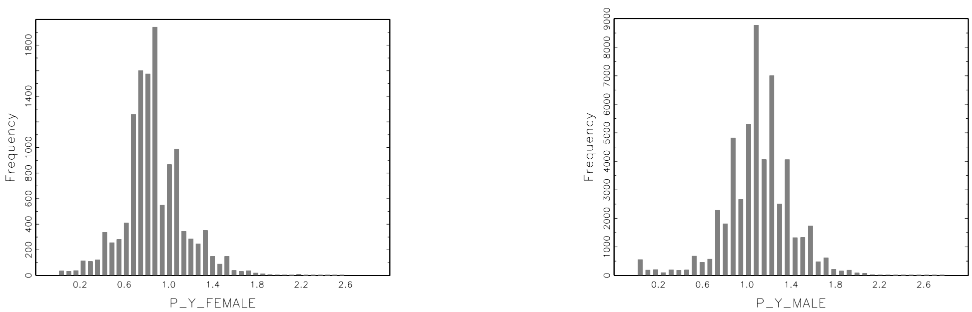
Right panel: female observations. Left panel: male observations.



By sex, left panel: duration 1. Right panel: duration 2. Solid: male. Dashed: female.



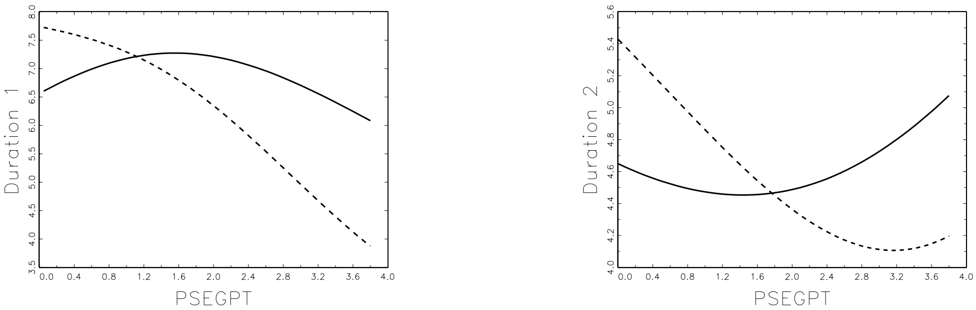
Right panel: female observations. Left panel: male observations.



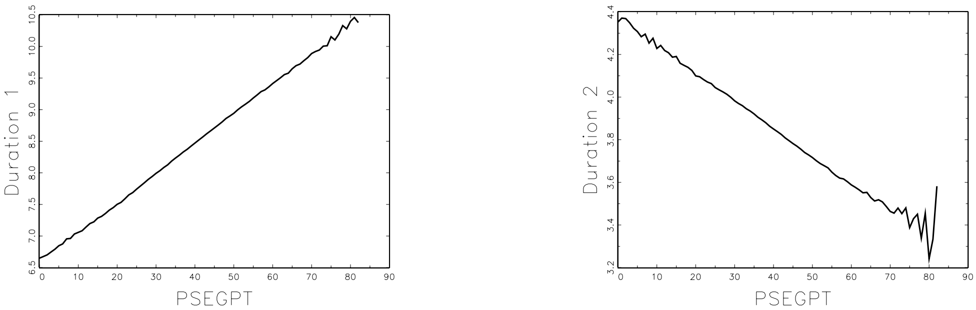
By sex, left panel: duration 1. Right panel: duration 2. Solid: male. Dashed: female.

Figure 1: Histograms

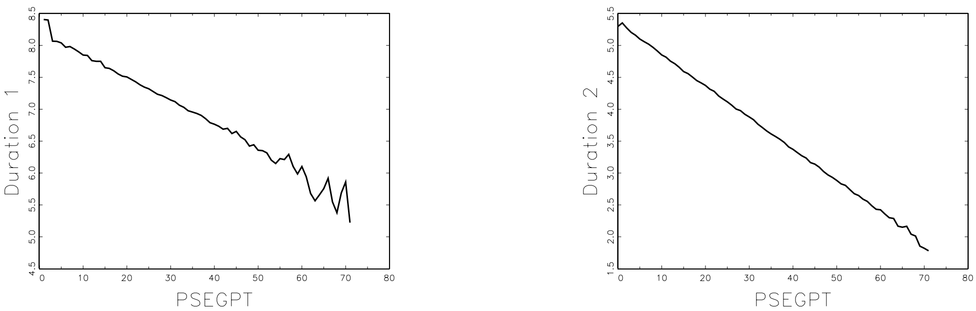
C Restricted Sample, BYVL ≥ 25



By sex, left panel: duration 1. Right panel: duration 2. Solid: male. Dashed: female.

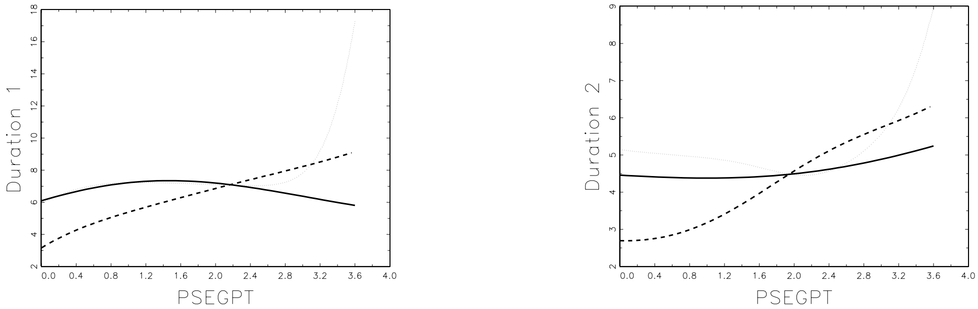


Only male obs., left panel: duration 1. Right panel: duration 2. Partially linear, PSEGPT instrumented.

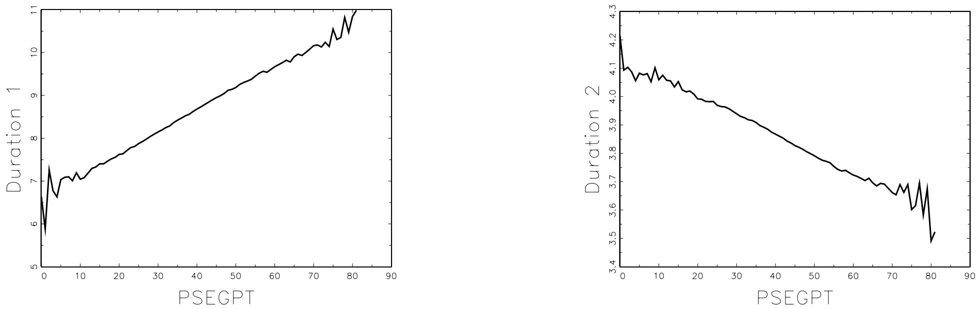


Only female obs., left panel: duration 1. Right panel: duration 2. Partially linear, PSEGPT instrumented.

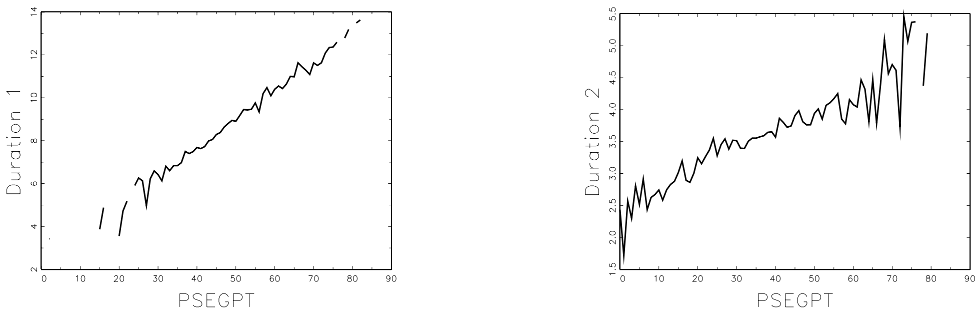
Figure 2: Explanatory Variable: PSEGPT, Stratification by Gender



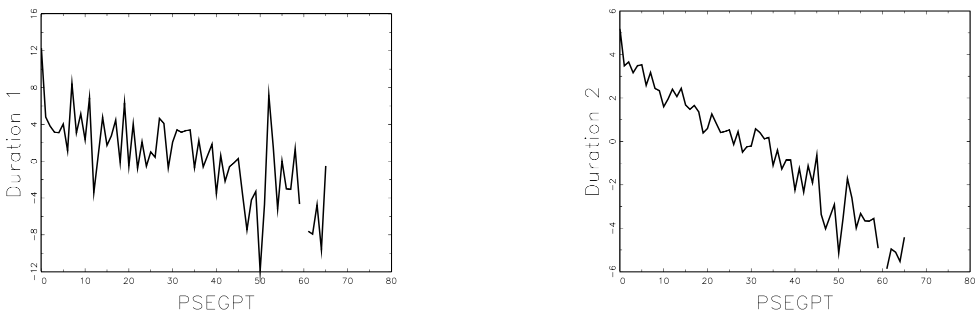
Left panel: duration 1. Right panel: duration 2. Solid: west. Dashed: east. Dotted: abroad. Male observations only.



Only West, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

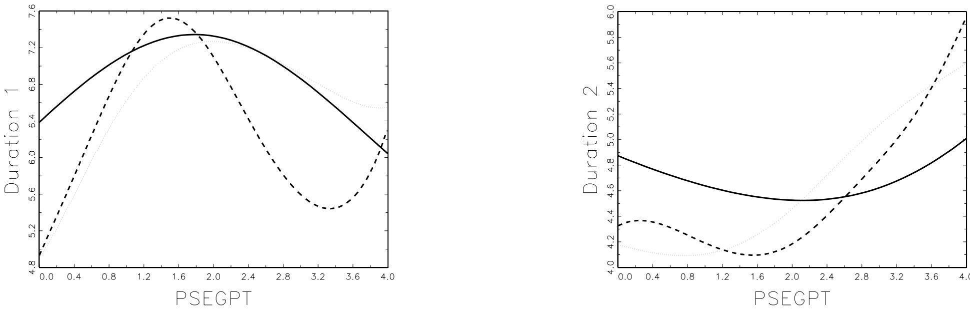


Only East, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

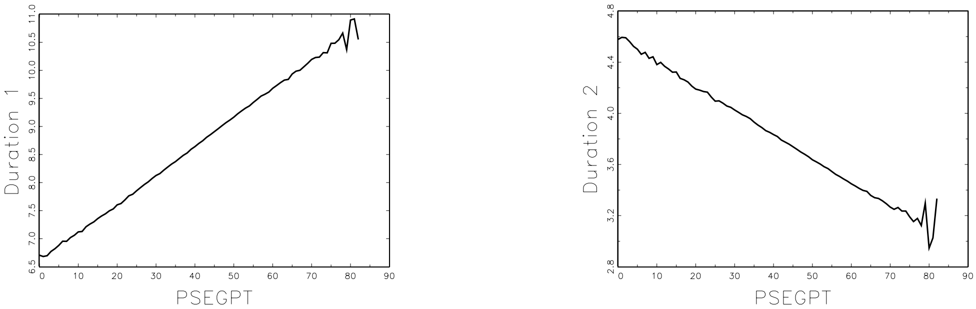


Only Abroad, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

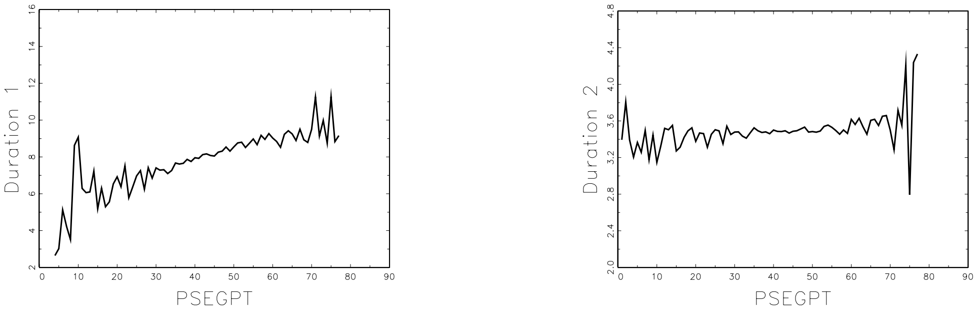
Figure 3: Explanatory Variable: PSEGPT, Stratification by Residence



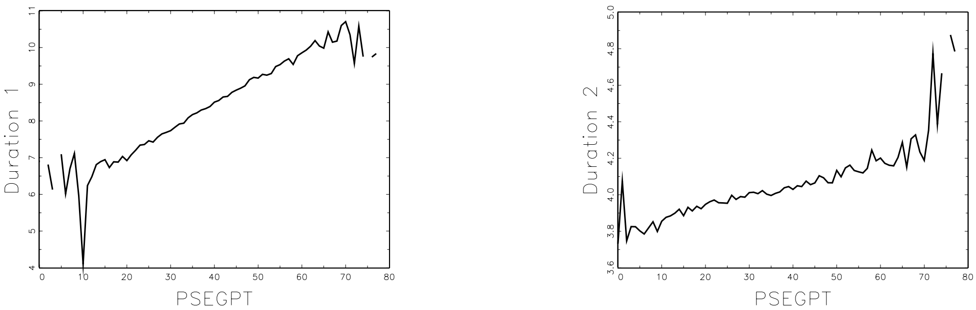
Left panel: duration 1. Right panel: duration 2. Solid: AJAZ = 0. Dashed: 0 < AJAZ ≤ 6. Dotted: AJAZ > 6. Male observations only.



Only AJAZ = 0, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

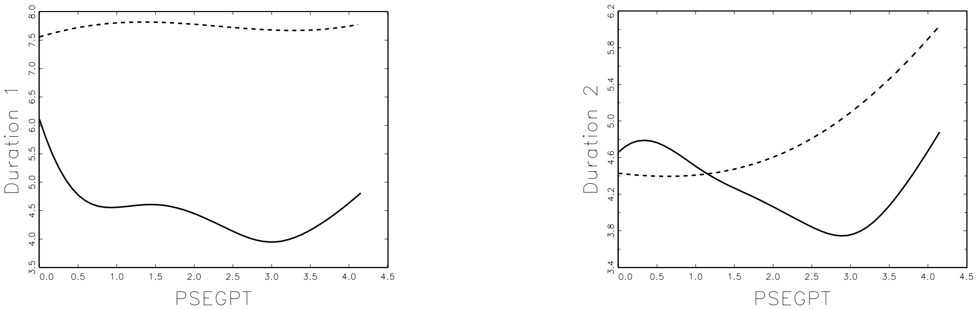


Only 0 < AJAZ ≤ 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

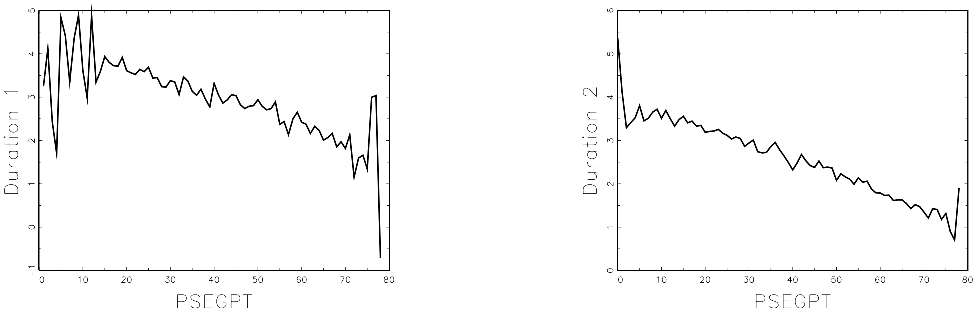


Only AJAZ > 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

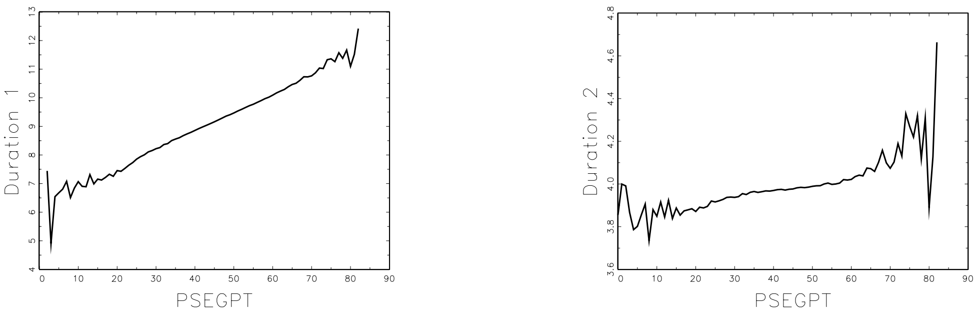
Figure 4: Explanatory Variable: PSEGPT, Stratification by Months in Unemployment



Left panel: duration 1. Right panel: duration 2. Solid: private health insurance.
Dashed: public health insurance. Male observations only.

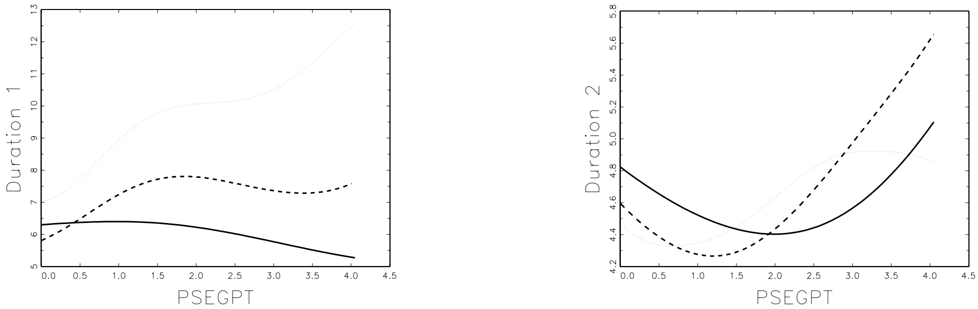


Only private insurance, left panel: duration 1. Right panel: duration 2. Male obs. only.
Partially linear, PSEGPT instrumented.

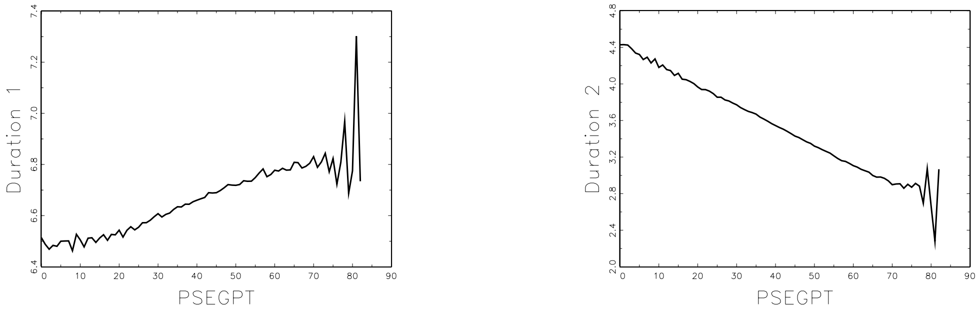


Only public insurance, left panel: duration 1. Right panel: duration 2. Male obs. only.
Partially linear, PSEGPT instrumented.

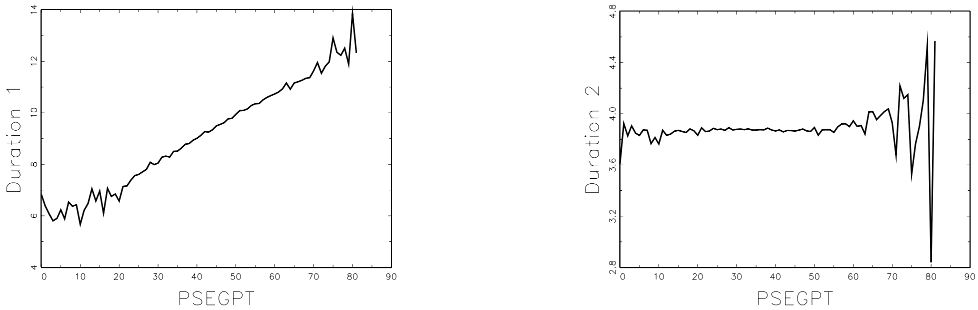
Figure 5: Explanatory Variable: PSEGPT, Stratification by Health Insurance



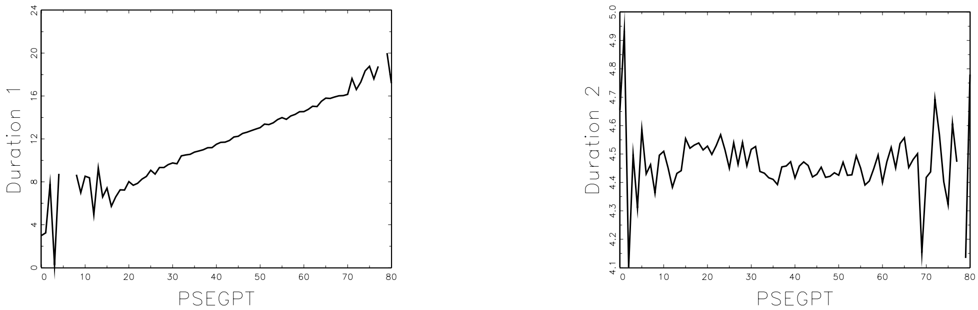
Left panel: duration 1. Right panel: duration 2. Solid: AUAZ = 0. Dashed: 0 < AUAZ ≤ 6. Dotted: AUAZ > 6. Male observations only.



Only AUAZ = 0, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

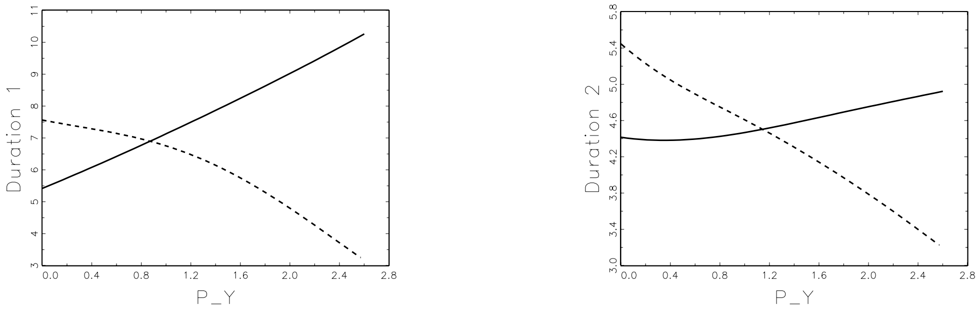


Only 0 < AUAZ ≤ 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

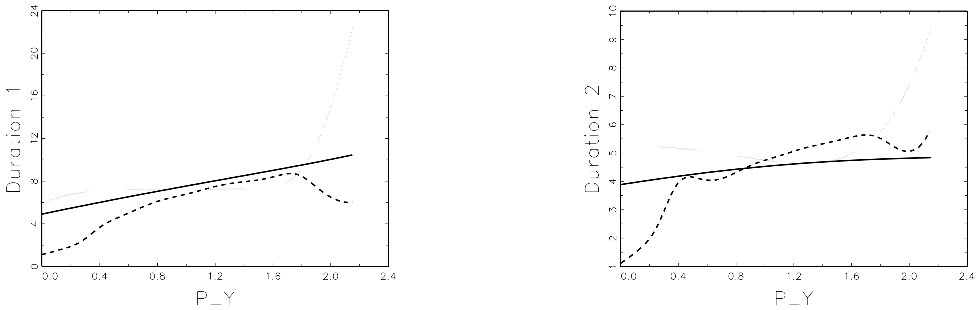


Only AUAZ > 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

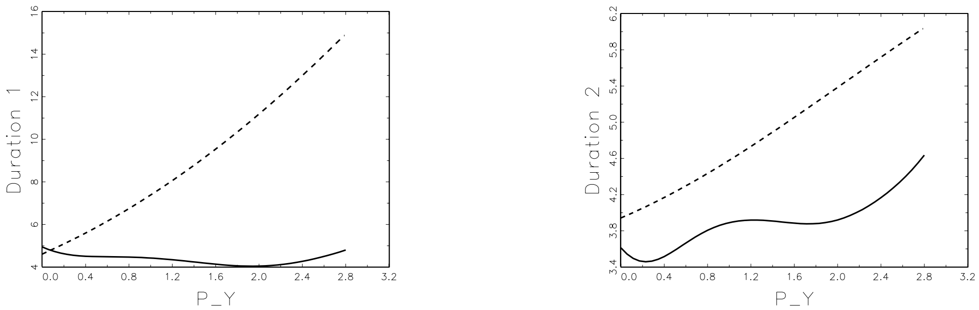
Figure 6: Explanatory Variable: PSEGPT, Stratification by Months in Ill-Health



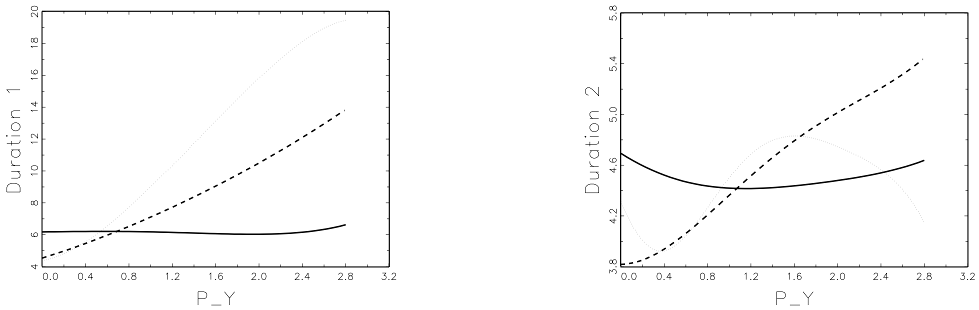
By sex, left panel: duration 1. Right panel: duration 2. Solid: male. Dashed: female.



By residence, left panel: duration 1. Right panel: duration 2. Solid: west. Dashed: east. Dotted: abroad. Male observations only.

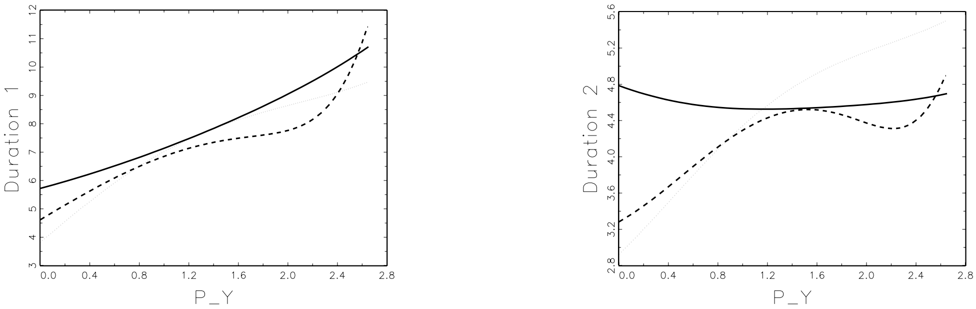


By health insurance, right panel: duration 1. Left panel: duration 2. Solid: private. Dashed: public.

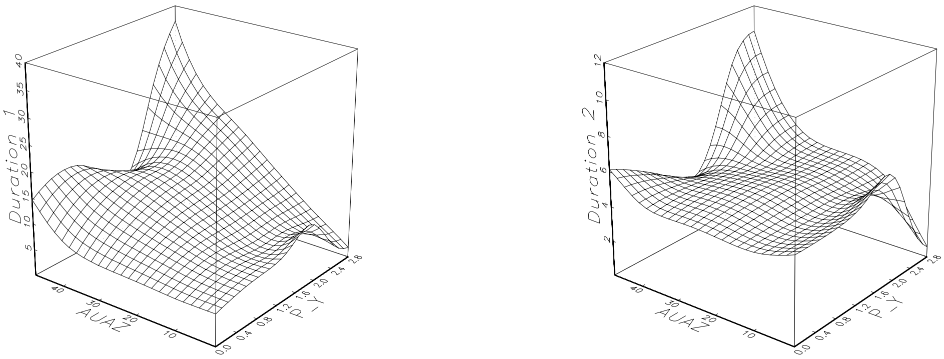


By months in illness, right panel: duration 1. Left panel: duration 2. Solid: no time in illness. Dashed: up to 6 months in illness. Dotted: More than 6 months in illness

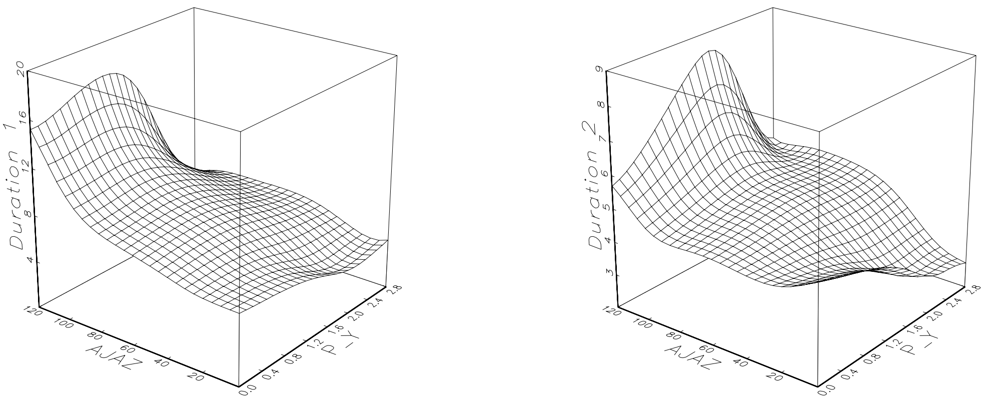
Figure 7: Explanatory Variable: PSEGPT/BYVL



By months in unemployment, right panel: duration 1. Left panel: duration 2. Solid: no time in unemployment. Dashed: up to 6 months in unemployment. Dotted: More than 6 months in unemployment



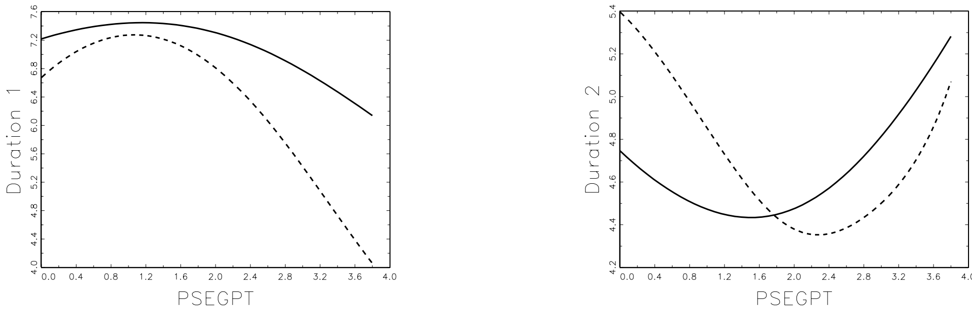
By time in illness, left panel: duration 1. Right panel: duration 2. Male observations only.



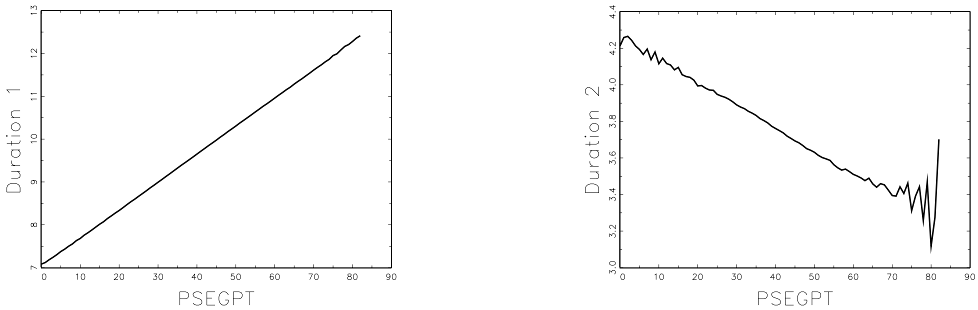
By time in unemployment, left panel: duration 1. Right panel: duration 2. Male observations only.

Figure 8: Explanatory Variable: PSEGPT/BYVL

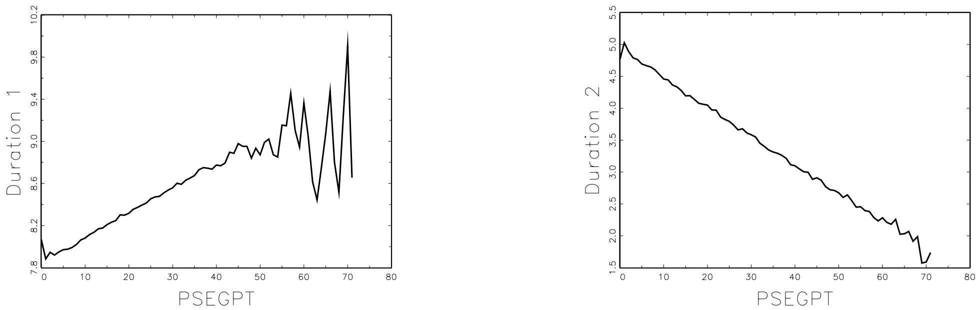
D Total Sample



By sex, left panel: duration 1. Right panel: duration 2. Solid: male. Dashed: female.

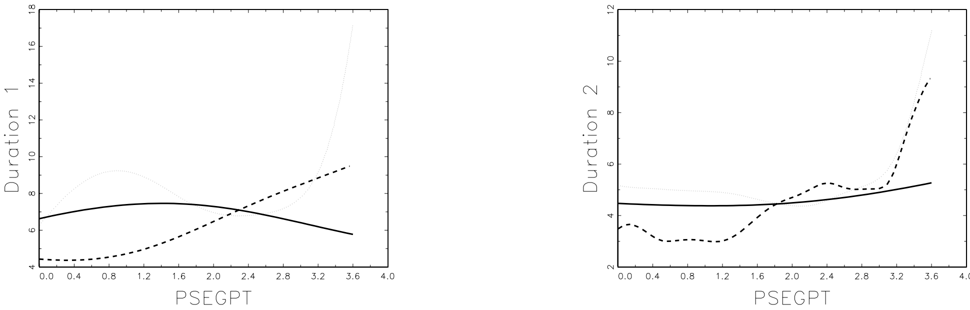


Only male obs., left panel: duration 1. Right panel: duration 2. Partially linear, PSEGPT instrumented.

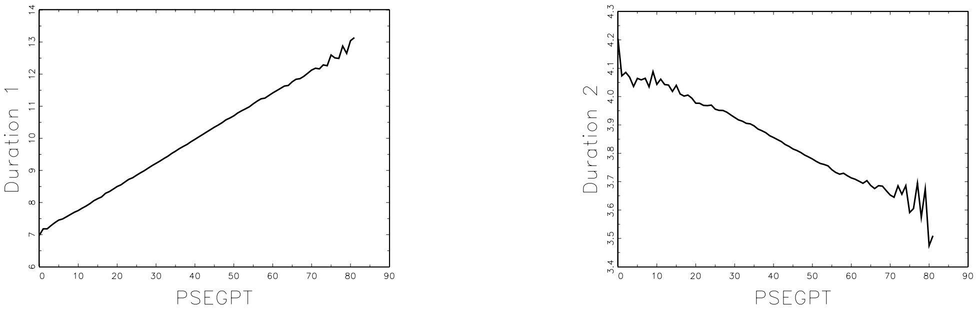


Only female obs., left panel: duration 1. Right panel: duration 2. Partially linear, PSEGPT instrumented.

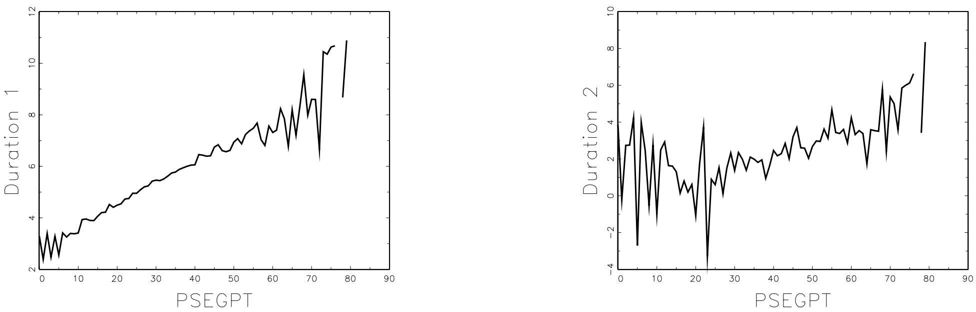
Figure 9: Explanatory Variable: PSEGPT, Stratification by Gender



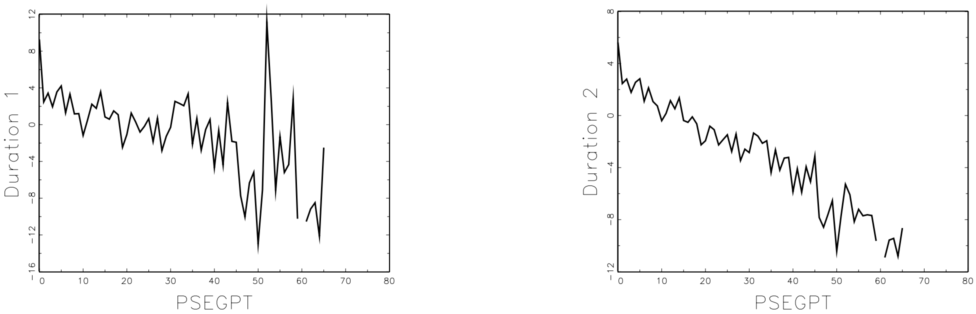
Left panel: duration 1. Right panel: duration 2. Solid: west. Dashed: east. Dotted: abroad. Male observations only.



Only West, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

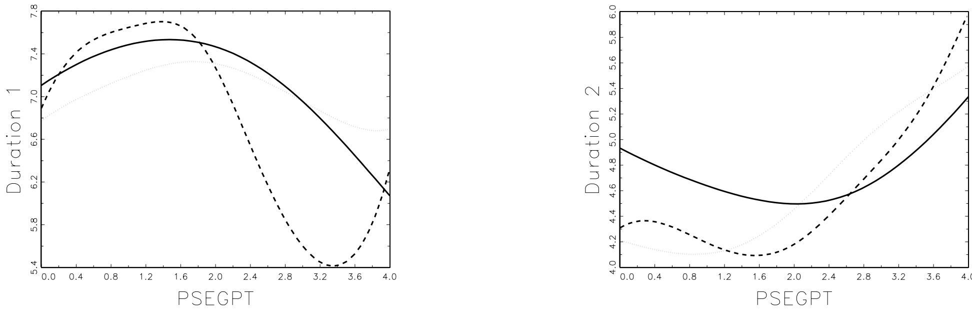


Only East, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

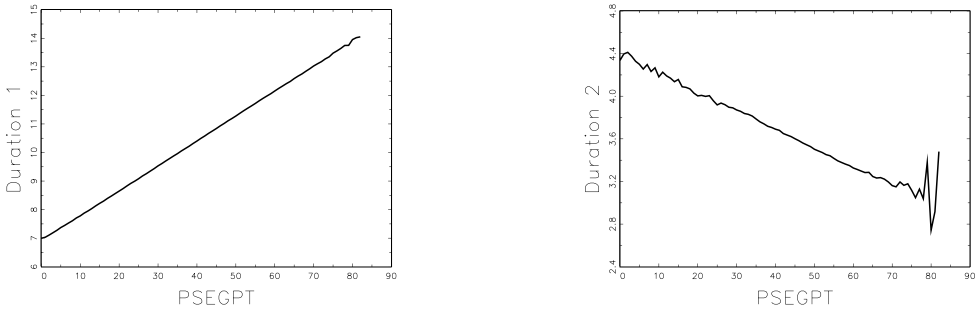


Only Abroad, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

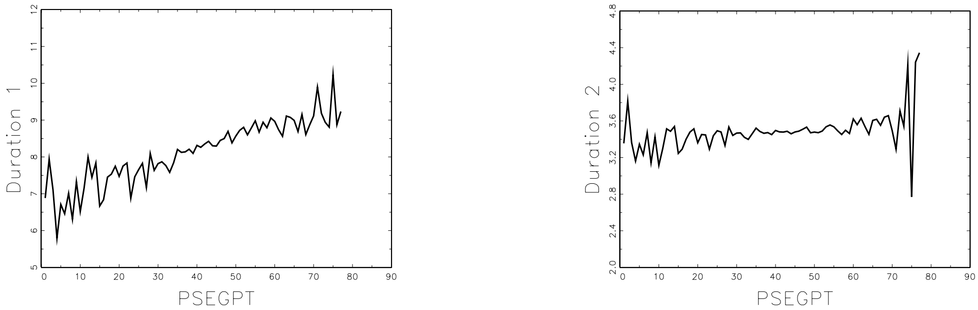
Figure 10: Explanatory Variable: PSEGPT, Stratification by Residence



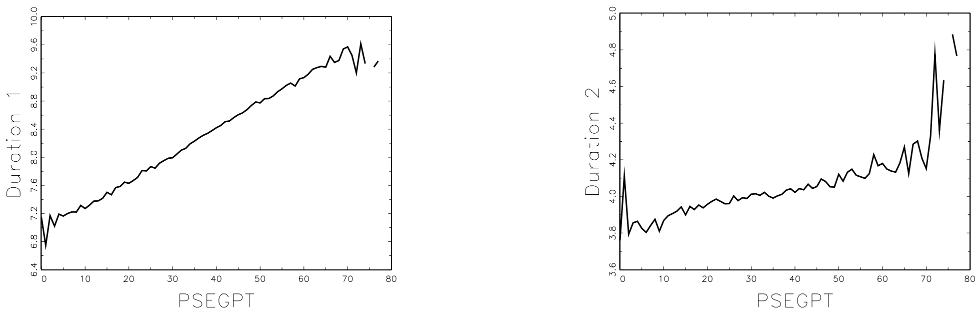
Left panel: duration 1. Right panel: duration 2. Solid: AJAZ = 0. Dashed: 0 < AJAZ ≤ 6. Dotted: AJAZ > 6. Male observations only.



Only AJAZ = 0, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

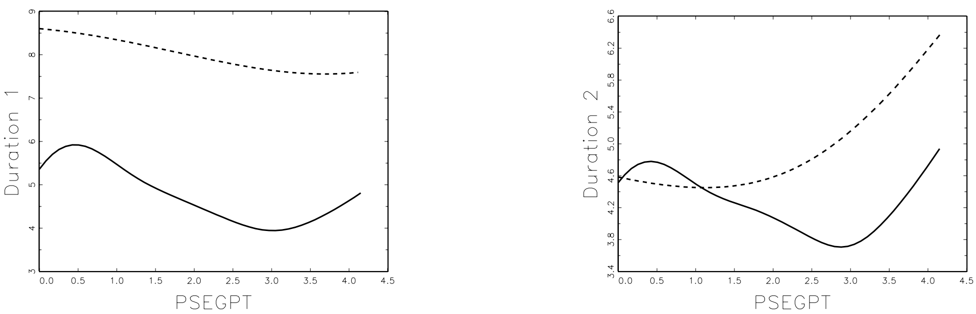


Only 0 < AJAZ ≤ 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

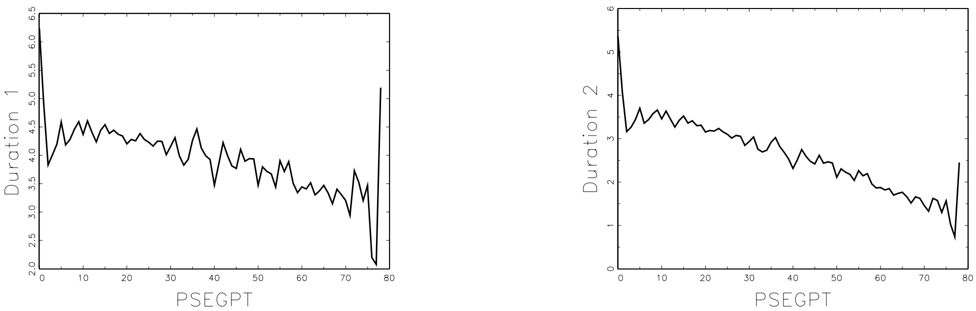


Only AJAZ > 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

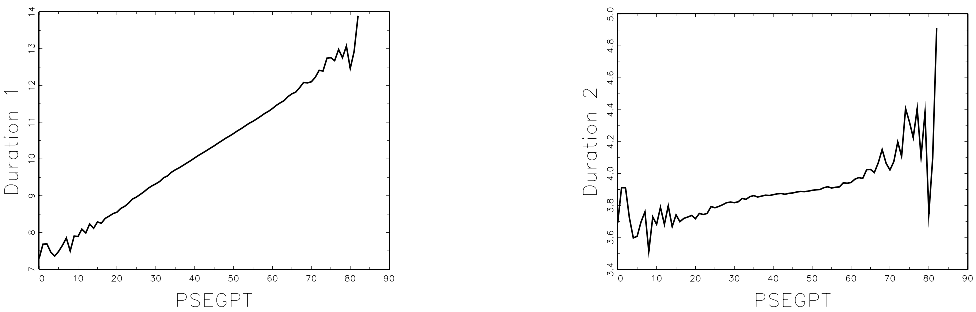
Figure 11: Explanatory Variable: PSEGPT, Stratification by Months in Unemployment



Left panel: duration 1. Right panel: duration 2. Solid: private health insurance.
Dashed: public health insurance. Male observations only.

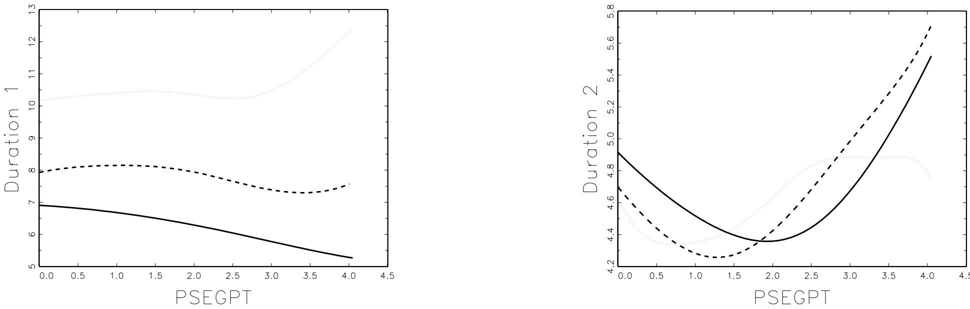


Only private insurance, left panel: duration 1. Right panel: duration 2. Male obs. only.
Partially linear, PSEGPT instrumented.

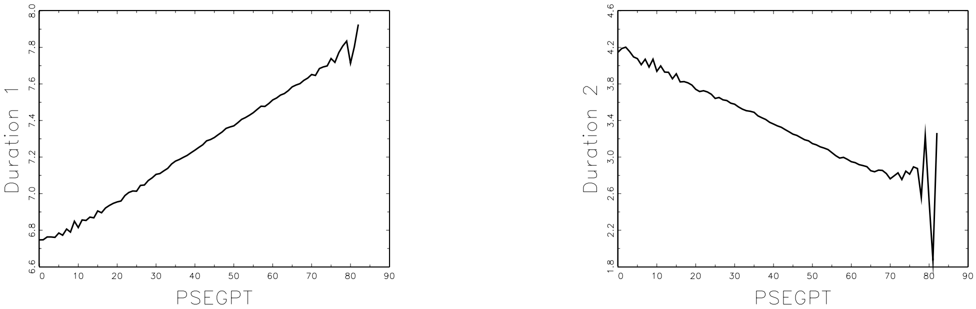


Only public insurance, left panel: duration 1. Right panel: duration 2. Male obs. only.
Partially linear, PSEGPT instrumented.

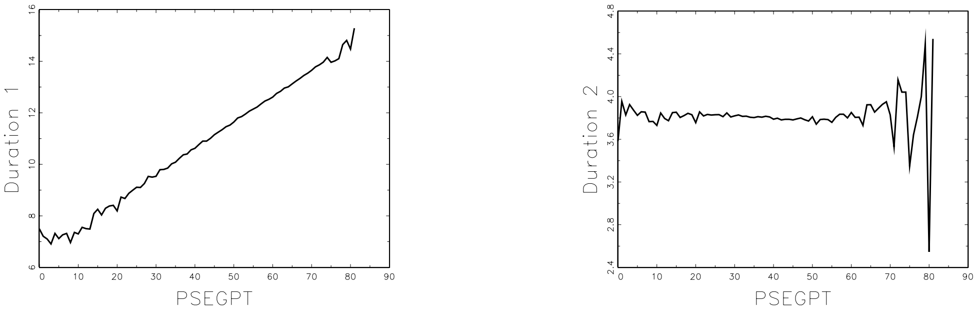
Figure 12: Explanatory Variable: PSEGPT, Stratification by Health Insurance



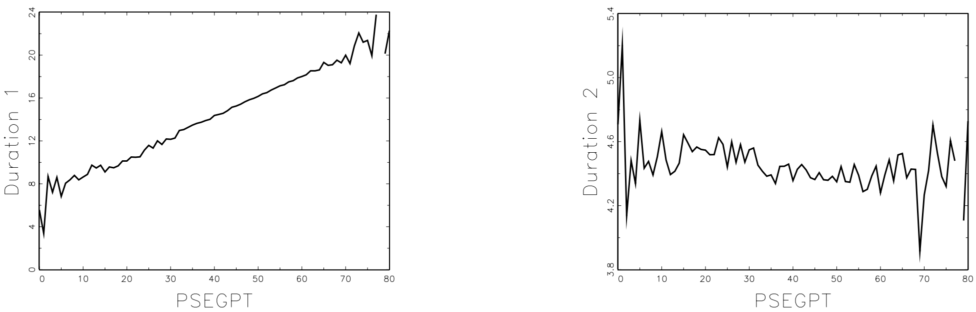
Left panel: duration 1. Right panel: duration 2. Solid: AUAZ = 0. Dashed: $0 <$ AUAZ ≤ 6. Dotted: AUAZ > 6. Male observations only.



Only AUAZ = 0, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

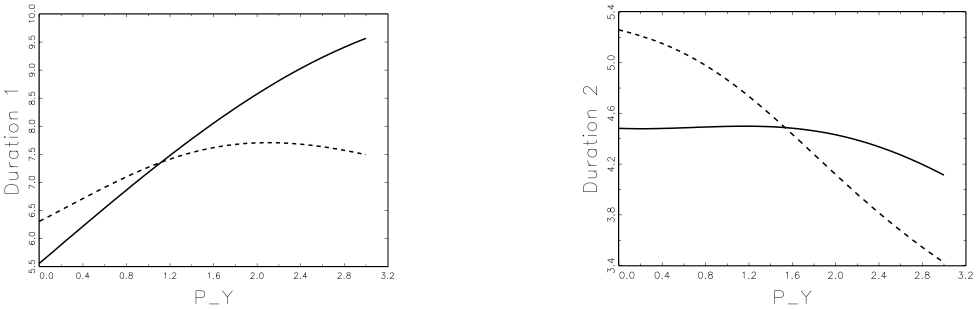


Only $0 <$ AUAZ ≤ 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

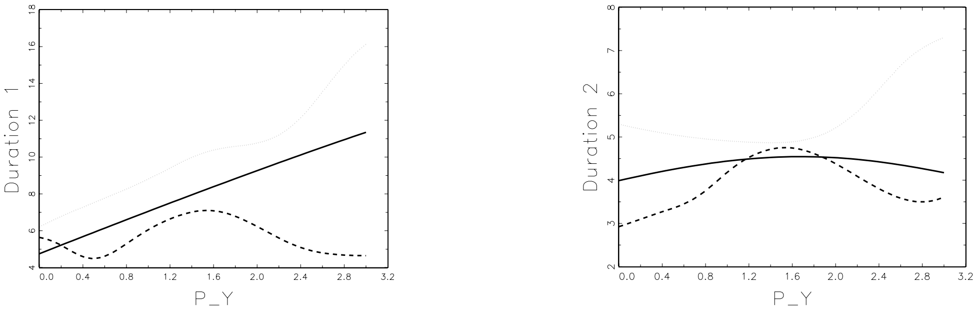


Only AUAZ > 6, left panel: duration 1. Right panel: duration 2. Male obs. only. Partially linear, PSEGPT instrumented.

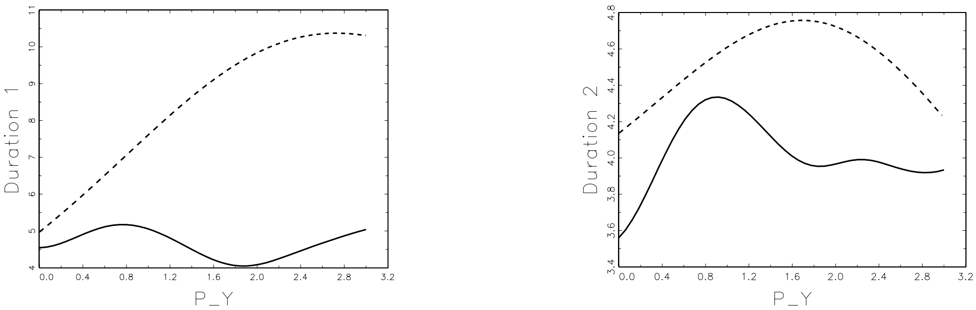
Figure 13: Explanatory Variable: PSEGPT, Stratification by Months in Ill-Health



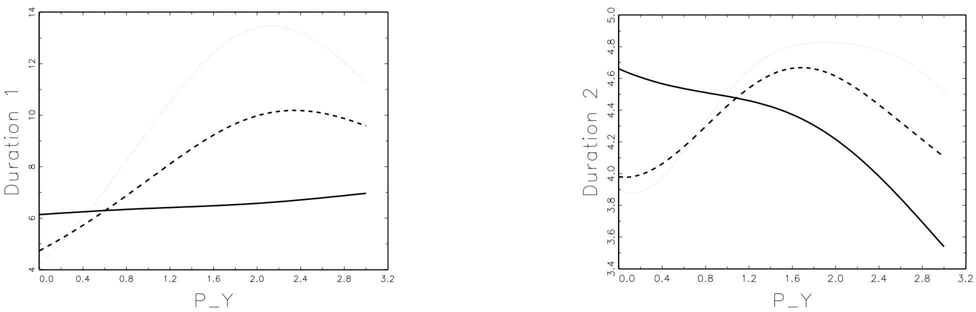
By sex, left panel: duration 1. Right panel: duration 2. Solid: male. Dashed: female.



By residence, left panel: duration 1. Right panel: duration 2. Solid: west. Dashed: east. Dotted: abroad. Male observations only.

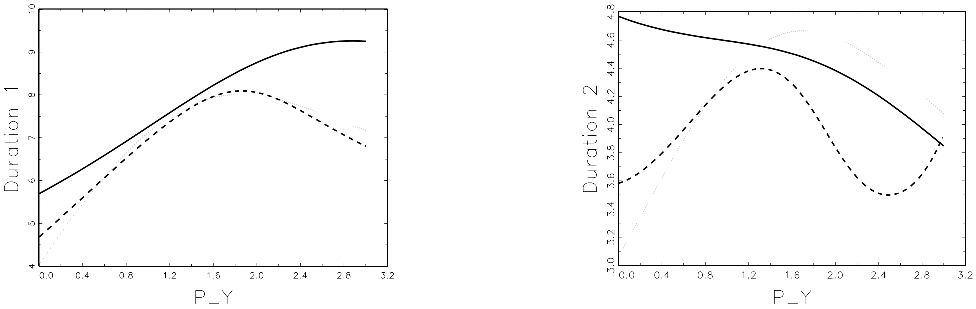


By health insurance, right panel: duration 1. Left panel: duration 2. Solid: private. Dashed: public.

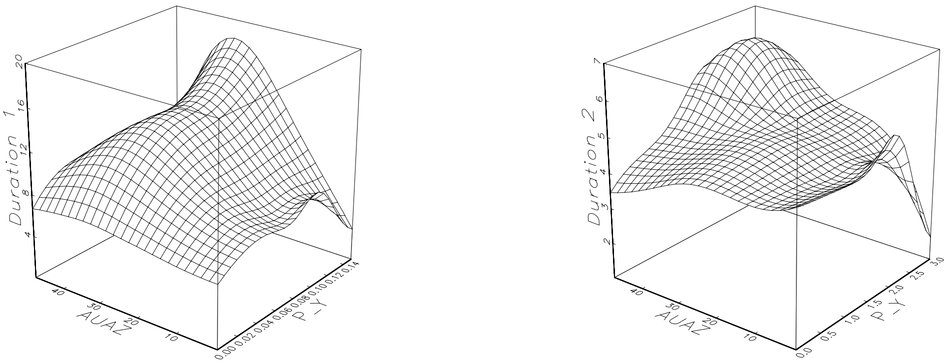


By months in illness, right panel: duration 1. Left panel: duration 2. Solid: no time in illness. Dashed: up to 6 months in illness. Dotted: More than 6 months in illness

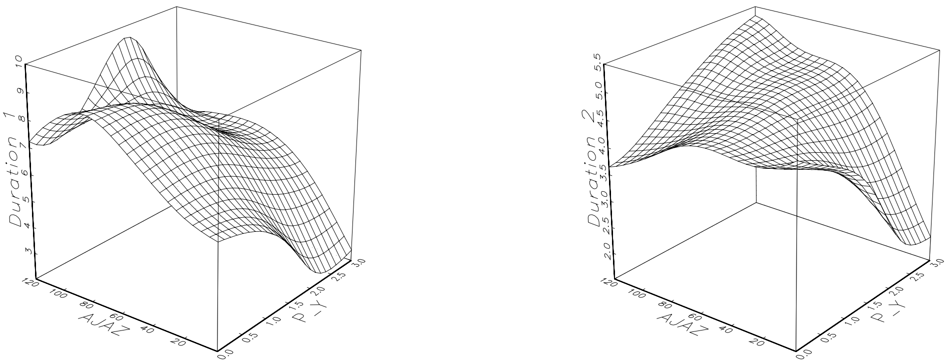
Figure 14: Explanatory Variable: PSEGPT/BYVL



By months in unemployment, right panel: duration 1. Left panel: duration 2. Solid: no time in unemployment. Dashed: up to 6 months in unemployment. Dotted: More than 6 months in unemployment



By time in illness, left panel: duration 1. Right panel: duration 2. Male observations only.



By time in unemployment, left panel: duration 1. Right panel: duration 2. Male observations only.

Figure 15: Explanatory Variable: PSEGPT/BYVL

References

Adams, P., M. D. Hurd, D. MacFadden, A. Merrill, and T. Ribeiro (2004): "Healthy, wealthy and wise? Tests for direct causal paths between health and socioeconomic status," *Journal of Econometrics*, 112, 3–56.

Angrist, J., and V. Lavy (1999): "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *The Quarterly Journal of Economics*, 114(2), 533–575.

Berkel, B., and A. Börsch-Supan (2004): "Pension Reform in Germany: The Impact on Retirement Decisions," *FinanzArchiv*, 60(3), 393–421.

Bloom, D. E., D. Canning, and M. Moore (2005): "The Effect of Iprovements in Health and Longevity on Optimal Retirement and Saving," *Harvard Initiative of Global Health Working Paper Series*, 2.

Bloom, D. E., D. Canning, M. Moore, and Y. Song (2006): "The Effect of Subjective Survival Probabilities on Retirement and Wealth in the United States," *NBER Working Paper*, 12688.

Blundell, R., and A. Duncan (1998): "Kernel Regression in Empirical Microeconomics," *The Journal of Human Resources*, 33(1), 62–87.

Breyer, F., and M. Kifmann (2002): "Incentives to retire later – a solution to the social security crisis?," *Journal of Pension Economics and Finance*, 1(2), 111–130.

——— (2004): "The German Retirement Benefit Formula: Drawbacks and Alternatives," *FinanzArchiv*, 60, 63–82.

Butrica, B. A., R. W. Johnson, K. E. Smith, and C. E. Steuerle (2006): "The Implicit Tax on Work at Older Ages," *National Tax Journal*, LVIV(2), 211–234.

Cleveland, W. S. (1974): "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74(368), 829–836.

Deutsche Rentenversicherung Bund (2005): "Rentenversicherung in Zeitreihen," *DRV Schriften*, 22.

DIAMOND, P. A. (2003): *Taxation, Incomplete Markets, and Social Security.* MIT Press, Cambrigde.

DIAMOND, P. A., AND J. A. MIRRLEES (1978): "A Model of Social Insurance with Variable Retirement," *Journal of Public Economics*, 10, 295–336.

FAN, J. (2000): "Prospects of Nonparametric Modeling," *Journal of the American Statistical Association*, 95(452), 1296–1300.

FELDSTEIN, M. S. (1974): "Social Security, Induced Retirement, and Aggregate Capital Accumulation," *Journal of Political Economy*, 82(5), 905–926.

FENGE, R., S. UEBELMESSER, AND M. WERDING (2006): "On the Optimum Timimg of Implicit Social Secrity Taxes Over the Life Cycle," *FinanzArchiv*, 62(1), 68–107.

GORSKI, M., T. KRIEGER, AND T. LANGE (2007): "Pensions, Education and Life Expectancy," *Working Paper*.

GROSSMAN, M. (1999): "The Human Capital Model of the Demand for Health," *NBER Working Paper*, 7078.

LAU, M. I., AND P. POUTVAARA (2006): "Social Security Incentives and Human Capital Investment," *Finnish Economic Papers*, 19(1), 16–24.

LAUTERBACH, K., M. LÜNGEN, B. STOLLENWERK, A. GERBER, AND G. KLEVER-DEICHERT (2006): "Zum Zusammenhang zwischen Einkommen und Lebenserwartung," *Studien zu Gesundheit, Medizin und Gesellschaft, Köln*, 01/2006.

LINDAHL, M. (2005): "Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income," *Journal of Human Resources*, 40(1), 144–168.

LIU, L., AND A. J. RETTENMAIER (2003): "Social Security Outcomes by Racial and Education Groups," *Southern Economic Journal*, 69(4), 842–864.

LOADER, C. (2004): "Smoothing: Local Regression Techniques," in *Handbook of Computational Statistics*, ed. by J. Gentle, W. Härdle, and Y. Mori. Springer.

Meer, J., D. L. Miller, and H. S. Rosen (2003): "Exploring the Health-Wealth Nexus," *Journal of Health Economics*, 22, 713–730.

Mittelhammer, R. C., G. G. Judge, and D. J. Miller (2000): *Econometric Foundations*. Cambridge University Press, Cambridge.

Schils, T. (2005): *Early Retirement Patterns in Europe: A comparative Panel Study*. Dutch University Press, Amsterdam.

Sheshinski, E. (1978): "A Model of Social Security and Retirement Decisions," *Journal of Public Economics*, 10, 337–360.

Speckman, P. (1988): "Kernel Smoothing in Partial Linear Models," *Journal of the Royal Statistical Society*, Series B, 50(3), 413–436.

von Gaudecker, H.-M., and R. D. Scholz (2006): "Lifetime Earnings and Life Expectancy," *MEA Discussion Paper*, 102-06.

Yatchew, A. (2003): *Semiparametric Regression for the Applied Econometrician*. Cambridge University Press, Cambridge.