UNIVERSITY OF COPENHAGEN

Essays on Labor Market Search and the Distribution of Incomes

Le Maire, Christian Daniel

Publication date:
2008

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Le Maire, C. D. (2008). *Essays on Labor Market Search and the Distribution of Incomes*. Department of Economics, University of Copenhagen. PhD Series, No. 131

Essays on Labor Market Search and the Distribution of Incomes

Daniel le Maire

Ph.D. Thesis

Department of Economics

University of Copenhagen

December 2008

Contents

1 Preface

The following five papers constitute my Ph.D. thesis. A lot of people have influenced my work and helped me during the three years. First of all, I wish to thank my advisor Karsten Albæk who convinced me to study Ph.D. and who has continuously been supporting me with lots of advises and guidance. Next, I wish to express my gratitude to fellow students. I began my BA studies in 1998 together with Christian Scheuer and Bertel Schjerning, and during my BA studies, master studies and Ph.D. studies I have been very lucky to work with both. Chapter 1 and 4 are coauthored with Christian Scheuer and chapter 2 with Bertel Schjerning.

The third paper is coauthored with Christian Dahl and Jakob Roland Munch. It has been very rewarding for me to work with Christian and discuss econometrics related to quantile regression. Furthermore, I have had the pleasure of working together with Jakob since 2001 where I was employed as student assistant in the Secretariat of the Danish Economic Council. I would also like to use this opportunity to thank my former colleagues in both the Danish Economic Council and the Danish Welfare Commission for shaping my thoughts on economics.

It has been rewarding as well as a pleasure to be part of the Centre of Applied Microeconometrics, and I wish to thank each and everyone who has been part of the centre. However, I wish to single out a few: I am extremely grateful for the all the help I have received from Mette Ejrnæs. Numerous times Mette has taken the time to discuss issues of econometrics both in relation to my papers and in relation to my teaching and has been most helpful with practical issues.

Furthermore, I am very grateful to Bo Honoré, Martin Browning, Peter Sørensen and Søren Leth-Petersen for detailed comments to several of my papers. Finally, I wish to thank John Kennes for many inspiring discussions on search theory.

As part of my Ph.D. studies I studied 10 months at Northwestern University. This was a great experience. Different early versions of the fourth chapter was handed in both Chris Taber's labor class and Dale Mortensen's labor class. I am grateful for the comments from both although especially grateful for several discussions with Dale Mortensen regarding the development of the theoretical part of this paper. Furthermore, I would like to express my gratitude for the financial support from the Tuborg Foundation, the Denmark-America Foundation, the Rudolph Als Foundation, and the Sasakawa Foundation, which made my stay at Northwestern possible.

Finally, I wish to thank my family and my girl friend Louise Rathsach Skouby for support during my three years of study.

Daniel le Maire
Copenhagen, August 2008.

2 Introduction and Summary

My thesis consists of five papers which are not all closely related. However, interpreted broadly all deal with the dispersion of earnings. In the first two papers the relationship between earnings and participation in the labor market and occupational choice is analyzed. In the third paper, the effects of the decentralization of wage bargaining on the wage distribution is examined, whereas the final two papers deal with assortative matching, which also has an effect on the earnings distribution.

The first paper (coauthored by Christian Scheuer) was originally conducted for the Danish Welfare Commission, and it seeks to examine the effects of financial incentives for persons receiving social assistance benefits. First, we use the within-individual variation in the recipients' spouse incomes to examine the employment effects of financial incentives. From this variation we find no employment effects. This result is consistent with consumption smoothing, but it could also be due to the workers receiving social assistance benefits being marginalized. Among other things, the wage distribution and the strong state dependence in their labor market status suggest that this group is marginalized. Second, we compute the net potential gain of working over receiving social assistance. Using this measure we find participation elasticities of 0.28-0.44 for men and in the range of 0.62-0.68 for women. However, it is likely that unobserved heterogeneity implies that these estimated elasticities are upward biased.

In the second paper (coauthored by Bertel Schjerning) we examine the importance of monetary incentives for the choice of becoming self-employed. First, we find that self-employed taxable income bunch at kink points in the tax system because self-employed can retain earnings and thereby transfer income across tax-years. Second, we find that both expected income level and income variance are important determinants in choice of occupation. Comparing men and women, we find that men put more emphasis on expected earnings level, while women appears more risk averse which contribute to explaining why fewer women are self-employed. Finally, our results suggest that non-western immigrants are marginalized into self-employment.

The third paper (coauthored by Christian M. Dahl and Jakob Roland Munch) studies how decentralization of wage bargaining from sector to firm level influences wage levels and wage dispersion. Existing studies rely on cross-section data and it is possible that the endogeneity of the wage bargaining arrangements, for example due to sorting, is not appropriately controlled for. We have the advantage of using a panel data set which covers part of the period of decentralization which facilitates identification of the effects of decentralization from changes that we claim are exogenous. Intuitively, when wages are negotiated at the local level we would expect a larger wage dispersion because firm- and individual-specific characteristics are more likely to enter the wage contracts. Consistent with this prediction we find that wages are more dispersed under firm-level bargaining compared to more centralized wage-setting systems. Furthermore, if local bargaining is known to imply more dispersed wages, we would expect that high ability workers sort into decentralized bargaining segments. Hence, we expect a positive correlation between local bargaining and unobserved ability and that appropriately controlling for unobserved heterogeneity should imply smaller estimated effects of different bargaining arrangements. This is also confirmed by our empirical results as the differences across wage-setting systems are reduced substantially when controlling for unobserved individual heterogeneity.

While the first three papers are empirical, the fourth paper (coauthored with Christian Scheuer) is both theoretical and empirical. Intuitively, one should think that high productive workers would choose to work in high productive firms, or interpreting it more general, that high wage workers would tend to work in high-wage firms. This is termed positive assortative matching. However, the existing empirical evidence suggests the opposite, that is negative assortative matching. In order to analyze this counter-intuitive result, we need a (search) model with i) worker flows between state of unemployment and different jobs, ii) positive assortative matching, and iii) a resulting log wage equation, which is additive-separable in worker and firm productivity. A central problem is that for assortative matching to arise, the canonical search model needs log-supermodularity of the production function, which implies that the wage equation is not additive separable. Our solution is to let workers determine how many jobs they apply to since then we can use a supermodular production function which delivers the wage equation needed. Besides continuous heterogeneity on both the worker and firm sides we add a match productivity effect. We show that when such match effect is present we obtain an estimated negative correlation between the worker and firm effects even though there is positive assortative matching (positive correlation) in the theoretical model. Furthermore, we find from empirical estimates on Danish matched employer-employee data that a match effect in the wages accounts for 15 per cent of the variation in the log wages. Finally, we find evidence of positive assortative matching in the Danish labor market.

The fifth and final paper builds on the previous paper by suggesting another way of achieving assortative matching in a search model when the production function is only supermodular. Specifically, we do not need workers to choose an optimal sample of firms to apply to. Instead, I show in a search model of the marriage market that we just need that both men and women to consider more than one candidate at a time.

3 Summary in Danish

Denne afhandling består af fem papirer. I det første papir (skrevet med Christian Scheuer) ønsker vi at undersøge effekten af økonomiske incitamenter på personer, der modtager kontanthjælp. Først undersøger vi, om der er effekter på beskæftigelsen, når partnerens indkomst varrierer over tid. Vi finder ingen deltagelseseffekter, men det kan skyldes, at individerne udglatter forbrug over tid, og muligvis også at kontanthjælpsmodtagere er marginaliseret og dermed ikke let kan finde et job. Lønfordelingen for den gruppe, vi betragter, hvor en stor del modtager løn lige over minimumslønnen, og den høje grad af persistens i modtagelsen af kontanthjælp, tyder netop på, at kontanthjælpsmodtagere er marginaliserede. For hver person udregner vi desuden den potentielle forskel i indkomst, når personen arbejder i forhold til, når personen modtager kontanthjælp. Ved at benytte denne forventede indkomstændring finder vi, at deltagelseselaticiteten er 0,28-0,44 for mænd og 0,62-0,68 for kvinder. Imidlertid argumenterer vi for, at det er sandsynligt, at uobserverbar heterogenitet betyder, at disse estimater er overvurderede.

I det andet papir (skrevet med Bertel Schjerning) undersøger vi, hvor stor en betydning økonomiske incitamenter har for valget om at blive selvstændig. Først og fremmest finder vi, at selvstændiges skattepligtige indkomst tenderer til at ligge lige under indkomstgrænsen for mellemskat og topskat, idet selvstændige kan overføre overskudet i virksomheden til et efterfølgende skatteår. For det andet finder vi, at både det forventede indkomstniveau og indkomstvariationen er vigtige parametre for valget om at blive selvstændig. Sammenligner vi mænd og kvinder, finder vi, at mænd lægger større vægt på indkomstniveauet, mens kvinder er mere risikoaverse og dermed foretrækker mindre indkomst variation, hvilket kan være med til at forklare, hvorfor færre kvinder bliver selvstændige. Tilslut viser vores resultater, at ikke-vestlige indvandrere maginaliseres til at blive selvstændige.

Det tredje papir (skrevet med Christian M. Dahl og Jakob Roland Munch) undersøger, hvorledes decentraliseringen af løndannelsen fra brancheniveau til virksomhedsniveau har påvirket lønniveauet og lønspredningen. Eksisterende studier benytter udelukkende data for et enkelt år, og det er dermed muligt at endogeneiteten i overenskomstområdet ikke afhjælpes. Vi benytter derimod et panel datasæt, der dækker en del af decentraliseringsperioden, hvormed vi kan benytte de eksogene skift i lønsystemet, som decentraliseringen har medført. Intuitivt må en mere decentral løndannelse betyde større lønspredning, da individuelle såvel som virksomhedsspecifikke karakteristika har større mulighed for at influere lønnen. Vi finder i overenstemmelse med denne forventning, at lønningerne har større spredning under decentral løndannelse. Hvis arbejderne ved dette, må vi forvente, at de mest produktive vælger at arbejde der, hvor der er størst lønspredning. Dermed er det forventelig, at der er en positiv korrelation mellem decentral forhandling og uobserverbar heterogenitet, og at effekterne af løndannelsen bliver mindre, når vi kontrollerer for den uobserverbare heterogenitet. Denne forventning bliver bekræftet af vores resultater, hvor effekten af løndannelsen bliver væsentlig mindre, når vi kontrollerer for uobserverbar heterogenitet.

Det fjerde papir (skrevet med Christian Scheuer) præsenterer en ny søgeteorimodel, hvor heterogenene arbejdere og virksomheder matcher med hinanden. Vi tillader, at arbejderne bestemmer, hvor mange jobs de ønsker at søge, hvilket medfører, at mere produktive arbejdere og mere produktive virksomheder har større sandsynlighed for at finde et match sammen (også kaldet positiv assortativ matching), på trods af at produktionsfunktionen kun er supermodulær. Foruden kontinuert heterogenitet på både arbejder- og virksomhedsside tilføjer vi et match produktivitetsled. Vi viser, at hvis der findes en sådan match effekt, kan vi estimere en negativ korrelation mellem arbejder- og virksomhedseffekterne, selv om der i virkeligheden er positiv assortativ matching (positiv korrelation) i den teoretiske model. Desuden finder vi på dansk mikro data, at match effekten udgør 15 procent af variationen i logaritmen til timelønnen. Tilslut finder vi også empirisk, at der er positiv assortativ matching på det danske arbejdsmarked.

Det femte og sidste papir bygger på det foregående papir, idet det foreslår en anden måde at opnå assortativ matching på, selv om produktionsfunktionen kun er supermodulær. I stedet for at lade arbejderen selv vælge det optimale antal job tilbud, han vil søge, viser vi i en søgemodel for ægteskabsmarkedet, at man bare behøver, at mænd og kvinder overvejer mere end en potentiel partner ad gangen.

Determinants of Labor Force Participation for Recipients of Social Assistance: A Panel Data Analysis for Denmark*

Daniel le Maire
University of Copenhagen and CAM

Christian Scheuer
Copenhagen Business School and CEBR

December 15, 2008

Abstract

In this paper we seek to examine the effects of economic incentives for recipients of social assistance. The workers receiving social assistance benefits are found to be marginalized and we find strong state dependence in their labor market status. We find no employment effects from the within-individual variation in the recipients' spouse incomes, and in contrast to Hyslop (1999) and Croda and Kyriazidou (2002) we argue that for the case of Denmark the spouse income is endogenous, so that we cannot use the between variation of spouse income to obtain meaningful estimates. Therefore, we also compute the net potential gain of working over receiving social assistance. Using this measure we find participation elasticities of 0.28-0.44 for men and in the range of 0.62-0.68 for women. Transforming the elasticities according to definition in the CGE-model DREAM, that is the percentage change in the number of recipients from a percentage change in the income, these amount to 0.09-0.17 and 0.13-0.20 for respectively men and women. However, we argue that it is likely that unobserved heterogeneity implies that both sets of estimated elasticities are upward biased.

Keywords: Labor force particpation, incentives, nonemployment, state dependence, heterogeneity.

JEL Classification: C23, C25, J21, J24.

*We gratefully acknowledge the comments we have received from Karsten Albæk, Martin Browning, Mette Ejrnæs, Jan Vognsen Hansen, Martin Ulrik Jensen, Henrik Jacobsen Kleven, Hans Christian Kongsted, Claus Thustrup Kreiner, Tove Birgitte Pedersen, Søren Leth-Petersen, Bertel Schjerning, Esben Anton Schultz, Anders Sørensen and seminar participants in the DGPE 2005 workshop. Finally, we thank Mikael Kirk for research assistance. All remaining errors are ours.

1 Introduction

In this paper we consider intertemporal labor supply and examine the effects of economic incentives on the labor market participation for recipients of social assistance.

The labor supply decision is usually decomposed into the intensive margin, that is the choice of hours, and the extensive margin, that is the participation decision. Even though it is generally believed (see e.g. Heckman (1993)) that the largest effects are to be found on the extensive margin, most empirical studies have focused on the intensive margin. Empirical studies for Denmark have typically found numerical small elasticities from wages on the amount of labor supplied; see e.g. Frederiksen et al. (2001). In fact, the studies that have found largest elasticities have used a Tobit framework and, hence, it seems as if the small elasticities at the intensive margin have been polluted by the jointly modeling of the participation decision.

In Denmark several types of benefits exist. Each of them is more or less directed to a distinct group of people. With respect to participation, recipients of sickness benefits and disablement benefits are for obvious reasons less interesting. Therefore, we restrict our attention to the recipients of social assistance – that is persons that are not eligible to unemployment benefits either because they are not members of an unemployment insurance fund or since they have not been in work recently.

We have access to a rich panel data set, and we present results from employing different panel data estimators. The first set of estimations follow Hyslop (1999) and Croda and Kyriazidou (2003) closely in the sense that we estimate the effect on the participation decision using spouse disposable income separated into permanent and temporary income.

Focusing solely on married women Hyslop (1999) and Croda and Kyriazidou (2003) argue that spouse income is exogenous due to differences in the participation pattern between men and women. Hyslop (1999) finds elasticities of respectively permanent and transitory spouse income of -0.2 and -0.04 for the US, while Croda and Kyriazidou (2003) for Germany find only very small effects from both permanent and transitory spouse income on the participation decision. However, in the case of Denmark the difference in participation between men and women is less pronounced (see e.g. Dex et al. (1995) for a cross-national comparison) and we should expect a similar picture to emerge for men and women among recipients of social assistance. Consequently, we also perform the estimations for men.

Instead of solely restricting our attention to the effect of spouse income on participation, we also estimate a set of models where we analyze the effect of the worker's own disposable income gap from working on the participation decision To compute these income gaps we make panel data selectivity predictions of the own wage income for each person. From this we can calculate the disposable income from working full-time and compare this to the disposable income when receiving social assistance.

The paper is organized as follows. In section 2, we present a theoretical labor market search model. Section 3 outlines the data used in the analysis, and section 4 the econometric methodology. In section 5 and 6, we examine the results when using respectively disposable spouse income and own predicted disposable income gaps. Section 7 concludes.

2 Theoretical Background

An important aspect of intertemporal labor supply is the persistence of labor market status. In this section we will briefly consider a stylized labor market search model, which can generate such pattern.

We extend a standard partial equilibrium search model with a stigmatizing effect of becoming nonemployed and by assuming an instant utility function, which is non-linear in income. The first assumption will similarly to Garibaldi and Wasmer (2004) imply state dependence, whereas the latter assumption will allow us to focus on the effect of the spouse's income, which we take to be exogenous to the worker in consideration.

Consider an indefinitely living worker, who is married and whose spouse's labor market status is exogenous. We assume a joint instant utility function $u(\cdot)$, where the worker's own labor income w and her spouse labor income s are perfect substitutes. Being employed or nonemployed, wage offers are distributed according to $F(w)$ and arrive with the exogenous Poisson rate λ. Workers who decline the new wage offer become non-employed.

When an employed worker becomes nonemployed the worker experiences a one-time stigmatization effect γ. When $h_{t-1} = \{0, 1\}$ denotes the labor market state of the previous period the value of being nonemployed $U(h_{t-1})$ is given by

$$rU(h_{t-1}) = u(s + b - \gamma h_{t-1}) + \lambda \left[\int_{-\infty}^{\infty} \max(W(x), U(0))\, dF(x) - U(h_{t-1}) \right] \tag{1}$$

where r is the discount rate and $W(w)$ is the value of being employed at wage w. The Bellman equation for an employed worker is

$$rW(w) = u(s + w) + \lambda \left[\int_{-\infty}^{\infty} \max(W(x), U(1))\, dF(x) - W(w) \right] \tag{2}$$

The reservation wage for an nonemployed R^U is the wage where the worker is indifferent between being unemployed and working, that is $U(0) = W\left(R^U\right)$

$$u(s + b) - u\left(s + R^U\right) = \lambda \int_{-\infty}^{\infty} \max(W(x), U(1))\, dF(x) - \lambda \int_{-\infty}^{\infty} \max(W(x), U(0))\, dF(x) \tag{3}$$

Similarly, the reservation wage for an employed R^W is when $U(1) = W\left(R^W\right)$

$$u(s + b - \gamma) - u\left(s + R^W\right) = \lambda \int_{-\infty}^{\infty} \max(W(x), U(1))\, dF(x) - \lambda \int_{-\infty}^{\infty} \max(W(x), U(0))\, dF(x) \tag{4}$$

Equating equations (3) and (4) gives us

$$u\left(s+R^{U}\right)-u\left(s+R^{W}\right)=u\left(s+b\right)-u\left(s+b-\gamma\right)$$

and performing a first-order Taylor series approximation in respectively $s+R^U$ and $s+b$ gives us

$$\begin{aligned} u'\left(s+R^{U}\right)\left(R^{W}-R^{U}\right) &= -u'\left(s+b\right)\gamma \Leftrightarrow \\ R^{W} &= R^{U}-\frac{u'\left(s+b\right)}{u'\left(s+R^{U}\right)}\gamma \end{aligned} \tag{5}$$

where it is obvious that $R^W < R^U$ which implies that employed persons have a higher probability of being employed in the subsequent time period.

Re-arranging equation (4) and integrating by parts gives us

$$u\left(s+b-\gamma\right)-u\left(s+R^{W}\right)=-\frac{\lambda}{r+\lambda}\int_{R^{W}}^{R^{U}}u'\left(s+x\right)F\left(x\right)dx$$

and by use of a Taylor series expansion in the point $s+R^W$ the reservation wage for a nonemployed is

$$R^{w}=b-\gamma+\frac{\lambda}{r+\lambda}\int_{R^{W}}^{R^{U}}\frac{u'\left(s+x\right)}{u'\left(s+R^{W}\right)}F\left(x\right)dx \tag{6}$$

With a strictly concave utility function the fraction inside the integral is always less than 1, but will be increasing towards 1 as the spouse's income increases and, hence, imply a higher reservation wage R^W. Moreover, since $R^U > b$ the difference between the two reservation wages R^U and R^W is declining in the spouse income.

In the present model state dependence is a result of the stigmatization effect, but also loss of skill by becoming non-employed can imply state dependence. Furthermore, in Garibaldi and Wasmer (2004) state dependence arises when search costs are higher for nonemployed workers and the value of home market production is stochastic. In addition to this, state dependence can arise if employers use nonemployment as a signal of low productivity. Finally, Hyslop (1999) argues that state dependence also can arise if the marginal utility of consumption is greater when working.

3 Data

We have access to an unbalanced panel data set for 1998-2003.[1] The data set is a representative 10 per cent sample of the Danish population. The variables originate from five databases. The first four databases, the Income Registry, the IDA database, the Housing Registry and the Health Insurance Registry are all maintained by Statistics Denmark. The fifth database is the

[1] In fact, we also have information for 1997, but these are only used to construct the lagged dependent variable used when estimating dynamic discrete choice models.

DREAM database of the Danish Ministry of Employment.

Our interest lies in the group of people in contact with the social assistance system. Social assistance benefits are lower than unemployment benefits and the recipients are persons that are not eligible to unemployment benefits either since they are not insured in an unemployment insurance fund or since they had not recently been in work for a sufficient long period of time. Hence, for most workers in the Danish labor market the relevant alternative to not work is unemployment benefits. Therefore, the characteristics of recipients of social assistance differ remarkably from workers in the labor force, and in this paper we only include persons that in at least one of the years covered, 1998-2003, have primarily been receiving social assistance. We define a recipient of social assistance as a person who has received social assistance benefits in the majority of the year, that is, at least 27 weeks in a given calendar year according to the DREAM Registry, which contains information on all transfers of public benefits on a weekly basis. For these persons we include information for all years available where the individual was either primarily receiving social assistance or primarily was in the labor force. In other words, persons who have been employed in all years are not included in our sample.

Since we are going to model the Danish tax system and compute the taxes paid in various scenarios, we further exclude self-employed, assisting wives, and people not fully taxable in Denmark in a given year from the sample. In addition to this, we have chosen not to include persons observed only once, since they cannot be used in the majority of our estimation procedures.

In Table 1 and 2 we present some characteristics of our sample divided by gender. Hourly gross wage and the partner's disposable income are in constant (1997) Danish kroner. Each table consists of six columns. The first column contains mean values for our full sample, while the next four columns have information on different sub-samples according to the individuals' transition pattern between social assistance and work. The sixth and final column is included to facilitate the understanding of how the characteristics of our sample differ from the labor force. Here, we show mean values for a 33 per cent sample of all Danes who have been in the labor force for at least one year in the period of 1998-2003.

Table 1: Mean values for males

	Full sample	Social assistance all years	Single transition from social assistance	Single transition from work	Multiple Transitions	Population Mean
Experience	4.91	3.13	4.88	8.12	6.58	17.27
	(5.95)	(4.78)	(5.89)	(6.86)	(6.05)	(9.87)
Age	35.89	36.65	35.42	35.89	34.16	40.64
	(9.73)	(10.02)	(9.71)	(9.37)	(9.01)	(10.72)
Neuro medicine(*)	0.02	0.02	0.01	0.02	0.01	0.01
	(0.24)	(0.27)	(0.17)	(0.28)	(0.17)	(0.39)
Psychiatry(*)	0.15	0.22	0.07	0.15	0.05	0.03
	(1.38)	(1.70)	(0.90)	(1.32)	(0.69)	(0.64)
General medical treatment(*)	8.73	10.04	6.76	8.98	7.03	4.91
	(17.80)	(20.10)	(13.13)	(18.69)	(13.75)	(9.86)
Unskilled	0.70	0.73	0.67	0.66	0.69	0.29
	(0.46)	(0.44)	(0.47)	(0.48)	(0.46)	(0.45)
Vocational training	0.23	0.19	0.24	0.29	0.26	0.46
	(0.42)	(0.39)	(0.43)	(0.45)	(0.44)	(0.50)
Short-cycle higher education	0.02	0.03	0.02	0.02	0.02	0.06
	(0.15)	(0.16)	(0.15)	(0.13)	(0.14)	(0.23)
Medium-cycle higher education	0.03	0.03	0.04	0.02	0.02	0.11
	(0.17)	(0.17)	(0.20)	(0.14)	(0.14)	(0.31)
Long-cycle higher education	0.02	0.03	0.02	0.02	0.01	0.08
	(0.15)	(0.16)	(0.15)	(0.13)	(0.12)	(0.28)
Immigrant	0.28	0.35	0.35	0.14	0.14	0.04
	(0.45)	(0.48)	(0.48)	(0.34)	(0.35)	(0.20)
Second generation immigrant	0.01	0.01	0.01	0.01	0.01	0.00
	(0.09)	(0.09)	(0.08)	(0.08)	(0.10)	(0.06)
Married	0.25	0.29	0.32	0.16	0.17	0.56
	(0.44)	(0.45)	(0.47)	(0.37)	(0.37)	(0.50)
Copenhagen	0.22	0.23	0.22	0.18	0.19	0.10
	(0.41)	(0.42)	(0.41)	(0.39)	(0.39)	(0.30)
Large city	0.19	0.22	0.18	0.16	0.15	0.14
	(0.39)	(0.42)	(0.39)	(0.37)	(0.35)	(0.35)
Rural area	0.48	0.45	0.48	0.54	0.53	0.64
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.48)
Children aged 0-6 years	0.28	0.31	0.32	0.21	0.23	0.30
	(0.67)	(0.74)	(0.68)	(0.56)	(0.57)	(0.63)
Children aged 7-17 years	0.29	0.33	0.33	0.19	0.24	0.42
	(0.75)	(0.84)	(0.76)	(0.58)	(0.64)	(0.76)
Owner	0.42	0.39	0.44	0.45	0.45	0.68
	(0.49)	(0.49)	(0.50)	(0.50)	(0.50)	(0.47)
Hourly gross wage	112.11	-	113.35	110.50	112.07	154.09
	(47.47)		(45.68)	(49.95)	(47.19)	(71.83)
Partner's disposable income	97968	96101	100072	98251	99981	133770
	(32749)	(29083)	(35136)	(36272)	(35258)	(60918)
No. Years in sample	4.67	4.29	4.98	4.78	5.33	-
	(1.44)	(1.63)	(1.23)	(1.23)	(0.87)	
No. Years of Social Assistance						
1	25.32	7.48	48.52	37.18	35.21	
2	18.05	10.55	22.62	26.39	25.09	
3	14.69	13.03	14.85	17.16	16.65	
4	13.92	16.27	9.54	12.12	14.65	
5	12.81	20.05	4.46	7.15	8.40	
6	15.21	32.61	0.00	0.00	0.00	
Sample Size	41,390	19,298	8,341	7,469	6,282	573,498

Standard deviations are in parentheses.
(*) These variables originate from the Health Insurance Registry and are computed as the yearly number of treatments with support from the Danish health system within the given area.

Table 1 shows that the full sample of men consists of 41,390 observations. Looking at the number of years the individuals received social assistance; about 25 per cent received social assistance in just one year during 1998-2003 while about 15 per cent received social assistance in the full period.

From the sample means a few things are worth noticing. Firstly, even though the mean age is approximately 36 years the mean value of experience (in years) is only approximately 5 years. Secondly, 70 per cent are unskilled compared to 29 per cent in the active population. The combination of lower experience and lower levels of education suggest a high degree of persistence in the labor market status, since males receiving social assistance are clearly less likely to become employed compared to the average male in the active population.

Furthermore, as the average experience indicates our sample is quite different from the average population suggesting that the group under consideration has a weak attachment to the labor market. The most prominent differences are that 28 per cent of the sample of men are immigrants compared to 4 per cent in the active Danish population. In addition to this, the average wage is 112 DKK per hour while the mean wage for private employed males was 154 DKK. Finally, only 42 per cent are homeowners (68 per cent is the active population)

We proceed by splitting the full sample into 4 subgroups; those receiving social assistance all years (19,298 persons), those who have a single transition from social assistance to work (8,341 persons), those having a single transition from work to social assistance (7,469 persons) and those who have more than one transition between work and social assistance (6,282 persons).

When breaking the sample down by transition patterns we notice that even though the labor market experience is quite different between the three groups of people who at some point have been on the labor market (4.9-8.1 years) the average wage for the people working is much alike (111-113 DKK) suggesting that the minimum wage restriction is binding for the majority of persons belonging to the group. Further, immigrants are much overrepresented among those receiving social assistance in all years and among those having only one single transition from social assistance (35 per cent compared to 14 per cent). This may explain why a larger fraction of these groups is married and why the groups on average have more children.

There is also a striking difference between the spouse's average disposable income as it is much higher for the active population. In particular, the partner income is lowest for those who are receiving social assistance in all years.

Table 2: Mean values for females

	Full sample	Social assistance all years	Single transition from social assistance	Single transition from work	Multiple Transitions	Population Mean
Experience	2.93	1.82	3.34	6.28	5.05	15.13
	(4.60)	(3.68)	(4.57)	(5.80)	(5.39)	(8.73)
Age	34.36	34.62	33.90	34.48	33.58	41.26
	(9.42)	(9.69)	(9.02)	(9.12)	(8.75)	(10.39)
Neuro medicine(*)	0.04	0.04	0.02	0.03	0.03	0.02
	(0.37)	(0.39)	(0.32)	(0.33)	(0.40)	(0.26)
Psychiatry(*)	0.22	0.26	0.13	0.29	0.09	0.07
	(1.86)	(2.04)	(1.24)	(2.28)	(0.96)	(1.21)
General medical treatment(*)	17.65	19.49	13.66	17.32	15.11	9.38
	(31.19)	(34.27)	(16.08)	(37.43)	(26.12)	(11.37)
Unskilled	0.78	0.83	0.72	0.68	0.71	0.29
	(0.42)	(0.38)	(0.45)	(0.47)	(0.45)	(0.45)
Vocational training	0.16	0.11	0.22	0.27	0.24	0.39
	(0.37)	(0.32)	(0.41)	(0.44)	(0.42)	(0.49)
Short-cycle higher education	0.02	0.02	0.02	0.01	0.01	0.05
	(0.13)	(0.13)	(0.15)	(0.10)	(0.08)	(0.22)
Medium-cycle higher education	0.03	0.03	0.03	0.03	0.03	0.20
	(0.17)	(0.17)	(0.18)	(0.18)	(0.16)	(0.40)
Long-cycle higher education	0.01	0.01	0.01	0.01	0.02	0.06
	(0.10)	(0.10)	(0.10)	(0.11)	(0.14)	(0.24)
Immigrant	0.32	0.40	0.30	0.11	0.12	0.03
	(0.47)	(0.49)	(0.46)	(0.32)	(0.33)	(0.18)
Second generation immigrant	0.01	0.01	0.01	0.01	0.00	0.00
	(0.08)	(0.08)	(0.10)	(0.10)	(0.05)	(0.05)
Married	0.32	0.37	0.33	0.19	0.18	0.62
	(0.47)	(0.48)	(0.47)	(0.39)	(0.39)	(0.49)
Copenhagen	0.18	0.18	0.15	0.18	0.18	0.10
	(0.38)	(0.39)	(0.36)	(0.38)	(0.38)	(0.31)
Large city	0.18	0.21	0.16	0.13	0.13	0.14
	(0.39)	(0.41)	(0.37)	(0.33)	(0.33)	(0.35)
Rural area	0.53	0.51	0.55	0.58	0.56	0.63
	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	(0.48)
Children aged 0-6 years	0.63	0.72	0.53	0.44	0.49	0.30
	(0.86)	(0.93)	(0.74)	(0.69)	(0.74)	(0.61)
Children aged 7-17 years	0.69	0.73	0.69	0.51	0.61	0.50
	(1.02)	(1.09)	(0.94)	(0.83)	(0.87)	(0.80)
Owner	0.28	0.23	0.34	0.38	0.36	0.71
	(0.45)	(0.42)	(0.47)	(0.48)	(0.48)	(0.45)
Hourly gross wage	100.99	-	100.59	100.60	102.53	122.83
	(41.81)		(36.82)	(44.72)	(49.80)	(47.46)
Partner's disposable income	108228	99405	123651	116514	121150	173982
	(43080)	(38738)	(44089)	(49192)	(45992)	(121702)
No. Years in sample	4.62	4.45	4.87	4.57	5.19	-
	(1.46)	(1.60)	(1.25)	(1.27)	(0.91)	
No. Years of Social Assistance						
1	19.91	6.44	42.41	37.41	34.08	
2	14.78	9.19	22.57	25.14	20.17	
3	14.19	11.37	18.69	16.13	20.11	
4	14.49	15.88	10.79	14.28	14.30	
5	14.57	19.73	5.54	7.04	11.33	
6	22.06	37.40	0.00	0.00	0.00	
Sample Size	40,170	23,698	8,450	4,686	3,336	532,727

Standard deviations are in parentheses.
(*) These variables originate from the Health Insurance Registry and are computed as the yearly number of treatments with support from the Danish health system within the given area.

The corresponding mean values for women are shown in Table 2. The pattern of Table 2 does to a large extent replicate the pattern in Table 1. Therefore, the most important information we retrieve is that also female recipients of social assistance seem to have a weak attachment to the labor market. The mean age is 34 years and the mean experience is only 2.9 years, 32 per cent are immigrants and 78 per cent unskilled. Again breaking down on different transition patterns we see that differences in experience are not reflected in the wage rate of those employed from the different groups, again suggesting the minimum wage restriction to be binding. Furthermore, the partner income is now somewhat lower for the group receiving social assistance all years. This may suggest that to some extent persons cohabiting are likely to have a similar attitude towards participating in the labor market, and that this effect might dominate the positive income effect exerted by partner earnings as modeled section 2.

4 The Econometric Framework

The estimations dealt with in this paper are reduced-form participation equations. There are different possible sources of the persistence in the individual labor market participation. Here, we study the importance of unobserved heterogeneity and state dependence in generating this persistence. We do not consider serial correlation in the time-varying error component, which similar to state dependence implies that transitory changes in the explanatory variable may have persistent effects on the dependent variable. However, in a similar analysis for Germany Croda and Kyriazidou (2003) find that serial correlation in the time-varying error component does not seem to matter. Hence, we disregarded such error components.

We estimate a number of binary response models. In the following we only shortly consider the econometric specifications, and refer to references mentioned below as well as textbooks such as Greene (2003) and Wooldridge (2002) for a more detailed treatment.

The first estimation considered is the pooled probit model

$$\Pr\left(y_{it}=1|X_{it}\right)=1\left(X_{it}\beta+\varepsilon_{it}\geq 0\right)=\Phi\left(X_{it}\beta\right) \tag{7}$$

where y_{it} is the binary variable for the participation decision of individual i at time t, X_{it} contains the explanatory variables, and the error term ε_{it} is assumed to be independent of the explanatory variables and distributed $iN\left(0,1\right)$.

The pooled probit estimator does not exploit the fact that we observe the persons again and again, and since it is very likely that the errors are correlated because the same persons are observed several times, we also estimate a random effects probit

$$\Pr\left(y_{it}=1|X_{i},\alpha_{i}\right)=1\left(X_{it}\beta+\alpha_{i}+\varepsilon_{it}\geq 0\right)=\Phi\left(X_{it}\beta+\alpha_{i}\right) \tag{8}$$

where the compound error-term $\alpha_i+\varepsilon_{it}$ is assumed to be independent of the explanatory variables and both terms are assumed to be normally distributed.

Absent state dependence transitory changes in X can only cause transitory changes in the dependent variable y. Allowing for state dependence is equivalent to allowing transitory changes in X to have permanent effect on y through the effects of the lagged dependent variable. In the absence of serial correlation this is the only way in which transitory changes in X can have persistent effects on y. Failing to allow for state dependence will bias the parameter estimates in presence of (true) state dependence. In addition, it is of interest whether employment is truly state-dependent, since this implies that becoming a recipient of social assistance deteriorates an individual's future labor market prospects. In order to address this issue, we estimate a dynamic random effects probit

$$\Pr\left(y_{it}=1|X_i,\alpha_i,y_{it-1}\right)=1\left(\gamma y_{it-1}+X_{it}\beta+\alpha_i+\varepsilon_{it}\geq 0\right)=\Phi\left(\gamma y_{it-1}+X_{it}\beta+\alpha_i\right) \quad (9)$$

Estimating a dynamic probit raises the question of how to treat the initial observations of the dependent variable y_{i0}. Heckman (1981) suggests approximating the conditional density of the initial dependent variable by estimating a probit using observations from the first year only and simultaneously specifying the unobserved heterogeneity conditional on the explanatory variables. We use the simpler estimation procedure for the dynamic correlated random effects probit outlined in Wooldridge (2005) where the approximation of the density of the initial dependent variable is left out. Instead the unobserved heterogeneity is allowed to be arbitrarily correlated with the initial dependent variable by inclusion of the initial observation y_{i0}. Furthermore, we model the unobservable individual component with the means of the time-varying explanatory variables as in Mundlak (1978) in order to be able to better interpret the effects of the time-varying variables. This implies that we assume that

$$\alpha_i=y_{i0}\rho+\bar{X}_i\lambda+\xi_i$$

where $\xi\sim N\left(0,\sigma_\xi^2\right)$.

We apply these estimators in two scenarios: First, we estimate the effect of changes in partner income and then we estimate the effect from changes in the worker's own potential disposable income gap from working. Besides these income variables X_{it} includes a quadratic term in age and experience, dummy indicators for level of education, living area, ethnic origin, variables for the number of children, variables for medical treatments as well as variables capturing the regional demand and supply conditions.

Hyslop argues that the spouse's income may have three different effects on the labor market participation, a) a direct effect as in the theoretical search model in section 2, b) an expectation effect as future spouse income increases or decreases may be anticipated, and c) a taste effect from the spouses having the same taste for work by assortative matching in the marriage market. In the presence of state dependence, the expectations of future outcomes of explanatory variables may affect the current labor market status. This we avoid by assuming that the partner income follows a stationary process. As in Hyslop (1999) stationarity is achieved by dividing spouse income into a permanent and a transitory component. The permanent income is just the average of the spouse's income in the period under consideration, while the transitory income is the yearly deviation from this mean. The taste for work effect can be expected to be in the individual unobservable component and to be correlated with the permanent spouse income, which makes the latter endogenous.

For transitory income on the other hand it is probably reasonable to assume no correlation with taste for work. If this assumption holds there will only be a direct income effect from transitory income. However, in a life-cycle model transitory income shocks should not have an effect on labor supply, since agents can smooth out consumption. Only when there exist credit market constraints transitory spouse income shocks will play a role for the labor supply.

We can take account of possible correlation between α_i and the time-varying explanatory variables by estimating correlated random effects and fixed effects estimators. Here we apply two fixed effects estimators which both use the logit specification and, hence, instead of assuming that the error-term ε_{it} is normally distributed we assume that it is logistically distributed. We need to observe a person in two periods when estimating the conditional maximum likelihood fixed effects logit. The idea is that only a person that changes state, that is for a two-period setting $y_{i1} + y_{i2} = 1$, contributes to the likelihood function. Therefore neither persons employed in all years or persons receiving social assistance in all years affect the likelihood function in the fixed effects logit. For $y_{i1} + y_{i2} = 1$ we have

$$\Pr\left(y_{i1} = 1 | X_i, \alpha_i, y_{i1} + y_{i2} = 1\right) = \frac{\exp\left(\left(X_{i2} - X_{i1}\right)\beta\right)}{1 + \exp\left(\left(X_{i2} - X_{i1}\right)\beta\right)} \tag{10}$$

As with the random effects probit model, naturally, we want to allow for state dependence. Chamberlain (1993) has shown that if individuals are only observed in three periods, the parameters of the dynamic fixed effects logit model are not identified. Subject to some regularity conditions Honoré and Kyriazidou (2000) have shown that the parameters are identified when we have four or more consecutive observations per individual. The basic idea follows that of the conditional likelihood approach. Consider the following events $A = \{y_{i0}, y_{i1} = 0, y_{i2} = 1, y_{i3}\}$ and $B = \{y_{i0}, y_{i1} = 1, y_{i2} = 0, y_{i3}\}$ where y_{i0} and y_{i3} are either 0 or 1. In this case we have

$$\begin{aligned}\Pr\left(A | X_i, \alpha_i\right) &= p_0\left(X_i, \alpha_i\right)^{y_{i0}} \left(1 - p_0\left(X_i, \alpha_i\right)\right)^{1-y_{i0}} \times \frac{1}{1 + \exp\left(X_{i1}\beta + \gamma y_{i0} + \alpha_i\right)} \\ &\quad \times \frac{\exp\left(X_{i2}\beta + \alpha_i\right)}{1 + \exp\left(X_{i2}\beta + \alpha_i\right)} \times \frac{\exp\left(y_{i3}X_{i3}\beta + y_{i3}\gamma + y_{i3}\alpha_i\right)}{1 + \exp\left(X_{i3}\beta + \gamma + \alpha_i\right)}\end{aligned} \tag{11}$$

while

$$
\begin{aligned}
\Pr(B|X_i,\alpha_i) &= p_0(X_i,\alpha_i)^{y_{i0}}(1-p_0(X_i,\alpha_i))^{1-y_{i0}} \times \frac{\exp(X_{i1}\beta+\gamma y_{i0}+\alpha_i)}{1+\exp(X_{i1}\beta+\gamma y_{i0}+\alpha_i)} \\
&\quad \times \frac{1}{1+\exp(X_{i2}\beta+\gamma+\alpha_i)} \times \frac{\exp(y_{i3}X_{i3}\beta+y_{i3}\alpha_i)}{1+\exp(X_{i3}\beta+\alpha_i)} \qquad (12)
\end{aligned}
$$

Noticing that if $X_{i2}=X_{i3}$ we can get rid of the α_i's so that we end up with

$$
\begin{aligned}
\Pr(A|X_i,\alpha_i,A\cup B,X_{i2}=X_{i3}) &= \frac{\Pr(A|X_i,\alpha_i,X_{i2}=X_{i3})}{\Pr(A|X_i,\alpha_i,X_{i2}=X_{i3})+\Pr(B|X_i,\alpha_i,X_{i2}=X_{i3})} \\
&= \frac{1}{1+\exp((X_{i1}-X_{i2})\beta+\gamma(y_{i0}-y_{i3}))} \qquad (13)
\end{aligned}
$$

Identification naturally extends to the case with more than 4 observations for each individual, see Honoré and Kyriazidou (2000). Identification in this case comes from all individuals changing state between two of the middle periods (that is any period but the first and last). Honoré and Kyriazidou (2000) propose to estimate β and γ by maximizing

$$
\sum_{i=1}^{n}\sum_{1\le t<s\le T-1}\left[\begin{array}{c} 1\{y_{it}+y_{is}=1\}K\left(\frac{X_{it+1}-X_{is+1}}{h_n}\right) \\ \ln\left(\frac{\exp((X_{it}-X_{is})b+g(y_{it-1}-y_{is+1})+g(y_{it+1}-y_{is-1})1\{s-t>1\})^{y_{it}}}{1+\exp((X_{it}-X_{is})b+g(y_{it-1}-y_{is+1})+g(y_{it+1}-y_{is-1})1\{s-t>1\})}\right)\end{array}\right] \qquad (14)
$$

over some compact set and where $K(\cdot)$ denotes a kernel density function which assigns a large probability for values where X_{it+1} and X_{is+1} are close. The advantage of this estimator is that it is completely agnostic about the nature of individual heterogeneity.

In the second part of the paper we examine the effect of the worker's own disposable income gain between receiving social assistance benefits and becoming employed. Since we do not observe the wage rate of recipients of social assistance we need to make a selectivity-corrected wage. We do this by estimating the Vella and Verbeek (1998, 1999) sample selection model, where both the equation of interest and the selection equation include random effects and error terms that are allowed to be correlated. We have

$$
\begin{aligned}
\ln w_{it} &= X_{it}\beta_1+\mu_i+\eta_{it} \\
y_{it} &= 1(\gamma y_{it-1}+X_{it}\beta_2+\alpha_i+\varepsilon_{it}\ge 0)=\Phi(\gamma y_{it-1}+X_{it}\beta_2+\alpha_i)
\end{aligned}
$$

where $\ln w_{it}$ is the log of the wage rate, $\mu_i \sim iN(0,\sigma_\mu^2)$, $\eta_{it}\sim iN(0,\sigma_\eta^2)$, $\alpha_i\sim iN(0,\sigma_\alpha^2)$, and $\varepsilon_{it}\sim iN(0,\sigma_\varepsilon^2)$. Denote the composite error $\upsilon_{it}=\mu_i+\eta_{it}$. Then, the conditional mean of the log of the wage rate is given by

$$
E[\ln w_{it}|X_i,y_{i0},y_i]=X_{it}\beta+\tau_1\upsilon_{it}+\tau_2\bar{\upsilon}_i \qquad (15)
$$

where the two last terms are the selection bias we correct for, and where $\tau_1 = \sigma_{\varepsilon\eta}/\sigma_\varepsilon^2$ and $\tau_2 = T\left(\sigma_{\alpha\mu} - \sigma_{\varepsilon\eta}\sigma_\mu^2/\sigma_\varepsilon^2\right)/\left(\sigma_\eta^2 + T\sigma_\mu^2\right)$ are constants to be estimated.

Similar to Vella and Verbeek (1998) we use the labor market state in the previous year as exclusion restriction and exploit that the earnings equation is static while the participation equation is dynamic. This aligns with the theoretical search model in section 2. For example, consider a negative health shock in the current period. The health shock implies a lower wage and through the reservation wage a lower probability of participating in the current period and, thereby, also in the subsequent period. Our identifying assumption is that in the subsequent period the past period's health shock does not affect the wage, since it is static. Using the lagged dependent variable as exclusion restriction is not necessary innocuous. For example, if state dependence arises because employers use unemployment as a signal of low productivity, it is obviously not reasonable to assume that the wage is not affected by the previous labor market state.

We compute the worker's own disposable income gap on a yearly basis by computing the disposable income from working full time and subtracting the potential disposable income from receiving only social assistance.

In order to compute the disposable income when working we begin by estimating the Vella-Verbeek sample selection model. The resulting wage prediction is used to compute the annual full-time wage income. Next, we use a detailed modeling of the Danish tax system, which takes the joint taxation of the spouses into account, in order to calculate the tax payments and finally we compute the disposable income.

For each person we also make a detailed computation of the social assistance benefits they would be entitled to on a yearly basis. Again, we subtract the appropriate tax payments and calculate the resulting disposable income.

Furthermore, in both scenarios we take full account of housing benefits for renters received from the government, which depend both on the tenant's family income and on the rent.[2]

With this information we compute the gap between the disposable income when working full-time and the disposable income of receiving social assistance. We do not take account of transportation costs and child care costs when we compute the income gaps.

5 Does Partner Income Create Work (dis-)Incentives?

In this section we investigate the importance of the spouse's disposable income for the participation decision by including permanent and transitory spouse income among the explanatory

[2] The dataset contains information for the yearly rent in 1999. Unfortunately, for only approximately 40 per cent of dwellings for rent we observe the yearly rent. Following Munch and Svarer (2002) we use Heckman's (1979) two-step procedure to predict the yearly rent. As exclusion restriction we use the number of apartments in the building, which has positive effect on the probability of observing a rent.

variables.

Table 3 gives the results for the estimations on the sample of male workers who at some point during the sample window primarily received social assistance. As expected, we find a significant hump-shaped effect from higher experience. More experience raises the probability of participation, but the effect declines with total years of experience. Controlling for this effect, age actually works in the opposite direction, so that older people ceteris paribus have lower participation probabilities. Having people with a vocational education as our reference category, we see that more education increases the probability of working.

We find that everything else being equal, immigrants have a higher probability of working. We think this is an effect of immigrants within the social assistance system being a more heterogenous group than the rest of the recipients. Newly arrived immigrants do not have access to unemployment benefits and this group is likely to contain both low and high productive individuals.

We have included the local gender specific unemployment rate and the number of regional vacancies normalized by the labor force to capture demand and supply effects in the regional labor market. Both variables are significant with expected signs, that is a positive effect of more vacancies and a negative effect of a higher unemployment rate.

Having children - and especially young children - lowers the probability of being employed. Finally, we have included three variables for the health status of the worker. These variables are defined as the yearly number of treatments with financial support from the Danish public health system within the given area. We find that the more neuro medicine treatments, or the more consultations with psychiatrists or doctors the lower the probability of being employed.

The first two specifications the pooled probit model and random effects probit model deliver similar results. Although the pooled probit model correctly predicts 71 per cent of the outcomes, the confusion matrices in Table 10 appendix A reveal that the pooled probit model only does a good job in fitting the $y = 0$ outcome, while the dynamic correlated random effects model in the third column performs much better in fitting also the $y = 1$ outcome. In the dynamic correlated random effects specification we approximate the unobserved individual part by the means of age, region dummies, children and health variables besides the random normal term. Besides the improved overall fit the picture does not change much. However, the coefficient to age becomes positive when the individual mean of age is part of the unobservable part. Similarly, the children variables become insignificant. The set of parameter estimates obtained from the fixed effects logit are similar to the various probit estimates, while we only can use the estimation results from the dynamic fixed effects to verify the presence of (true) state dependence.

The final column in Table 3 shows that using an alternative cut-off level of 42 weeks for the definition of being a recipient of social assistance compared to the 27 weeks used in all other estimations does not matter for the results.

Table 3: Participation probability, partner income, males

	Pooled probit	Random effects probit	Dynamic correlated random effects probit	Fixed effects logit	Dynamic fixed effects logit	Dynamic correlated random effects probit, 42 weeks
Spouse' temporary income	-0.008	-0.009	-0.007	-0.032	-0.017	-0.013
	(0.007)	(0.009)	(0.008)	(0.016)*	(0.052)	(0.009)
Spouse's permanent income	0.030	0.044	0.027	-	-	0.024
	(0.004)**	(0.007)**	(0.005)**	-	-	(0.006)**
Age	-0.071	-0.108	0.084	-	-	0.071
	(0.011)**	(0.019)**	(0.021)**	-	-	(0.025)**
Age squared/100	0.041	0.068	0.006	-	-	0.000
	(0.014)**	(0.025)**	(0.017)	-	-	(0.000)
Experience	0.202	0.313	0.118	0.785	0.018	0.101
	(0.006)**	(0.012)**	(0.009)**	(0.070)**	(0.117)	(0.010)**
Experience squared/100	-0.587	-0.952	-0.343	-2.938	-	-0.003
	(0.025)**	(0.050)**	(0.033)**	(0.247)**	-	(0.000)**
Unskilled	-0.117	-0.139	-0.091	-	-	-0.090
	(0.027)**	(0.053)**	(0.034)**	-	-	(0.039)*
Short-cycle higher education	0.097	0.261	0.148	-	-	0.116
	(0.068)	(0.132)*	(0.085)	-	-	(0.094)
Medium-cycle higher education	0.161	0.208	0.174	-	-	0.169
	(0.059)**	(0.119)	(0.076)*	-	-	(0.082)*
Long-cycle higher education	0.182	0.302	0.115	-	-	0.185
	(0.069)**	(0.138)*	(0.089)	-	-	(0.093)*
Immigrant	0.336	0.406	0.235	-	-	0.151
	(0.032)**	(0.062)**	(0.041)**	-	-	(0.046)**
Second generation immigrant	-0.039	0.072	0.023	-	-	0.088
	(0.156)	(0.262)	(0.182)	-	-	(0.222)
Copenhagen	-0.046	-0.012	0.101	0.660	-	0.084
	(0.046)	(0.085)	(0.221)	(0.396)	-	(0.263)
Large city	-0.101	-0.186	-0.247	-0.144	-	-0.453
	(0.046)*	(0.087)*	(0.279)	(0.502)	-	(0.317)
Rural area	-0.070	-0.125	-0.158	0.141	-	-0.131
	(0.038)	(0.072)	(0.234)	(0.424)	-	(0.274)
Children aged 0-6 years	-0.128	-0.167	-0.054	-0.207	-0.746	-0.056
	(0.014)**	(0.025)**	(0.040)	(0.072)**	(0.223)**	(0.046)
Children aged 7-17 years	-0.034	-0.050	0.012	-0.054	-	-0.004
	(0.013)**	(0.023)*	(0.042)	(0.078)	-	(0.048)
Unemp. on municipality and gender	-0.062	-0.077	-0.052	-0.148	-	-0.047
	(0.008)**	(0.014)**	(0.010)**	(0.040)**	-	(0.011)**
Regional vacancies	2.603	2.876	1.908	3.574	-	0.845
	(0.788)**	(1.404)*	(0.976)	(5.154)	-	(1.129)
Neuro medicine	-0.187	-0.246	-0.183	-0.370	-	-0.150
	(0.060)**	(0.084)**	(0.085)*	(0.191)	-	(0.081)
Psychiatry	-0.036	-0.028	0.001	0.010	-	0.023
	(0.011)**	(0.015)	(0.019)	(0.035)	-	(0.021)
General medical treatment	-0.005	-0.004	-0.003	-0.004	-	-0.004
	(0.001)**	(0.001)**	(0.001)*	(0.002)	-	(0.002)
y_0 (initial participation)	-	-	0.014	-	-	-0.041
	-	-	(0.037)	-	-	(0.044)
Lagged participation	-	-	1.312	-	2.648	1.292
	-	-	(0.034)**	-	(0.215)**	(0.040)**
Observations	15,383	15,383	15,383	6,761	3,652	11,827
Number of persons	5,212	5,212	5,212	1,626	694	4,067

Standard errors in parentheses. All equations besides the dynamic fixed effects logit include time-dummies.
The correlated random effects contain the means of age, region dummies, children and health variables.
* significant at 5 per cent; ** significant at 1 per cent

The spouse's temporary income is insignificant in all estimations, but the fixed effects logit, where the coefficient is significantly negative. In the absence of major tax or benefit reforms the within change in the spouse's income – the transitory part – is limited. If it takes a certain amount of change in partner income to affect the labor market status, we could fear that we actually observe too little variation in our data to obtain significant estimates. In absence of credit constraints the insignificant parameter estimates suggest that individuals smooth out consumption.

The permanent income only varies between individuals and, hence, the coefficient cannot be identified in the fixed effects specifications. In the rest of the estimations the spouse permanent income is significantly positive. Hence, it seems to be the case that the spouses have the same taste for work by assortative matching in the marriage market. In other words, the positive sign implies that Hyslop's taste effect dominates and, therefore, that the permanent spouse income is endogenous to the employment decision.

Table 4: Participation probability, partner income, females

	Pooled probit	Random effects probit	Dynamic correlated random effects probit	Fixed effects logit	Dynamic fixed effects logit	Dynamic correlated random effects probit, 42 weeks
Spouse' temporary income	0.033	0.041	0.036	0.056	-0.003	0.031
	(0.005)**	(0.006)**	(0.006)**	(0.012)**	(0.063)	(0.006)**
Spouse's permanent income	0.060	0.076	0.054	-	-	0.053
	(0.003)**	(0.005)**	(0.003)**	-	-	(0.004)**
Age	-0.012	0.006	0.179	-	-	0.177
	(0.010)	(0.019)	(0.021)**	-	-	(0.023)**
Age squared/100	-0.033	-0.080	-0.060	-	-	-0.001
	(0.014)*	(0.026)**	(0.018)**	-	-	(0.000)**
Experience	0.252	0.389	0.144	0.708	0.095	0.135
	(0.007)**	(0.015)**	(0.011)**	(0.079)**	(0.221)	(0.013)**
Experience squared/100	-0.940	-1.517	-0.539	-3.570	-	-0.005
	(0.035)**	(0.067)**	(0.050)**	(0.365)**	-	(0.001)**
Unskilled	-0.348	-0.610	-0.371	-	-	-0.345
	(0.028)**	(0.055)**	(0.037)**	-	-	(0.042)**
Short-cycle higher education	0.090	0.059	0.063	-	-	0.052
	(0.081)	−(0.168)	−(0.106)	-	-	(0.119)
Medium-cycle higher education	-0.107	-0.151	-0.084	-	-	-0.159
	(0.066)	−(0.131)	−(0.086)	-	-	(0.101)
Long-cycle higher education	-0.245	-0.401	-0.327	-	-	-0.370
	(0.121)*	(0.226)	(0.152)*	-	-	(0.177)*
Immigrant	0.208	0.178	0.085	-	-	0.027
	(0.030)**	(0.058)**	(0.039)*	-	-	(0.043)
Second generation immigrant	0.316	0.479	0.224	-	-	0.617
	(0.141)*	(0.261)	(0.180)	-	-	(0.197)**
Copenhagen	-0.031	-0.043	-0.156	-0.203	-	0.014
	(0.044)	(0.084)	(0.242)	(0.439)	-	(0.282)
Large city	-0.075	-0.134	-0.586	-0.731	-	-0.708
	(0.045)	(0.086)	(0.273)*	(0.483)	-	(0.317)*
Rural area	0.050	0.025	-0.098	-0.198	-	-0.140
	(0.038)	(0.072)	(0.220)	(0.374)	-	0.258
Children aged 0-6 years	-0.278	-0.369	-0.196	-0.570	0.113	-0.247
	(0.015)**	(0.025)**	(0.041)**	(0.082)**	(0.305)	(0.0460)**
Children aged 7-17 years	-0.059	-0.096	-0.107	-0.339	-	-0.127
	(0.012)**	(0.022)**	(0.042)*	(0.085)**	-	(0.045)**
Unemp. on municipality and gender	-0.037	-0.052	-0.037	-0.096	-	-0.044
	(0.007)**	(0.013)**	(0.009)**	(0.042)*	-	(0.010)**
Regional vacancies	2.468	2.598	2.276	-5.571	-	1.622
	(0.741)**	(1.341)	(0.944)*	(4.882)	-	(1.073)
Neuro medicine	-0.086	-0.108	-0.079	-0.117	-	-0.089
	(0.033)**	(0.046)*	(0.049)	(0.095)	-	(0.048)
Psychiatry	-0.056	-0.057	-0.021	-0.043	-	-0.015
	(0.010)**	(0.014)**	(0.016)	(0.033)	-	(0.018)
General medical treatment	-0.006	-0.005	-0.001	0.001	-	-0.003
	(0.001)**	(0.001)**	(0.001)	(0.002)	-	(0.001) **
y_0 (initial participation)	-	-	-0.202	-	-	-0.144
	-	-	(0.039)**	-	-	(0.044)**
Lagged participation	-	-	1.537	-	3.752	1.461
	-	-	(0.035)**	-	(0.294)**	(0.039)**
Observations	21,149	21,149	21,149	7,048	3,335	17,284
Number of persons	6,985	6,986	6,986	1,736	638	5,808

Standard deviations are in parentheses. All equations besides the dynamic fixed effects logit include time-dummies. The correlated random effects contain the means of age, region dummies, children and health variables.
* significant at 5 per cent; ** significant at 1 per cent

The results for women are shown in Table 4. There are some noticeable differences in the parameter estimates compared to men. First, whereas men with higher education have a higher probability of being employed, the comparable effect for women is insignificant or even negative.

Second, there is no significant difference in the employment probability between living in the capital of Copenhagen and in rural areas. For women we find significantly positive effects of both temporary and permanent income, which again reflect the taste effect endogeneity.

Table 5 presents elasticities for the previous estimations. From the dynamic random effects probit model it can be seen that participation in itself increases the probability of also being in the labor market next year by about 40 per cent. In a recent paper Ahmad (2007) also estimates reduced-form participation equations for immigrants receiving social assistance benefits and argues that there is a large degree of state dependence among immigrants in Denmark.

Table 5: Elasticities, partner income

	Pooled probit	Random effects probit	Dynamic random effects probit	Dynamic random effects probit, 42 weeks
Males				
Lagged dependent variable	-	-	0.405	0.398
	-	-	(0.020)**	(0.023)**
Spouse's temporary income	-0.002	-0.002	-0.002	-0.004
	(0.014)	(0.013)	(0.016)	(0.019)
Spouse's permanent income	0.339	0.358	0.303	0.276
	(0.052)**	(0.068)**	(0.067)**	(0.081)**
Females				
Lagged dependent variable	-	-	0.408	0.379
	-	-	(0.022)**	(0.024)**
Spouse's temporary income	0.007	0.006	0.007	0.006
	(0.042)	(0.032)	(0.045)	(0.051)
Spouse's permanent income	0.918	0.791	0.805	0.798
	(0.054)**	(0.067)**	(0.068)**	(0.075)**

Standard deviations are in parentheses.
* significant at 5 per cent; ** significant at 1 per cent

Unlike both Hyslop's (1999) and Croda and Kyriazidou's (2003) results for married women we find that the permanent part of the spouse income is endogenous for both men and women and, therefore, that it is not useful for examining the effects of financial incentives for the participation decision. The reason seems to be that the difference in the participation rate between the genders is smaller in Denmark than in Germany and the US. Additionally, the share of women in part-time work is much higher in Germany. With respect to the transitory part of the spouse's income we similar to Hyslop (1999) and Croda and Kyriazidou (2003) find very small and insignificant effects, which suggest that workers are not credit constrained and can smooth out consumption.

6 Does Own Income Create Work Incentives?

The results from the previous section show the weakness of using spouse's income as predictor for participation, namely that it may be endogenous. Furthermore, depending on how income is used in a family, it may be the case that the response is larger for the own disposable income gain from working. Hence, the results from the previous section do not necessarily imply that there are low participation elasticities for recipients of social assistance. Instead in this section we use predicted own disposable income gaps to investigate the importance of financial incentives for the participation decision.

Table 6 shows the second step of the Vella-Verbeek sample selection model. It is striking that most of the variables are insignificant. For example for women there is no significant wage difference between being unskilled and having a university degree. This seems to suggest that for the sample of social assistance recipients the human capital model works poorly. This is also reflected in the low R^2 of $7-9$ per cent. Using Danish register data it is far from unusual to obtain an explanatory power of more than 30 per cent. As a matter of fact, by considering all the persons for whom we observe a wage rate we obtain an R^2 of about 20 per cent using only experience and its square as well as schooling to explain the variation of the log wages.[3] In order to understand the results in Table 6 one might again refer to Table 1 and 2 to see that the level of the mean wage implies, that the minimum wage constraint is binding for a large part of the recipients of social assistance, perhaps suggesting that these people are unable to receive a wage matching their marginal productivity measured by the usual Mincer explanatory variables.

For men the selection into the sub-sample of persons for whom we observe the wage is positive with respect to the time-invariant unobservables. This is as expected because this suggests, that people with higher ability earn more and have a higher probability of finding employment. For women we find no significant selection in terms of time-invariant unobservables.

[3] These results are not shown.

Table 6: Wage equation, random effects

	Male	Female
Age	0.017	0.005
	(0.003)**	(0.004)
Age squared/100	-0.022	-0.005
	(0.004)**	(0.005)
Experience	0.010	0.006
	(0.003)**	(0.003)
Experience squared/100	-0.023	-0.012
	(0.009)*	(0.013)
Unskilled	-0.021	0.002
	(0.009)*	(0.010)
Short-cycle higher education	0.027	0.024
	(0.028)	(0.036)
Medium-cycle higher education	0.110	0.114
	(0.025)**	(0.024)**
Long-cycle higher education	0.171	0.072
	(0.030)**	(0.042)
Immigrant	-0.004	-0.002
	(0.012)	(0.013)
Second generation immigrant	-0.016	0.024
	(0.042)	(0.044)
Copenhagen	0.023	0.042
	(0.014)	(0.015)**
Large city	-0.017	0.003
	(0.015)	(0.017)
Rural area	-0.021	-0.011
	(0.011)	(0.013)
Children aged 0-6 years	0.012	-0.010
	(0.006)*	(0.006)
Children aged 7-17 years	-0.003	-0.006
	(0.005)	(0.005)
Unemp. on municipality and gender	-0.004	-0.006
	(0.002)	(0.003)*
Regional vacancies	-0.558	0.020
	(0.245)*	(0.282)
Neuro medicine	0.026	-0.004
	−(0.014)	(0.009)
Psychiatry	0.008	-0.002
	(0.003)**	(0.003)
General medical treatment	0.000	0.000
	(0.000)	(0.000)
Mean (generalized residual)	0.028	0.003
	(0.004)**	(0.005)
Generalized residual	-0.035	-0.032
	(0.007)**	(0.007)**
Variance of individual specific error	0.213	0.201
Variance of time-varying error	0.224	0.219
R-squared	0.087	0.075
Observations	11,124	7,823
Persons	4,569	3,485

Note: Standard deviations are in parentheses.
Both equations include time-dummies.
*significant at 5 per cent; **significant at 1 per cent

Table 7: Participation probability, own income, males

	Pooled probit	Random effects probit	Dynamic random effects probit
Own income gap/10000	0.031	0.044	0.045
	(0.003)**	(0.005)**	(0.004)**
Age	-0.067	-0.082	0.040
	(0.006)**	(0.010)**	(0.013)**
Age squared/100	0.037	0.041	0.007
	(0.008)**	(0.013)**	(0.009)
Experience	0.185	0.263	0.097
	(0.004)**	(0.007)**	(0.005)**
Experience squared/100	-0.526	-0.780	-0.285
	(0.016)**	(0.030)**	(0.020)**
Unskilled	-0.127	-0.178	-0.088
	(0.017)**	(0.032)**	(0.020)**
Short-cycle higher education	0.028	0.076	-0.007
	(0.048)	(0.088)	(0.057)
Medium-cycle higher education	0.166	0.222	0.110
	(0.043)**	(0.080)**	(0.051)*
Long-cycle higher education	0.258	0.299	0.077
	(0.050)**	(0.093)**	(0.060)
Immigrant	0.285	0.338	0.219
	(0.020)**	(0.037)**	(0.025)**
Second generation immigrant	-0.009	0.039	0.001
	(0.076)	(0.135)	(0.089)
Copenhagen	0.011	0.037	0.025
	(0.027)	(0.046)	(0.085)
Large city	-0.089	-0.118	-0.105
	(0.028)**	(0.049)*	(0.111)
Rural area	-0.030	-0.041	0.024
	−(0.022)	−(0.040)	−(0.089)
Children aged 0-6 years	0.003	0.002	0.025
	−(0.012)	−(0.019)	−(0.026)
Children aged 7-17 years	0.076	0.106	0.122
	(0.010)**	(0.017)**	(0.028)**
Unemp. on municipality and gender	0.049	0.064	0.042
	(0.005)**	(0.008)**	(0.006)**
Regional vacancies	2.127	1.965	1.632
	(0.483)**	(0.814)*	(0.571)**
Neuro medicine	0.077	0.084	-0.068
	(0.032)*	(0.041)*	(0.043)
Psychiatry	0.032	0.032	-0.016
	(0.006)**	(0.008)**	(0.010)
General medical treatment	0.006	0.005	-0.002
	(0.000)**	(0.001)**	(0.001)**
y_0 (initial participation)	0.000	0.000	0.079
	(0.000)	(0.000)	(0.022)**
Lagged participation	0.000	0.000	1.325
	(0.000)	(0.000)	(0.021)**
Observations	-	-	41,390
Number of persons	-	-	10,689

Standard deviations are in parentheses. All equations include time-dummies.
The correlated random effects contain the means of age, region dummies, children and health variables.
* significant at 5 per cent; ** significant at 1 per cent

Table 8: Participation probability, own income, females

	Pooled probit	Random effects probit	Dynamic random effects probit
Own income gap/10000	0.045	0.059	0.049
	(0.002)**	(0.003)**	(0.003)**
Age	-0.014	0.001	0.173
	(0.007)*	(0.013)	(0.016)**
Age squared/100	-0.030	-0.071	-0.047
	(0.009)**	(0.017)**	(0.012)**
Experience	0.216	0.320	0.111
	(0.005)**	(0.010)**	(0.007)**
Experience squared/100	-0.772	-1.188	-0.412
	(0.023)**	(0.047)**	(0.031)**
Unskilled	-0.337	-0.617	-0.345
	(0.020)**	(0.039)**	(0.026)**
Short-cycle higher education	-0.039	-0.134	-0.070
	(0.062)	(0.125)	(0.081)
Medium-cycle higher education	-0.195	-0.247	-0.202
	(0.046)**	(0.089)**	(0.058)**
Long-cycle higher education	0.079	-0.133	-0.159
	(0.073)	(0.144)	(0.094)
Immigrant	0.041	0.003	-0.031
	(0.021)	(0.042)	(0.028)
Second generation immigrant	0.033	0.083	-0.007
	(0.086)	(0.164)	(0.108)
Copenhagen	-0.091	-0.131	-0.109
	(0.029)**	(0.054)*	(0.115)
Large city	-0.090	-0.145	-0.356
	(0.031)**	(0.059)*	(0.144)*
Rural area	0.028	0.014	-0.082
	(0.026)	(0.049)	(0.114)
Children aged 0-6 years	-0.153	-0.206	-0.128
	(0.011)**	(0.019)**	(0.028)**
Children aged 7-17 years	0.019	0.006	-0.032
	(0.009)*	(0.017)	(0.028)
Unemp. on municipality and gender	-0.026	-0.039	-0.024
	(0.005)**	(0.009)**	(0.006)**
Regional vacancies	1.009	1.120	0.387
	(0.540)	(0.960)	(0.684)
Neuro medicine	-0.073	-0.081	-0.065
	(0.023)**	(0.031)**	(0.032)*
Psychiatry	-0.045	-0.051	-0.037
	(0.006)**	(0.009)**	(0.010)**
General medical treatment	-0.003	-0.002	0.000
	(0.000)**	(0.000)**	(0.001)
y_0 (initial participation)	0.000	0.000	-0.075
	(0.000)	(0.000)	(0.027)**
Lagged participation	-	-	1.503
	-	-	(0.025)**
Observations	40,170	40,170	40,170
Number of persons	10,586	10,586	10,586

Standard errors in parentheses. All equations include time-dummies.
The correlated random effects contain the means of age, region dummies, children and health variables.
* significant at 5 per cent; ** significant at 1 per cent

The results from the participation equation where we include the worker's own disposable income gain from working is presented in Table 7 and 8. Overall, the picture is similar compared to Table 3 and 4 although there are some changes due to the fact that Table 7 and 8 also include singles. When we estimate on both couples and singles we find that only children aged 0-6 years have a negative influence on the employment probability and for men we even find significant positive coefficients to the number of children aged 7-17 years.

For women there are two more changes. First, the coefficient to medium-cycle higher eduation becomes significantly negative, although this does not change the conclusion that there is an insignificant or negative effect on the participation probability of having a longer education. Thus, for workers with a weak attachment to the labor market and low experience, vocational education, which is the reference category, might be the most favourable education. Furthermore, the signaling effect of a social assistance spell to the employer might be more severe for educations with an overall very low unemployment rate.

Second, immigrants are no longer more likely to become employed than natives. This suggests that single female immigrants do not participate to the same degree as married female immigrants. However, the reason might be that single mothers to a larger extent are paid additional benefits in an ad-hoc way, which depends on the local government which, therefore, are impossible to model. This implies that the calculated income gap for single female immigrants is overvalued, which in turn may imply that they seem to be less likely to participate.

Own disposable income gap is significantly positive in both estimation frameworks and for both genders, and thus, having a higher income gap raises the probability of participation.

Table 9: Elasticities, own income gap

	Pooled probit	Random effects probit	Dynamic random effects probit
Males			
Lagged dependent variable	-	-	0.422
	-	-	(0.012)
Own income gap	0.276	0.305	0.422
	(0.034)	(0.037)	(0.044)
Change in own income gap, 5,000 DKK	0.018	0.020	0.028
Females			
Lagged dependent variable	-	-	0.414
	-	-	(0.015)
Own income gap	0.655	0.618	0.678
	(0.033)	(0.037)	(0.044)
Change in own income gap, 5,000 DKK	0.052	0.049	0.054

Standard deviations are in parentheses.
* significant at 5 per cent; ** significant at 1 per cent

Table 9 presents the elasticities from changing own disposable income gap on the participation probability. The elasticities for men are in the range of $0.28 - 0.44$, while the elasticities for women are roughly double the size. The estimated elasticities imply that increasing the disposable income with $5,000$ DKK delivers an increase in the particpation probability of 0.025 for men and 0.05 for women. Another way of measuring the participation elasticities, which for example is used in the CGE model DREAM, is the percentage change in the number of recipients of social assistance from a percentage change in the income gap. In Appendix B we show how to convert the elasticities and for men the elasticity is $0.09 - 0.17$, while we for women obtain $0.13 - 0.20$.

In the absence of major benefit and tax policy changes in the considered period we do not want to identify the effect of the disposable income gap solely from within-individual variation. Therefore, we do not include the mean of disposable income in the linear approximation for the unobservable individual effect. However, this implies that we cannot appropriately control for unobserved heterogeneity, which may bias the results. If receiving social assistance is voluntary, but joy of working is positively correlated with productivity, the elasticity will be upward biased.

It is obviously pointless to consider the effects of economic incentives if receiving benefits is completely involuntary. Even though this extreme is not the case, it is likely that some recipients of social assistance would prefer to be employed, but have a low probability of getting a job due to very low productivities. In this case our estimated model confuses low incentives with low participation probability and the estimated elasticity will also tend to be upward biased. In orther words, the effects from an actual tax or benefit reform will be smaller than our income elasticity estimates.

There exist a few other Danish studies also estimating participation income elasticities. Graversen (1996) estimates participation elasticities for workers in the labor force from a natural experiment study of the 1987 tax reform. He finds participation elasticities in the range of $0.2 - 0.7$ for single women and 0.05 for married women. Pedersen and Smith (2002) also computes disposable income gaps from working, but do as Graversen focus on unemployed workers in the labor force. They use survey information on the unemployed worker's expected wage and use information on employed workers' transportation costs and child care costs. For 1996 they find an income elasticity of 0.3 for men and 0.7 for women. Pedersen and Smith do not solve the potential endogeneity problem of the worker's disposable income. The elasticities are strikingly similar in our study and in Pedersen and Smith, which should imply that the effect of financial incentives are similar among the unemployed workers in the labor force and among recipients of social assistance. However, since the share of very low productivity workers is larger among the repicipents of social assistance than the recipents of unemployment benefits, the share who is involuntary out of employment is probably largest among the repicipents of social assistance. Hence, it is probably the case that our estimates for the recipients of social assistance have the largest bias.

In contrast to the two mentioned studies, Toomet (2005) focus on social assistance and exploits the fact that the social assistance benefits is increased by 70 per cent when recipients without children turns 25 years, while the benefit level remains constant at the age of 25 years for recipients with children. Hence, Toomet uses the latter group as control group. From a difference-in-difference estimation a participation elasticity of income of 0.4 is found for women while the income effect for men is found to be insignificant.

In a recent survey-based study, Graversen and Tinggaard (2005) examine the effects of the implementation of the social assistance ceiling in 2004.[4] Approximately 1,000 social assistance recipients were interviewed just before the implementation and again nine months after the implementation of the social assistance benefit ceiling. Although the ceiling has reduced the amount of social assistance received and, hence, was expected to provide larger incentives, Graversen and Tinggaard conclude that there seems to be no effect on participation and on whether the recipient search or not. Moreover, Graversen and Tinggaard only find very modest effects on the search intensity.

Finally, in a recent study Graversen (2006) uses the variation in social assistance benefits from the implementation of the social assistance ceiling. A pooled probit model for participation is estimated for each month for 18 months. Graversen compute the income gap from working similar to here although he uses a median wage for all persons rather than a predicted wage. This way Graversen avoids the potential endogeneity of the wages, but the framework only allows Graversen to determine effects from differences in benefits in the short term. No significant employment effects are found from this analysis.

7 Conclusion

In this paper we have examined the effects of economic incentives on the labor market participation for recipients of social assistance. A simple examination of the characteristics of recipients of social assistance reveals the group has weak attachment to the labor market. This conclusion is further strengthened by the poor performance of the human capital model and the large degree of state dependence in the employment status. Hence, to some extent we would believe that this group of people is involuntary nonemployed and that only small effects from economic incentives are to be expected. We find no employment effects of incentives from the within variation in the spouse's income. Recently, Graversen and Tinggard (2005) and Graversen (2006) have evaluated the short-term effect of the social assistance ceiling. They find no or very small effects.

However, from estimations where we use predicted disposable income gain from working

[4] The social assistance ceiling aimed at providing economic incentives for married social assistance recipients, by setting a reduced maximum of social assistance benefits that a household can receive. It came into effect January 1, 2004.

we obtain elasticities in the range of $0.28 - 0.44$ for men and in the range of $0.62 - 0.68$ for women. These elasticities are of similar magnitude as those Pedersen and Smith (2002) find for workers in the labor force. However, neither we, nor Pedersen and Smith are able to sufficiently control for unobserved heterogeneity of the disposable income gap. Therefore, we believe that the estimates are likely to be upward biased. When converting the elasticities to percentage change in the number of recipients of social assistance from a percentage change in the income gap as used in the CGE model DREAM the estimated elasticities correspond to $0.09 - 0.17$ for men and $0.13 - 0.20$ for women.

References

[1] Ahmad, N. (2007): Dynamic Labour Market Behaviour of Immigrants in Denmark, working paper, University of Aarhus.

[2] Chamberlain, G. (1993): Feedback in Panel Data Models, unpublished manuscript, Department of Economics, Harvard University.

[3] Croda, E. and E. Kyriazidou (2003): Intertemporal Labor Force Participation of Married Women in Germany: A Panel Data Analysis, Working Paper.

[4] Danish Economic Council (2004): Dansk økonomi efterår 2004, Copenhagen.

[5] Danish Welfare Commission (2005): Fremtidens velfærd – vores valg, Copenhagen .

[6] Dex, S., S. Gustafsson, N. Smith and T. Callan (1995): Cross-National Comparisons of the Labour Force Participation of Women Married to Unemployed Men, *Oxford Economic Papers* **47**, pp. 611-635.

[7] Frederiksen, A., E.K. Graversen and N. Smith (2001): Overtime Work, Dual Job Holding and Taxation, IZA discussion paper 323, Bonn.

[8] Garibaldi, P. and with E. Wasmer (2004): Raising Female Employment: Reflections and Policy Tools, *Journal of the European Economic Association* **2**, pp. 320-330.

[9] Graversen, B.K. (2006): Making Work Pay: Is There an Employment Effect for Disadvantaged families, working paper, The Danish National Institute of Social Research.

[10] Graversen, E.K. (1996): Measuring Labour Supply Responses to Tax Changes by Use of Exogenous Tax Reforms, CLS working paper 96-17, University of Aarhus and Aarhus Business School.

[11] Graversen, B.K. and K. Tinggaard (2005): Loft over ydelser – Evaluering af loftet over ydelser til Kontanthjælpsmodtagere, working paper, The Danish National Institute of Social Research.

[12] Greene, W.H. (2003): *Econometric Analysis*, Prentice Hall.

[13] Heckman, J.J. (1979): Sample Selection Bias as a Specification Error, *Econometrica* **46**, pp. 931-959.

[14] Heckman, J.J. (1981), The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process, in C. F. Manski and D. McFadden (eds.): *Structural Analysis of Discrete Data with Econometric Applications*, pp. 179–195, MIT Press.

[15] Heckman, J.J. (1993): What Have Been Learned about Labor Supply in the Past Twenty Years, *American Economic Review* **47**, pp. 116-121.

[16] Honoré, B. E., and E. Kyriazidou (2000), Panel Data Discrete Choice Models with Lagged Dependent Variables, *Econometrica* **68**, 839–874.

[17] Hyslop, D. (1999): State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women, *Econometrica* **67**, 1255-1294.

[18] Mundlak, Y. (1978): On the Pooling of Time Series and Cross Section Data, *Econometrica* **46**, pp. 69–85.

[19] Munch, J.R. and M. Svarer (2002): Rent Control and Tenancy Duration, *Journal of Urban Economics* **52**, pp. 542-560.

[20] Pedersen, P. J. and N. Smith (2002): Unemployment traps and financial disincentives to work, *European Sociological Review* **18**, pp. 271-288.

[21] Toomet, O. (2005): Does an Increase in Unemployment Income Lead to Longer Unemployment Spells? Evidence Using Danish Unemployment Assistance Data, working paper, University of Aarhus.

[22] Vella, F. and M. Verbeek (1998): Whose Wages do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men, *Journal of Applied Econometrics* **13**, pp. 163-183.

[23] Vella, F. and M. Verbeek (1999): Two-Step Estimation of Panel Data Models with Censored Endogenous Variable and Selection Bias, *Journal of Econometrics* **90**, pp. 239-263.

[24] Wooldridge, J.M. (2002): *Econometric Analysis of Cross-Section and Panel Data*, MIT Press.

[25] Wooldridge, J.M. (2005): Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Effects, *Journal of Applied Econometrics* **20-1**, pp. 39-54.

Appendix

Appendix A: Goodness of Fit

Table 10: Confusion matrices, spouse income, males

Pooled probit					Dynamic random effects probit			
	Actual					Actual		
Predicted	0	1	Total		Predicted	0	1	Total
0	9,234	3,248	12,482		0	9,192	1,704	10,896
1	1,250	1,651	2,901		1	1,292	3,195	4,487
Total	10,484	4,899	15,383		Total	10,484	4,899	15,383

Table 11: Confusion matrices, spouse income, females

Pooled probit					Dynamic random effects probit			
	Actual					Actual		
Predicted	0	1	Total		Predicted	0	1	Total
0	15,593	3,380	18,973		0	15,609	1,942	17,551
1	959	1,217	2,176		1	943	2,655	3,598
Total	16,552	4,597	21,149		Total	16,552	4,597	21,149

Table 12: Confusion matrices, own income gap, males

Pooled probit					Dynamic random effects probit			
	Actual					Actual		
Predicted	0	1	Total		Predicted	0	1	Total
0	26,877	9,640	36,517		0	25,068	4,276	29,344
1	2,172	2,701	4,873		1	3,981	8,065	12,046
Total	29,049	12,341	41,390		Total	29,049	12,341	41,390

Table 13: Confusion matrices, own income gap, females

Pooled probit					Dynamic random effects probit			
	Actual					Actual		
Predicted	0	1	Total		Predicted	0	1	Total
0	29,905	7,181	37,086		0	29,254	3,821	33,075
1	1,487	1,597	3,084		1	2,138	4,957	7,095
Total	31,392	8,778	40,170		Total	31,392	8,778	40,170

Appendix B: Elasticities Measured as Population Changes

Elasticities are most often defined as the percentage change in participation probability for a percentage change in income gap. Hence, we can define

$$\varepsilon_p = \frac{\partial p}{\partial y}\frac{y}{p} \tag{16}$$

where y is the disposable income gap and p is the participation probability. Another way of thinking about participation elasticities is the percentage change in the number of recipients of social assistance from a percentage change in the income gap. We can express this alternative elasticity as

$$\varepsilon_U = -\frac{\partial U}{\partial y}\frac{y}{U} \tag{17}$$

where U is the number of persons receiving social assistance. Letting N denote the total population, which is constant, and noticing that $U = N(1-p)$ we have that

$$\varepsilon_U = -\frac{-\partial p \cdot N}{\partial y}\frac{y}{N(1-p)} = \frac{p}{(1-p)}\varepsilon_p \tag{18}$$

Since ε_U is the definition used in two recent tax reform simulations on the CGE model DREAM (cf. Danish Economic Council (2004) and Danish Welfare Commission (2005)) we compute these elasticities from our estimations using the predicted participation probability $\hat{p}$, for each individual in our samples.

From (18), we see that only for a participation rate of exactly 50 per cent, the two elasticities are equal. For our sample we have participation rates somewhat below 50 percent, since our samples consist of only people who have been in contact with the social assistance system.

Earnings, Uncertainty, and the Self-Employment Choice*

Daniel le Maire
University of Copenhagen and CAM

Bertel Schjerning
University of Copenhagen and CAM

July 29, 2008

Abstract

This paper investigates the relationship between self-employment choice, expected earnings, and uncertainty. Several interesting results emerge from our analysis on Danish longitudinal register data: Firstly, self-employed (taxable) personal income bunch at kink points in the tax system since self-employed can retain earnings and thereby transfer income across tax-years. Secondly, expected income level and income variance are important determinants in choice of occupation. Thirdly, men put more emphasis on expected earnings level, while women appears more risk averse, which contribute to explain why fewer women are self-employed. Finally, our results suggest that non-western immigrants are marginalized into self-employment.

JEL codes: J16, J24, J31, C33, C35.

*We gratefully acknowledge the comments we have received from Karsten Albæk, Martin Browning, Mette Ejrnæs, Steen Jørgensen, Ulrich Kaiser, Claire Kelly, Nikolaj Malchow-Møller, Jacob Roland Munch, Søren Leth-Petersen, Jacob Palsgaard Petersen, Michael Svarer and Lise Vesterlund. The usual disclaimer applies.

1 Introduction

Compared to wage work, self-employment is a fundamentally different occupation with respect to the type and source of income. While wage workers receive a wage which is subject to a relatively small level of uncertainty, self-employed individuals often face considerably more variation in their income. Moreover, since self-employed typically use own wealth to finance their business, they bear the risk associated with starting up the firm. Therefore, the expected income and the uncertainty of this income are likely to be important determinants of an individual's occupational choice.

The main objective of this paper is to investigate the relationship between the occupational choice and the distributions of associated monetary gains in different occupations. Specifically, we analyze how existing earnings differentials and differences in income uncertainty can explain observed occupational choices. A particular focus is on explaining why fewer women choose to become self-employed.

If individuals are risk averse, we would expect that the self-employed should be compensated for facing higher income uncertainty. However, earnings-differentials may arise for other reasons than risk compensation: Hamilton (2000) argues that cross-sectional earnings differentials may arise due to i) different earnings-experience profiles, ii) self-selection, and iii) non-pecuniary benefits. Hamilton finds that mean and median incomes are lower in self-employment than in wage-employment in the US, although those in the higher income brackets earn more in self-employment than in wage-employment. Hamilton concludes that individuals choose self-employment primarily because of non-pecuniary benefits.

An alternative (or complementary) explanation is that those who choose to become self-employed may be less risk-averse than the typical wage employed. They may even be risk-lovers. In a recent paper by Elston, Harrison, and Rutström (2006), experiments are used to characterize the attitudes to risk among entrepreneurs. Their main finding is that full-time entrepreneurs are less risk-averse and exhibit a significant joy of winning compared to non-entrepreneurs and part-time entrepreneurs.

Yet another explanation relates to the individual's subjective assessment of the probability of success. While Coelho and de Meza (2006) provide experimental evidence that entrepreneurs tend to overestimate their chance of success, Elston, Harrison, and Rutström (2006) do not find systematic judgmental error of profitability. However, it is found that part-time entrepreneurs are reluctant to enter markets where profitability is based on their perception of their relative skill ability.

Evidence from existing Danish questionnaire surveys shows that men focus more on the expected income level than women when choosing occupation, whereas women emphasize non-pecuniary benefits (Statistics Denmark, 1999; and Kjeldsen and Nielsen, 2000). Thus, 90 per cent of the women who had a child in the age of 0-2 years at the time of the business start-up state that an important reason for becoming self-employed was to make it easier to combine family life and work.

With respect to risk aversion, Byrnes, Miller, and Schafer (1999) analyze 150 psychological studies of risk-taking behavior, and find that in 14 out of 16 tasks, women are more risk-averse. However, according to Croson and Gneezy (2004) the evidence of women being more risk-averse is less clear in the economics literature which has typically focused on financial risk.

Several studies have suggested that overconfidence is part of human nature, e.g. Svenson (1980) reports that 90 per cent of Swedish drivers rate themselves above average. Recently and in relation to occupational choice, Niederle and Vesterlund (2006) find from the conduction of experiments that more women than men prefer to work under a non-competitive piece-rate compensation system rather than under a competitive tournament compensation scheme even though women are found to be as productive as men. Niederle and Vesterlund (2006) conclude that the reason for this difference is that men are too overconfident and enjoy competition more. In other words, too many low productivity men enter the competitive tournament, while productive women do not enter enough.

To evaluate whether the self-employed actually are compensated for their risk-taking, individual level information about the expected income (and the expected distribution of income) in both self-employment and wage-employment is required. To obtain this information, we estimate earnings functions for self-employed and wage-employed separately. However, individuals would be expected to select themselves into the type of occupation where they are most productive. Therefore, we estimate earnings functions for each occupational choice, using the dynamic panel data sample selection model of Vella and Verbeek (1998, 1999). This also allows us to disentangle the role of unobserved heterogeneity and state dependence in the occupational choices. We find evidence of state dependence in the occupational choices.

The estimated earnings functions are then in turn used to predict an individual's income (and the uncertainty of this income) in different occupations. The random components of the model are partitioned into transitory and permanent shocks, which in turn are used to create occupational and education specific measures of income variance (uncertainty) and skewness (the risk/chance of very low/high incomes).

Rather that rather than characterizing the entrepreneur, we directly evaluate the impact of earnings on the choice of becoming self-employed, wage-employed or unemployed by examining the roles of expected earnings, risk aversion and over-confidence. This is done for each gender separately. Our results complement existing evidence from experimental economics, providing an potential explanations for the substantial gender gap in the probability of choosing to become self-employed.

We use a large longitudinal data set based on Danish register data from 1980 to 1996, providing us with detailed individual information about income, wealth, education, labor market status (occupation), region of residence, and immigration status. Since the panel covers more than 15 years, we can track long sequences of individual occupational choices and, thereby, appropriately investigate the dynamics of the self-employment choice.

Our results point to a large role for monetary aspects when choosing occupation. As expected, people prefer the sector with the highest expected income and lowest expected variance and, thus, on average appear risk-averse. We find that men put more emphasis on the earnings level, while women appear more risk-averse, which could be one of the crucial reasons why fewer women are self-employed. We do not find evidence of overconfidence. If anything, women instead seem to under-estimate their chance of success compared to men.

The explanatory power of the occupational choice model is quite impressive considering that we only include the predicted income level, variance and skewness. However, we explain much less of the variation in the realized occupational choices for the group of non-western immigrants. Immigrants are interesting with respect to occupational choice since they are more likely to start up their own business than natives. We find that immigrants put much less emphasis on the earnings level. These findings provide additional evidence for immigrants being marginalized into self-employment as Blume, Ejrnæs, Nielsen, and Würtz (2005) suggest. From their analysis on Danish transition data it is found that most non-western immigrants entering self-employment come from unemployment and that they do not use self-employment as a stepping stone for becoming wage-employed.

The rest of the paper is organized as follows: In section 2 we describe the data used in the analysis. In section 3 we formulate the econometric specification. In section 4, we present the results. Section 5 concludes.

2 Data

The data we use in this paper is an unbalanced panel data set for 1980-1996. The data is a representative 10 per cent sample extract drawn from the Integrated Database for labor market Research (IDA) and the Danish Income Registry (IKR) both maintained by Statistics Denmark. IDA and IKR are both longitudinal data based on register data for all individuals in Denmark. Since data originates from administrative records covering the entire Danish population there is only natural attrition in the data, i.e. birth, death and migration of individuals. The occupational status is observed once a year (the last week of November). We divide the labor market status into three states; self-employed, wage-employed, and unemployed. Since the panel covers more than 15 years, we have the possibility to track individuals over long time periods (before, during and after self-employment) and, thereby, appropriately control for the dynamics of the occupational choice. These high-quality Danish data contains very detailed individual information concerning, e.g., income, wealth, education, labor market status, region

FIGURE 1: DISTRIBUTION OF DISPOSABLE INCOME IN 1996

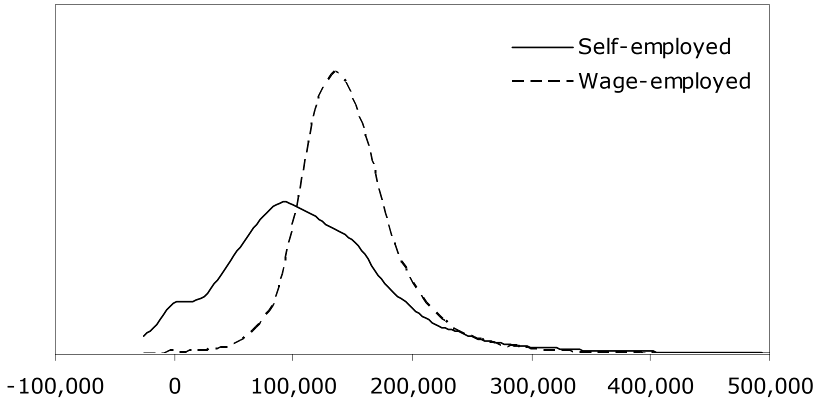


TABLE 1: INCOME DISTRIBUTIONS IN 1996, SELECTED PERCENTILES

Percentile	Personal Income		Disposable Income	
	Self-empl.	Wage-empl.	Self-empl.	Wage-empl.
5	0	111,055	-12,529	85,848
10	7,745	135,380	12,687	99,698
25	75,980	167,222	61,119	119,634
40	120,678	190,022	88,076	133,592
50	148,590	205,274	104,021	142,430
60	183,967	221,686	121,745	151,914
75	242,834	253,018	152,229	169,066
90	345,187	325,042	202,619	202,231
95	458,649	387,028	248,060	229,819

of residence, and immigration status. Moreover, the data also includes the same information for cohabitants allowing us to aggregate variables to the household level.

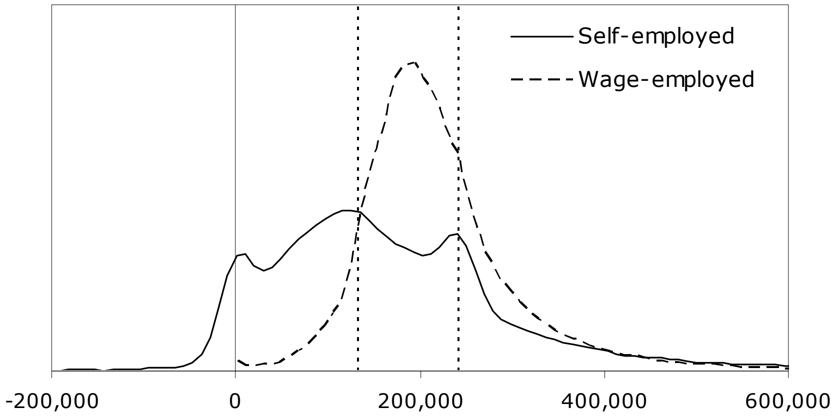
In order to avoid distortions in the results due to retirement patterns and educational attainment we restrict the sample to include persons aged 30-55 years only. This leaves us with $2,424,694$ observations in total of which $1,130,635$ are women.

For the analysis of occupational choice we need to decide on an income measure to use. One obvious candidate is disposable income since this measure is closely related to current consumption possibilities and, hence, utility.[1]

Figure 1 shows kernel densities for the disposable income for self-employed and for wage-employed in 1996. Both distributions are right-skewed with the distribution of incomes from self-employed being most right skewed. From both Figure 1 and Table 1 it can be seen that the mean disposable income for self-employed is considerably below the mean income for wage-employed. However, due to the skewness the 90th percentile earns more in self-employment

[1]We compute the gross income including wage-income, capital income, labor market contributions (since 1994), taxable and non-taxable benefits. In order to obtain the disposable income we subtract the tax payments.

FIGURE 2: DISTRIBUTION OF PERSONAL INCOME IN 1996



than the equivalent in wage-employment.

Figure 1 and Table 1 confirm the US evidence presented in Hamilton (2000), who also finds that mean and median incomes are lower in self-employment, but that those in the higher percentiles earn more in self-employment relative to wage-employment.

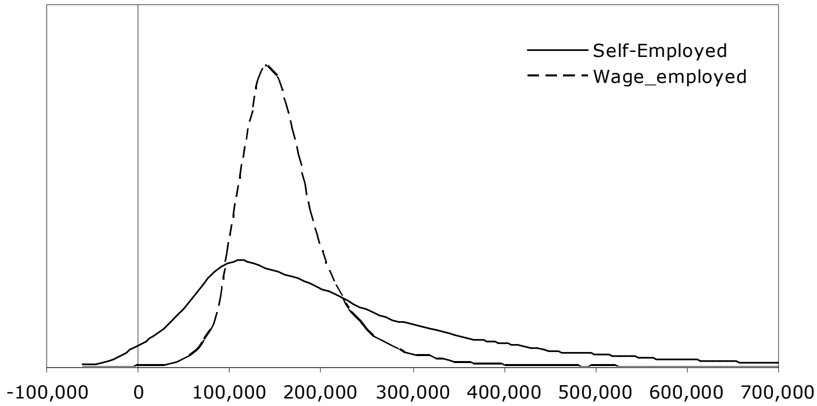
In Figure 2 we have depicted the (taxable) personal income for respectively wage-employed and self-employed together with two dotted vertical lines indicating where the medium and upper tax brackets set in. In contrast to wage-employed, self-employed tend to bunch just below where the tax brackets set in. This can be due to self-employed being in charge of their own working time, but it may also reflect that self-employed are building up inventories and capital stocks or have other means of extracting income from their firm (possibly also in the grey area between firm economics and personal economics). Finally, an institutional feature ("Virksomhedsordningen") allows self-employed to retain earnings in the firm.

The bunching at the tax brackets suggests that adding retained earnings (less of taxes) to the disposable income constitutes a better income measure for self-employed and we only use this income measure in the rest of the paper. As shown in Figure 3 we find that the unconditional mean and median incomes are larger in self-employment than in wage-employment in contrast to the US evidence in Hamilton (2000) and in contrast to when applying the narrow income measure.

3 Econometric Specification

The organization of this section, can be summarized as follows: First, we consider the estimation of conditional earnings functions using Vella and Verbeek (1998, 1999) sample selection model for panel data. Hereafter, we construct income uncertainty and skewness measures. Finally, using measures for expected income, uncertainty and skewness, we model the occupational

FIGURE 3: DISTRIBUTION OF DISPOSABLE INCOME PLUS RETAINED EARNINGS



choice in a conditional logit model.

3.1 Earnings Conditional on Occupational Choice

For each person we separately predict the disposable income including retained earnings from being self-employed and being wage employed. The chosen income measure is disposable income including retained earnings. We use unemployment benefits for the group of unemployed. For each occupation we model earnings as a simple log-linear mincer earnings equation

$$\ln y_{nt}^* = x_{nt}\beta + \alpha_n + \varepsilon_{nt} \tag{1}$$

where n indexes individuals $(n = 1, .., N)$ and t indexes time $(t = 1, ..., T)$; y_{nt}^* is annual disposable income plus retained earnings, β is a vector of unknown coefficients to be estimated, x_{nt} is a vector covariates, α_n represents unobserved heterogeneity and ε_{nt} is a normally distributed disturbance.

Since we observe earnings for the chosen occupational status only, the conditional earnings functions will in general be estimated on a non-random selected sample. There are several arguments, why self-selection may be an issue in the present context. In the Roy (1951) model the individual ex-ante knows her sector-specific productivity, and will select herself into the sector, where she is most productive. Furthermore, if the incomes in the two sectors are highly correlated, the most productive persons will select the sector with the largest dispersion of sector specific abilities, while the least productive will select the sector with the smallest dispersion.

The Danish labor market is characterized by a compressed wage structure as a consequence of the generous unemployment benefit level and a high degree of organization on both employer and worker sides. As argued by Malchow-Møller, Markusen, and Skaksen (2005) such institutional arrangement may well imply that the most productive are not paid according to their marginal product and, therefore, the most able may select themselves into self-employment. On the other hand, the least productive may not have a sufficiently high productivity to earn the minimum wage in paid employment. Consequently, marginalization may also push the least productive into self-employment. Blume, Ejrnæs, Nielsen, and Würtz (2005) argue that this indeed is the case for non-western immigrants in the Danish labor market.

Yet another type of selection, ex-post self-selection, arises in leaning models such as Jovanovic (1979) and Jovanovic (1984), where persons have no ex-ante knowledge of their productivity, but consecutively observe output realizations. Persons experiencing poor output realizations will quit and search for a new match.

To control for the selection problem we use the Vella and Verbeek (1998, 1999) dynamic panel data application of Heckman's two-step sample selection model. The selection is modelled as a dynamic random effects probit, which allows us to separate two sources of persistence in the occupational choice: Persistence as a result of unobserved heterogeneity and (true) state dependence. Since we do not observe the first occupational choice, we cannot assume that the initial observation of the occupational is truly exogenous. We use the Wooldridge (2005) way of handling the initial conditions problem and, thus, allow the unobserved heterogeneity to be correlated with the initial dependent variable.

We will now briefly explain the model.[2] We consider a model consisting of two equations, where the parameters of equation (1) are of primary interest, while the selection equation below is a reduced form equation for the occupational choice. The selection part of the model can be summarized as

$$\begin{aligned} d_{nt}^* &= x_{nt}\gamma + \phi d_{nt-1} + \mu_n + \eta_{nt} && (2)\\ d_{nt} &= \mathbf{1}\left(d_{nt}^* > 0\right) && (3)\\ \ln y_{nt} &= \ln y_{nt}^* \text{ if } d_{nt} = \mathbf{1} && (4)\\ &= 0 \text{ (unobserved) otherwise} \end{aligned}$$

where y_{nt}^* and d_{nt}^* are latent endogenous variables with observed counterparts y_{nt} and d_{nt}.

The equation of interest is assumed to have the usual error component structure, where $\alpha_n \sim iN\left(0, \sigma_\alpha^2\right)$ and $\varepsilon_{nt} \sim iN\left(0, \sigma_\varepsilon^2\right)$. For the selection equation we allow for unobserved heterogeneity through random individual effects, such that the selection equation has the following two-component error structure $\mu_n \sim iN\left(0, \sigma_\mu^2\right)$ and $\eta_{nt} \sim iN\left(0, \sigma_\eta^2\right)$. We allow for correlation between the individual effects as well as correlation between the idiosyncratic disturbances, that is $\operatorname{cov}(\alpha_n, \mu_n) = \sigma_{\alpha\mu} \neq 0$ and $\operatorname{cov}(\varepsilon_{nt}, \eta_{nt}) = \sigma_{\varepsilon\eta} \neq 0$. Finally, denote $\xi_{nt} = \alpha_n + \varepsilon_{nt}$, $\upsilon_{nt} = \mu_n + \eta_{nt}$, $\mathbf{x}_n = [x_{n1}, ..., x_{nT}]'$ and let $\boldsymbol{\upsilon}_n$ be a T vector of υ_{nt}.

[2]For a detailed treatment of the model see Vella and Verbeek (1999).

Assume now

$$\boldsymbol{\upsilon}_n|x_n \sim iN\left(0, \sigma_\mu^2 \mathbf{ii}' + \sigma_\eta^2 \mathbf{I}\right) \tag{5}$$

$$E\left[\xi_{nt}|x_n, \upsilon_n\right] = \tau_1 \upsilon_{nt} + \tau_2 \bar{\upsilon}_n \tag{6}$$

where $\bar{\upsilon}_n = T^{-1}\sum_{t=1}^{T}\upsilon_{nt}$ and where $\tau_1 = \sigma_{\varepsilon\eta}/\sigma_\varepsilon^2$ and $\tau_2 = T\left(\sigma_{\alpha\mu} - \sigma_{\varepsilon\eta}\sigma_\mu^2/\sigma_\varepsilon^2\right)/\left(\sigma_\eta^2 + T\sigma_\mu^2\right)$ are constants to be estimated and $\mathbf{i}$ is a column of ones. Note that equation (6) imposes strict exogeneity of x_{nt}, such that errors are assumed to be independent of future and lagged values of x_{nt}. To estimate the conditional mean for the dependent variable in the equation of interest, we condition on the chosen occupation

$$E\left[\ln y_{nt}|x_n, d_{n0}, d_n\right] = x_{nt}\beta + E\left[\xi_{nt}|x_n, d_{n0}, d_n\right]$$

where $E\left[\xi_{nt}|x_n, d_{n0}, d_n\right]$ is the selection bias induced by correlation between the errors in the two equations.

Under these assumptions, it can be shown that the conditional mean of the error-term from the selection equation, $E\left[\upsilon_{nt}|x_n, d_{n0}, d_n\right]$ can be estimated by the following expression

$$\tilde{\upsilon}_{nt} = \frac{1}{\int f\left(d_n|x_n, \mu_n\right) f\left(\mu_n\right) d\mu_n} \int \left(\mu_n + E\left[\eta_{nt}|x_n, \mu_n\right]\right) f\left(d_n|x_n, \mu_n\right) f\left(\mu_n\right) d\mu_n \tag{7}$$

This expression can be approximated by quadrature methods or simulation. Once we have estimated the reduced form parameters for the selection equation, we can easily simulate the conditional error $\tilde{\upsilon}_{nt}$.[3]

After computing $\tilde{\upsilon}_{nt}$ and the individual specific means $\bar{\upsilon}_n = \frac{1}{T_n}\sum_t^{T_n}\tilde{\upsilon}_{nt}$ we can estimate the following equation by the simple linear random effects model

$$\ln y_{nt} = x_{nt}\beta + \tilde{\upsilon}_{nt}\theta_1 + \bar{\upsilon}_n\theta_2 + \alpha_n + \varepsilon_{nt}$$

3.2 Uncertainty and Skewness Measures

For each category in our disaggregated education breakdown shown in Table A.1 we estimate the occupational-specific measures of variance and skewness of the income processes. This is done separately for men and women.

We divide the uncertainty into a permanent part relating to the variance of the individual time-constant α_n and into a transitory uncertainty relating to the time-varying error-term.[4]

[3] The procedure is summarized in algorithm 1 in the appendix.

[4] Recently, Diaz-Serrano, Hartog, and Nielsen (2003) have used a similar approach in the context of educational choice.

Among the covariates in x_{nt} we have included 28 educational dummies.[5] We define $a_n = \exp(\alpha_n)$ and $e_{nt} = \exp(\varepsilon_{nt})$ and compute the variance R and the skewness K for each education type l by

$$R_l^{\alpha} = \frac{1}{N_l}\sum_{n=1}^{N_l}(a_{nl} - \bar{a}_l)^2 \qquad K_l^{\alpha} = \frac{1}{N_l}\sum_{n=1}^{N_l}(a_{nl} - \bar{a}_l)^3$$

$$R_l^{\varepsilon} = \frac{1}{T}\frac{1}{N_l}\sum_{n=1}^{N_j}(e_{nlt} - \bar{e}_l)^2 \qquad K_l^{\varepsilon} = \frac{1}{T}\frac{1}{N_l}\sum_{n=1}^{N_l}(e_{nlt} - \bar{e}_l)^3$$

By averaging the residuals only on education groups, we effectively assume that the income uncertainty does not depend on for example experience, which is obviously an approximation. Averaging the incomes on other variables as well is not feasible with the detailed education break-down used.

For an unemployed there is no or very little uncertainty regarding income. Consequently, we set the variance and skewness equal to zero.

3.3 A Model of Occupational Choice

The behavioral framework underlying the occupational choice model is simple: We assume that individuals each period associate each occupation with a continuous random utility function, U_{nit} , where each occupation is indexed by $i \in [se, we, ue]$. Each period individuals choose between self-employment (se), wage-employment (we) and unemployment (ue) to maximize the U_{nit}.[6] Random utility is assumed to be a linear function of occupational specific earnings, and the variance and skewness of permanent and transitory income shocks. Hence, the random utility function can be written as

$$U_{nit} = \mathbf{x}_{nit}\boldsymbol{\beta} + \delta_i + \epsilon_{nit} \quad \text{with} \quad n = 1, ..., N \quad \text{and } t = 1, .., T$$

where δ_i is a choice-specific constant, $\mathbf{x}_{nit} = \left[\hat{Y}_{nit}, R_l^{\alpha}, K_l^{\alpha}, R_l^{\varepsilon}, K_l^{\varepsilon}\right]$ denotes the set of attributes associated with each occupation, $\boldsymbol{\beta}$ is a vector of coefficients related to the the choice specific attributes $\mathbf{x}_{int}$. The error component ϵ_{nit} is assumed to be individual-, choice,- and time specific and distributed according to a Type I extreme value distribution. With this distributional assumption, we end up with McFadden's well known Conditional Logit model for discrete choices.

[5] In the IDA database there are 1,750 different educations, but in order to secure representativity we operate with 28 education groups only (see Table A.1 in the Appendix). We have aimed at securing representativity by not making a too disaggregated educational break-down, but on the other hand aimed at selecting as homogeneous groups as possible.

[6] Even though each individual maximizes utility each period by choosing occupations this need not be equivalent to maximization of life-time utility given by a discounted sum of period utility. However, this simplification is needed to make the model operational.

4 Results

4.1 Self-selection and Earnings Differentials

In this section, we investigate the extent to which earnings differentials can be explained by individuals self-selecting themselves into the different occupations. To account for the potentially important selection problems, we estimate the model sample selection model of Vella and Verbeek (1998,1999). First, we estimate the selection equation given by equation (2) and equation (3) by a dynamic random effects probit. Hereafter, we estimate the parameters in the conditional earnings function in equation (1).

Since the choice of labor market state differs considerably between the genders, the sample correction and the prediction of incomes are done separately for men and women. Additionally, the existence of wage differentials between the genders suggests that it would be appropriate to run the wage equations separately.

The results from the selection equations given in table A.2 suggest that the impact of the lagged dependent variable is positive and highly significant, indicating the presence of substantial state dependence. State dependence can be a result of cost of and uncertainty of labor market transitions and is likely to be amplified for transitions into self-employment in the presence of start-up costs. In an intertemporal model of occupational choice Schjerning (2005) shows that the combination of irreversible start-up costs and income uncertainty introduces an option value of being self-employed. To avoid potential start-up costs associated with later re-entry, the self-employed is willing to wait until good times occur rather than temporarily leaving self-employment. This introduces a value of waiting and consequently we will see later entry and later exit.

The magnitude of state dependence for the self-employed is substansial: Being self-employed in the previous period increases the probability of being self-employed in the current period from 1.2 to 41.5 per cent for females and from 4.0 per cent to 48.7 per cent for men. As a comparison, the marginal effect from previous wage-employment is 19.9 per cent for females and 24.0 per cent for men.

The results from the selection equation suggest that, in general, the probability of being self-employed varies much between the educational categories both with respect to length and type of education. Although the picture is quite mixed, it seems to be the case that unskilled and some groups of highest education are the most likely to become self-employed. The latter is due to the fact that the self-employed include professionals such as practitioner doctors, dentists, lawyers and accountants.

The estimated earnings equations are given in table A.3. As dependent variable we use the disposable income including retained earnings. We allow for unobserved heterogeneity in the form of random effects, and we control for the usual socio-demographic variables. We find positive coefficients on marriage for men, while they are negative for women. The origin variables have the expected signs and magnitudes, i.e., non-western immigrants earn considerably less than western immigrants, second generation immigrants and natives. It is striking that non-western immigrants are more likely to become self-employed even though they should expect a much lower income in self-employment compared to wage-employment.

We find the usual hump-shaped effect of age, which obviously captures labor market experience. We have included dummies for each education from our detailed educational break-down shown in the appendix. The general picture is as expected that the longer education, the higher disposable income. As one would presume, the returns to education differ remarkably between the educations. For example, the returns to humanities are lower compared to social sciences at each length of education reflecting the relatively higher unemployment rate that may lead to accepting jobs below the educational level.

If education is a signal, so that employers use education to screen potential workers, we would expect lower returns to education in self-employment. There does not seem to be much evidence for the signalling hypothesis.

Since we do not wish to rely on the non-linearity of the selection equation to identify the selection effects in the income equations we need to exclude at least one variable. We use the lagged dependent variable, household wealth, dummies for children in the household and a dummy for the spouse being self-employed.

The inclusion of the correction terms account for the selection bias induced by the correlation between unobservables in the selection model and earnings equations. The coefficients to the correction terms υ_{nt} and $\bar{\upsilon}_n$ are statistically significant in all four regressions. In the case of men, the coefficient on both correction terms are negative, implying that the marginalization on average dominates. Taken literally, we have that those in wage-employment will tend to earn more in self-employment than those already self-employed.

In contrast to this, the coefficient on the individual specific correction term, $\bar{\upsilon}_n$ is positive in the self-employment earnings equation for women implying that those in wage-employment have a lower self-employment potential than the currently self-employed. Since the income is measured on a yearly basis a possible explanation for the positive selection into self-employment relates to differences in the hours of work between wage-employed and self-employed. About 20 per cent of female wage-employed work part-time and if this fraction is larger than the corresponding for women in self-selection a positive selection into self-employment will, on average, emerge. In recent work by Carrasco and Ejrnæs (2003) it is in fact argued that the relative low share of female self-employed in Denmark can be explained by the relative high level of flexibility in the Danish labor market providing the possibility to work part time in paid employment. Similar arguments apply to women planning to have children, as the opportunities for paid maternity leave are better in wage-employment. Another explanation

TABLE 2: CHOICE OF LABOR MARKET STATUS (CONDITIONAL LOGIT MODEL)

			Distribution of Permanent Chock's				MRS			
Subsample	Mean Earnings		Variance		Skewness		Variance/ Mean	Skewnes/ Mean	Sample size	R^2
All	1.58	(0.007)**	-0.15	(0.004)**	-0.02	(0.001)**	-0.10	-0.010	7,274,082	51
Non-western immigrants	1.08	(0.044)**	0.05	(0.025)	-0.02	(0.005)**	0.04	-0.016	119,412	23
Women	0.40	(0.017)**	-0.22	(0.011)**	-0.04	(0.001)**	-0.56	-0.091	3,391,905	59
Non-western immigrants	0.81	(0.095)**	-0.34	(0.070)**	0.02	(0.009)*	-0.42	0.022	45,564	29
Married	1.00	(0.023)**	-0.27	(0.014)**	-0.03	(0.002)**	-0.27	-0.032	2,514,129	61
HH. Wealth(t-1)>500.000	0.85	(0.048)**	-0.14	(0.022)**	-0.01	(0.003)**	-0.16	-0.013	493,101	64
Father self-employed	1.29	(0.148)**	0.20	(0.030)**	-0.06	(0.009)**	0.16	-0.043	64,728	64
age<40	0.40	(0.026)**	-0.11	(0.013)**	-0.05	(0.002)**	-0.27	-0.118	1,353,195	60
age>45	0.58	(0.030)**	-0.37	(0.025)**	-0.02	(0.002)**	-0.63	-0.038	1,148,580	57
Men	2.81	(0.016)**	-0.10	(0.005)**	0.00	(0.001)	-0.04	-0.001	3,882,177	44
Non-western immigrants	1.81	(0.088)**	0.14	(0.027)**	-0.03	(0.006)**	0.08	-0.018	73,848	20
Married	2.39	(0.020)**	-0.15	(0.006)**	0.01	(0.001)**	-0.06	0.003	2,761,398	48
HH. Wealth(t-1)>500.000	1.42	(0.044)**	-0.18	(0.012)**	0.06	(0.002)**	-0.13	0.045	474,390	43
Father self-employed	3.77	(0.110)**	-0.10	(0.023)**	-0.02	(0.007)**	-0.03	-0.006	91,098	40
age<40	3.33	(0.027)**	-0.07	(0.007)**	-0.02	(0.002)**	-0.02	-0.006	1,538,208	50
age>45	2.53	(0.027)**	-0.10	(0.009)**	0.01	(0.002)**	-0.04	0.004	1,348,191	39

Notes: Standard errors are in parentheses. * significant at 5%; ** significant at 1%. Other controls: Occupational specific constants and measures of the temporary components of estimated chocks (skewness and variance)

might be glass-ceiling effects in wage-employment, see e.g. Albrecht, Björklund, and Vroman (2003) for Swedish evidence.

4.2 The Occupational Choice

The occupational choice model is estimated for a several different subsamples. The results from these estimations are shown in Table 2. Each row in the Table corresponds to the results for a different subsample. The figures in the Table show the effects of the mean, variance (uncertainty) and skewness (i.e., in this case the chance of very high incomes) of predicted earnings conditional on the occupational choice. Note that in the estimations, variance and skewness of both transitory and permanent shocks are included in the model. For expositional purposes, however, we only report variance and skewness of the permanent income component in the Table.[7]

The coefficients to mean earnings gives the marginal utility of expected income, while the

[7] Since the earnings equations were estimated with age variables and time-dummies there is no aggregate time variation left in the error-terms, but still individual specific variation occurs. Alvarez, Browning, and Ejrnæs (2002) find that Danish income processes are particularly heterogenous. Steep income-experience profiles imply a large variance, but when controlling for the income level, we should due to income smoothing expect that a flat income-tenure profile is preferred. However, a steeper wage-experience profile may indicate greater possibilities such as promotion for wage-employed and business expansion for self-employed, which may explain the positice coefficient to the variance of the temporary income shocks. The skewness of the time-varying part does not seem to play any role.

(negative of the) coefficient to the variance can be interpreted as the marginal (dis-) utility of income uncertainty. To give an example, a positive coefficient to expected income is associated with individuals consistently choosing occupations with higher levels of expected earnings, while a negative coefficient to variance emerges when individuals choose occupations with little income uncertainty.

To make results comparable across different subsamples, we compute the marginal rate of substitution (MRS) between the variance and mean earnings and between the skewness and mean earnings. The MRS can be interpreted as the rate at which you are willing to trade off more uncertainty for higher income.[8] These results are shown in the right part of the Table.

Considering the full sample (the first row in the Table), the results point to a large role for monetary aspects when choosing occupation. As expected, people's choice of occupation is positively affected by expected (mean) earnings and negatively by a higher variance of the income. Thus, on average, people appear risk averse. These findings are found to be robust to various sample decompositions.

Turning to the differences between the genders, we find that men put more emphasis on the earnings level, while women appear more risk averse. This is reflected in the much lower value of the MRS estimate for women. This could be one of the main reasons why fewer women choose to become self-employed.

Women seem to be behaving in a less risk averse manner when household wealth exceeds DKK 500,000. This is perfectly consistent with standard models of intertemporal behavior that find that the degree of effective risk aversion is decreasing in wealth; see, e.g., Deaton (1991), Carroll (1997), and Schjerning (2005).

The finding that married women appear less risk averse than other women is also fully consistent with models from the literature on family economics that point to risk sharing as being a potentially important economic gain from marriage, see e.g. Hess (2004).

A similar variation in men's attitudes towards risk is not found. An interesting finding, however, is that the MRS between income uncertainty and expected earnings is virtually zero compared to females. This confirms the evidence from Danish questionnaires, referred to above, which pointed to men putting much more emphasis on monetary gains (expected income) than women.

Finally, a positive coefficient to skewness is interpreted as being consistent with evidence of overconfidence. If people systematically prefer occupations with a high degree of skewness (a chance of very high incomes) it may be due unrealistic, strong beliefs in their own ability.

For the full sample, a negative coefficient to skewness is found. Hence, on average, there is no evidence of overconfidence. The more detailed results with respect to this behavioral

[8] This normalization is important since estimates from two different subsamples are not directly comparable due to differences in the variances of the unobserved factors.

hypothesis are mixed and inconclusive. If anything, men behave somewhat more overconfidently than females. This result match those found from experimental studies.

It is striking that we in the model for immigrants only can explain 17 per cent compared to 50 per cent in the other models. Moreover, the coefficient to income is much lower than in the other conditional logit models. Hence, other important (unobserved) factors, such as lack of opportunities in the ordinary labor market and non-pecuniary benefits may be much more relevant in explaining their occupational choice. Hence, the low explanatory power, and the lower coefficient to income points to self-employment being the last resort due to marginalization in wage-employment. We also find that non-western immigrants appear less risk averse. This may be due to marginalization forcing immigrants to accept insecure and low paid occupations, but it can also be a consequence of cultural differences in the attitudes towards self-employment.

5 Conclusions

This paper uses high quality Danish longitudinal register data, to investigate the relationship between self-employment choice, expected earnings and income uncertainty. We proceed in the following steps: Firstly, we estimate of conditional earnings functions using the sample selection model of Vella and Verbeek (1998, 1999). Secondly, using measures for expected income, uncertainty and skewness, we model the occupational choice in a conditional logit model.

Comparing earnings distributions based on different income measures, we find that i) the dispersion of incomes is in general much larger for the self-employed and ii) Danish self-employed earn more than wage-employed when retained earnings are included in the income measure. Contrary to wage-workers, self-employed (taxable) personal income bunch at kink points in the tax system since self-employed (unlike wage workers) has the possibility to retain earnings and thereby transfer income across years. The progressive Danish income tax system provides strong incentives to make such transfers.

Several experimental studies have found that while men are more competitive, women are more risk averse. In the context of occupational choice, we find that men put more emphasis on the income level, while women seem to be more risk-averse. This result is found to be robust to various sample decompositions.

Linking the behavioral results from the experimental literature with income distributions in self-employment and wage-employment may explain why fewer women become self-employed. We find that part of the gender gap can be explained by gender differences in the trade-offs between income level and the variance of incomes. However, we find no effect of skewness of incomes.

Non-western immigrants are overrepresented in self-employment. The occupational choice model performs considerably worse for this group and we find smaller effects of income level and variance. Furthermore, the sample selection model shows that non-western are more likely to become self-employed even though they should expect a much lower income in self-employment than native Danes. This suggests that non-western immigrants are marginalized into self-employment

References

ALBRECHT, J., A. BJÖRKLUND, AND S. VROMAN (2003): "Is There a Glass Celing in Sweden?," *Journal of Labor Economics*, 21(1), 145–177.

ALVAREZ, J., M. BROWNING, AND M. EJRNÆS (2002): "Modelling Income Processes with Lots of Heterogeneity," CAM Working Paper 2002-01.

BHAT, C. (2001): "Quasi-Random Maximum Simulated Likelihood Estimation of the Multinomial Mixed Logit Model," *Transportation Research*, B 35, 677–693.

BLUME, K., M. EJRNÆS, H. S. NIELSEN, AND A. WÜRTZ (2005): "Labor Market Transitions of Immigrants with Emphasis on Marginalization and Self-Employment," working paper.

BYRNES, J. P., D. C. MILLER, AND W. D. SCHAFER (1999): "Gender Differences in Risk-Taking: A Meta-Analysis," *Psychological Bulletin*, 125, 367–383.

CARRASCO, R., AND M. EJRNÆS (2003): "Self-Employment in Denmark and Spain: Institution, Economic Conditions and Gender Differences," CAM Working Paper 2003-06.

CARROLL, C. D. (1997): "Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quaterly Journal of Economics*, 112(1), 1–56.

COELHO, M., AND D. DE MEZA (2006): "Self Deception, Self Selection, Self Destruction: An Experimental Investigation of Adverse Selection," working paper.

CROSON, R., AND U. GNEEZY (2004): "Gender Differences in Preferences," working paper.

DEATON, A. (1991): "Savings and Liquidity Constraints," *Econometrica*, 59, 1121–1248.

DIAZ-SERRANO, L., J. HARTOG, AND H. S. NIELSEN (2003): "Compensating Wage Differentials for Schooling Risk in Denmark," IZA Discussion Paper no. 963.

ELSTON, J. A., G. W. HARRISON, AND E. E. RUTSTRÖM (2006): "Characterizing the Entrepreneur Using Field Experiments," working paper.

HAMILTON, B. H. (2000): "Does Entrepreneurhip Pay? An Empirical Analysis of the Returns to Self-Employment," *Journal of Political Economy*, 108, 604–631.

HESS, G. D. (2004): "Marriage and Consumption Insurance: What's Love Got to Do with It?," *Journal of Political Economy*, 112, 290–318.

JOVANOVIC, B. (1979): "Job Matching and theTheory of Turnover," *Journal of Political Economy*, 87, 972–990.

——— (1984): "Matching, Turnover and Unemployment, Journal of Political Economy," *Journal of Political Economy*, 92, 108–122.

MALCHOW-MØLLER, N., J. R. MARKUSEN, AND J. R. SKAKSEN (2005): "Labour Market Institutions, Learning, and Self-Employment," CEBR discussion paper 2006-24.

NIEDERLE, M., AND L. VESTERLUND (2006): "Do Women Shy Away from Competition? Do Men Compete Too Much?," working paper.

ROY, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

SCHJERNING, B. (2005): "A Dynamic Model for Occupational Choice, Savings and Formation of Entrepreneurial Human Capital under Liquidity Constraints and Income Uncertainty," working paper.

STATISTICSDENMARK (1999): "Generel Erhvervsstatistk," Statistiske Efterretninger 1999:19.

SVENSON, O. (1980): "Are We All Less Risky and More Skillful Than our Fellow Drivers?," *Acta Psychologica*, 47, 143–148.

TRAIN, K. (2003): *Discrete Choice Models with Simulation*. Cambridge University Press, Cambridge.

VELLA, F., AND M. VERBEEK (1998): "Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men," *Journal of Applied Econometrics*, 13(2), 163–183.

VELLA, F., AND M. VERBEEK (1999): "Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias," *Journal of Econometrics*, 90, 239–263.

WOOLDRIDGE, J. M. (2005): "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Effects," *Journal of Applied Econometrics*, 20(1), 39–54.

.

6 Appendix

Algorithm 1 *Estimation of conditional error-term from the selection equation,* $E\left[\upsilon_{nt}|x_{nt},d_{n0},d_{nt}\right]$

1. *For a given set of parameter values* $\theta_1 = (\gamma, \phi, \sigma_\mu)$ *take a draw from* μ_n^r *from* $f(\mu_n|\sigma_\mu) = N(0, \sigma_\mu)$ *and calculate the likelihood for individual i conditional on the draw*

$$f\left(d_n, d_{i0}|x_{n,}\mu_n^r\right) = \prod_{t=1}^{T_n} f\left(d_{nt}|x_{nt,}\mu_n^r\right) f\left(d_{n0}|x_{nt,}\mu_n^r\right)$$

where $f\left(d_{nt}|x_{nt,}\mu_n^r\right) = \Phi_{nt}d_{nt}+(1-\Phi_{nt})(1-d_{nt})$ *and where* $\Phi_{nt} \equiv \Phi\left(x_{nt}\gamma + \phi d_{nt-1} + \mu_n^r\right)$

2. *Repeat many times and average the results to obtain the Simulated Log Likelihood function (SLL)*

$$SLL = \ln \frac{1}{R}\sum_{r}^{R} f\left(d_n, d_{i0}|x_{n,}\mu_n^r\right)$$

3. *Choose* θ_1^{MSL} *so that SLL is maximized*

4. *Given the MSL estimates from the fist stage regression* θ_1^{MSL}, *we can easily simulate* $\hat{\upsilon}_{nt}$. *Take R draws from* $f\left(\mu_n|\sigma_\mu^{MSL}\right)$ *and calculate the simulated counterpart of* $\hat{\upsilon}_{nt}$

$$\tilde{\upsilon}_{nt} = \frac{1}{\frac{1}{R}\sum_r^R f\left(d_n, d_{i0}|x_{n,}\mu_n^r\right)} \frac{1}{R}\sum_{r}^{R}\left(\mu_n^r + E\left[\eta_{nt}|x_n, \mu_n^r\right]\right) f\left(d_n, d_{i0}|x_{n,}\mu_n^r\right)$$

where $E\left[\eta_{nt}|x_n, \mu_n^r\right] = \frac{d_{nt}\phi_{nt}}{\Phi_{nt}}$-$\frac{(1-d_{nt})\phi_{nt}}{1-\Phi_{nt}}$ *is the cross-sectional generalized residual for the probit model and where* $\phi_{nt} \equiv \phi\left(x_{nt}\gamma + \phi d_{nt-1} + \mu_n^r\right)$

To improve coverage of the integrals and reduce simulation noise, we use Halton Draws.[9] .

[9] Halton draws provides a superior coverage as it induces negative correlation across individuals. In the context of discrete choice models, Bhat (2001) found in a Mixed Logit Model, that 100 Halton draws provided more precise results than 1000 standard pseudo random draw. Train (2003) provide a comprehensive and excellent treatment of several variance reduction techniques.

TABLE A.1: MEAN, VARIANCE AND SKEWNESS (EDUCATIONAL BREAKDOWN)

		# observations		Mean disposable income		Variance				Skewness			
						Transitory effect		Permanent effect		Transitory effect		Permanent effect	
		se	we	se	we	se	we	se	we	se	we	se	we
Missing education		3,867	29,167	209,691	163,169	35.633	0.049	2.196	0.224	47.9	11.8	6.1	3.9
Primary School	*Basic school*	80,596	657,151	217,842	133,278	35.476	0.037	2.296	0.099	236.9	36.9	25.8	7.5
Secondary school	*General*	4,509	43,962	267,294	166,541	0.660	0.039	3.216	0.223	25.5	2.1	6.9	3.6
	Commercial and technical	1,362	12,906	353,176	218,601	1.062	0.100	6.970	0.806	11.9	20.8	6.9	13.1
Vocational training	*Shop assistents*	24,343	345,545	232,634	151,299	2.342	0.034	0.974	0.122	104.9	9.8	5.3	6.2
	Building and construction	18,769	113,541	225,676	170,977	0.571	0.024	0.554	0.072	28.6	5.1	7.5	4.4
	Metal	17,887	147,347	240,837	177,130	0.219	0.028	0.784	0.075	11.2	19.9	5.7	4.5
	Graphic	2,061	18,304	277,758	202,616	3.609	0.029	2.278	0.161	22.1	4.5	7.2	8.3
	Technical	2,606	33,878	158,173	129,978	1.342	0.035	1.147	0.097	18.8	14.5	3.0	2.7
	Service and transport	10,725	23,617	124,194	131,480	1.225	0.051	0.478	0.095	42.1	22.1	1.7	1.5
	Food	20,383	49,809	314,211	170,244	1.155	0.032	0.764	0.091	83.8	5.1	5.3	4.7
	Health care	2,254	67,273	138,455	117,762	0.659	0.026	0.896	0.052	14.2	5.2	3.3	1.2
Post secondary	*Humanities and social sciences*	1,576	18,953	185,609	153,600	0.224	0.039	2.013	0.141	2.1	2.5	3.6	4.5
	Technical	3,847	32,146	246,378	181,377	0.202	0.040	0.594	0.112	3.7	15.2	3.0	6.7
	Agriculture	917	8,504	306,245	157,736	0.218	0.039	0.697	0.066	3.2	5.1	1.1	1.1
	Health care	206	10,378	114,543	129,392	1.163	0.024	0.812	0.051	4.5	2.0	0.3	0.9
	Police and defence	327	15,524	234,734	202,600	0.194	0.020	0.621	0.079	3.7	6.2	3.2	3.8
Higher education short cycle	*Humanities*	3,033	157,352	174,274	156,319	7.537	0.021	0.893	0.054	48.4	2.6	3.1	1.5
	Social sciences	1,780	22,332	491,580	251,898	0.251	0.073	1.292	0.270	3.2	15.7	2.5	3.8
	Technical	3,338	40,915	308,715	246,404	0.483	0.041	1.327	0.126	17.3	8.7	8.4	10.5
	Health care	1,962	65,014	168,011	137,621	0.095	0.026	0.659	0.057	2.0	2.7	4.8	4.7
	Food, agriculture and transport	730	15,661	276,414	216,992	0.883	0.032	2.391	0.073	19.7	2.4	5.1	1.7
	BA	469	5,205	336,057	251,710	0.828	0.087	1.234	0.320	8.4	11.9	1.5	2.4
Higher education MA level	*Humanities*	578	22,592	209,513	190,457	0.162	0.038	0.868	0.068	2.3	26.6	1.7	1.5
	Natural sciences	159	9,727	226,141	217,852	0.179	0.023	1.104	0.070	2.8	1.2	4.0	3.8
	Social sciences	3,589	25,269	476,302	251,683	0.321	0.047	0.778	0.190	19.3	6.9	3.5	5.4
	Technical	1,912	18,252	334,441	255,821	0.553	0.048	2.193	0.161	7.1	17.9	10.3	7.0
	Food	1,431	5,250	358,845	228,455	0.122	0.024	0.446	0.073	5.8	2.1	7.6	2.4
	Health care	7,069	18,412	441,567	260,069	0.074	0.038	0.199	0.085	2.4	13.1	3.0	1.3

TABLE A.2: SELECTION EQUATIONS (RESULTS FROM A BINARY PROBIT WITH RANDOM EFFECTS)

		Males				Females			
		Self-employment		Wage-employment		Self-employment		Wage-employment	
		Coefficient	Std.	Coefficient	Std.	Coefficient	Std.	Coefficient	Std.
Lagged dependent, y(t-1)		2.590	0.007	1.415	0.004	2.651	0.010	1.226	0.005
Initial dependent, y(0)		1.962	0.018	1.918	0.010	1.641	0.023	1.540	0.011
Age		0.705	0.071	-0.157	0.044	0.731	0.096	0.974	0.047
Age squared		-0.827	0.082	0.067	0.050	-0.822	0.110	-1.231	0.053
Wealth (in mio dkr, 1996 prices)		0.014	0.002	0.001	0.001	0.005	0.002	0.008	0.001
No. of children aged 0-6		0.049	0.006	-0.019	0.004	0.049	0.010	-0.098	0.006
No. of children aged 7-17		0.030	0.004	-0.011	0.003	0.012	0.006	0.001	0.003
Married		-0.006	0.009	0.112	0.007	-0.047	0.011	0.064	0.007
Immigrant (western)		0.043	0.030	-0.155	0.023	0.132	0.032	-0.159	0.021
Immigrant (non-western)		0.230	0.028	-0.481	0.020	0.124	0.037	-0.529	0.024
Second generation immigrants		0.092	0.084	-0.111	0.070	0.132	0.109	0.037	0.079
Spouse self-employed		0.453	0.013	0.137	0.005	0.399	0.011	0.152	0.005
Regional	*Copenhagen*	0.009	0.017	-0.131	0.012	0.033	0.022	-0.097	0.013
dummies	*Large city*	0.012	0.016	-0.096	0.012	0.022	0.020	-0.161	0.012
	Rural	0.097	0.013	-0.123	0.009	0.120	0.015	-0.197	0.009
Missing education		0.066	0.029	-0.030	0.023	0.152	0.037	-0.109	0.025
Secondary	*General*	0.061	0.026	0.073	0.020	0.133	0.034	0.093	0.022
school	*Commercial and technical*	0.070	0.043	0.153	0.038	0.002	0.070	0.190	0.042
Vocational	*Shop assistents*	-0.038	0.014	0.176	0.012	-0.023	0.013	0.155	0.008
training	*Building and construction*	-0.025	0.014	0.067	0.011	0.180	0.084	-0.011	0.067
	Metal	-0.072	0.014	0.152	0.011	0.403	0.139	0.063	0.103
	Graphic	-0.025	0.036	0.041	0.028	0.252	0.082	-0.182	0.055
	Technical	-0.010	0.046	0.087	0.033	0.071	0.029	0.000	0.019
	Service and transport	0.188	0.041	-0.145	0.032	0.500	0.025	-0.370	0.020
	Food	0.175	0.018	-0.156	0.015	0.110	0.057	0.051	0.037
	Health care	-0.143	0.086	0.418	0.066	-0.099	0.022	0.351	0.013
Post	*Humanities and social sciences*	0.171	0.069	0.007	0.059	0.176	0.036	0.063	0.023
secondary	*Technical*	-0.056	0.026	0.191	0.021	0.057	0.054	0.223	0.035
	Agriculture	-0.050	0.060	0.126	0.047	-0.172	0.084	0.439	0.054
	Health care	-0.751	0.361	0.478	0.153	-0.204	0.064	0.443	0.038
	Police and defence	-0.425	0.053	0.763	0.039	-0.422	0.322	0.532	0.122
Higher	*Humanities*	-0.365	0.026	0.587	0.018	-0.221	0.019	0.552	0.012
education	*Social sciences*	0.003	0.035	0.256	0.029	-0.097	0.057	0.418	0.037
short cycle	*Technical*	-0.048	0.024	0.302	0.019	0.139	0.118	0.073	0.078
	Health care	0.153	0.077	0.285	0.065	-0.177	0.024	0.785	0.016
	Food, agriculture and transportat	-0.258	0.043	0.478	0.033	-0.095	0.114	0.277	0.073
	BA	0.043	0.067	0.170	0.054	-0.238	0.189	0.324	0.094
Higher	*Humanities*	-0.358	0.047	0.510	0.032	-0.044	0.055	0.297	0.032
education	*Natural sciences*	-0.390	0.069	0.616	0.051	-0.221	0.122	0.352	0.065
MA level	*Social sciences*	0.198	0.029	0.073	0.023	0.195	0.051	0.218	0.034
	Technical	0.021	0.033	0.264	0.027	0.112	0.101	0.110	0.072
	Food	0.223	0.053	0.034	0.042	0.395	0.094	-0.008	0.062
	Health care	0.587	0.031	-0.333	0.027	0.563	0.039	0.048	0.029
	Year dummies	Yes		Yes		Yes		Yes	
	Constant	-4.327	0.153	-0.661	0.098	-4.627	0.208	-2.555	0.104
	s_{μ}	0.760	0.009	0.838	0.005	0.626	0.011	0.730	0.005
Number of observations		1288888		1288888		1126960		1126960	
Number of individuals		136990		136990		122749		122749	
Log-likelihood		-116985.808		-287641.544		-59297.355		-243222.453	

TABLE A.3: EARNINGS EQUATIONS (CORRECTED FOR SAMPLE SELECTION BIAS AND UNOBSERVED HETEROGENEITY)

		Males				Females			
		Self-employment		Wage-employment		Self-employment		Wage-employment	
		Coefficient	Std.	Coefficient	Std.	Coefficient	Std.	Coefficient	Std.
Age (divided by 10)		0.707	0.030	0.323	0.004	0.282	0.081	0.069	0.005
Age squared (divided by 1000)		-0.834	0.034	-0.350	0.005	-0.410	0.092	-0.111	0.005
Married		0.110	0.006	0.063	0.001	-0.285	0.015	-0.101	0.001
Immigrant (western)		-0.278	0.031	-0.081	0.006	-0.264	0.062	-0.046	0.006
Immigrant (non-western)		-0.608	0.029	-0.187	0.006	-0.543	0.071	-0.065	0.008
Second generation immigrants		0.008	0.086	-0.027	0.014	-0.098	0.230	0.009	0.018
Regional	*Copenhagen*	-0.239	0.015	-0.085	0.002	-0.141	0.034	-0.044	0.002
dummies	*Large city*	0.019	0.016	-0.039	0.002	-0.033	0.036	-0.062	0.002
	Rural	0.041	0.012	0.014	0.002	-0.072	0.027	-0.042	0.002
Missing education		-0.068	0.031	0.086	0.006	-0.098	0.072	0.063	0.007
Secondary	*General*	0.079	0.029	0.017	0.005	0.093	0.062	0.009	0.005
school	*Commercial and technical*	0.258	0.047	0.238	0.009	-0.056	0.131	0.114	0.010
Vocationa	*Shop assistents*	0.196	0.016	0.185	0.003	0.107	0.029	0.093	0.002
training	*Building and construction*	0.025	0.015	0.066	0.003	-0.152	0.194	0.052	0.017
	Metal	0.071	0.015	0.087	0.003	0.175	0.288	0.101	0.027
	Graphic	0.148	0.041	0.223	0.007	0.276	0.172	0.171	0.017
	Technical	0.038	0.051	0.042	0.009	-0.113	0.061	0.036	0.005
	Service and transport	0.002	0.036	0.084	0.009	-0.013	0.038	-0.028	0.007
	Food	0.319	0.016	0.085	0.004	0.182	0.099	0.051	0.008
	Health care	0.144	0.112	0.043	0.012	0.049	0.051	0.039	0.003
Post	*Humanities and social sciences*	0.086	0.072	0.209	0.013	0.219	0.067	0.189	0.006
secondary	*Technical*	0.139	0.030	0.178	0.005	0.098	0.121	0.073	0.009
	Agriculture	0.316	0.064	0.218	0.013	0.071	0.171	0.135	0.013
	Health care	0.141	0.452	-0.021	0.027	-0.044	0.149	0.113	0.009
	Police and defence	0.242	0.076	0.208	0.007	0.215	0.755	0.086	0.022
Higher	*Humanities*	0.133	0.040	0.189	0.004	0.122	0.052	0.224	0.003
education	*Social sciences*	0.660	0.041	0.426	0.006	0.443	0.151	0.276	0.008
short cycl	*Technical*	0.300	0.029	0.401	0.005	0.324	0.207	0.385	0.020
	Health care	0.279	0.097	0.180	0.014	0.468	0.058	0.219	0.004
	Food, agriculture and transport	0.181	0.061	0.341	0.008	0.299	0.255	0.191	0.017
	BA	0.127	0.068	0.314	0.011	0.484	0.313	0.163	0.015
Higher	*Humanities*	0.233	0.080	0.336	0.007	0.247	0.109	0.445	0.007
education	*Natural sciences*	0.354	0.114	0.387	0.009	-0.066	0.378	0.476	0.015
MA level	*Social sciences*	0.773	0.034	0.486	0.006	0.719	0.109	0.476	0.007
	Technical	0.376	0.039	0.477	0.007	0.564	0.198	0.435	0.020
	Food	0.634	0.054	0.428	0.013	0.613	0.200	0.420	0.023
	Health care	0.959	0.028	0.663	0.008	1.257	0.069	0.627	0.010
$s_{a\mu}$		-0.062	0.003	-0.112	0.001	-0.024	0.005	-0.076	0.002
$s_{e?}$		-0.012	0.004	0.010	0.001	0.092	0.008	0.013	0.001
Constant		15.109	0.068	15.783	0.010	15.779	0.180	16.323	0.011

Summary Statistics

	Males: Self-employment	Males: Wage-employment	Females: Self-employment	Females: Wage-employment
Number of observations	168782	1037089	47496	994058
Number of individuals	27544	122754	10456	117511
s_a	0.6656	0.2929	0.9831	0.3092
s_e	0.4557	0.1737	0.5914	0.1869
Fraction of variance due to individual specific error	0.68	0.74	0.73	0.73
R-squared	0.12	0.20	0.11	0.23

Wage Dispersion and Decentralization of Wage Bargaining*

Christian M. Dahl†
University of Aarhus

Daniel le Maire‡
University of Copenhagen and CAM

Jakob Roland Munch§
University of Copenhagen and EPRU

July, 2008

Abstract

This paper studies how decentralization of wage bargaining from sector to firm level influences wage levels and wage dispersion. We use a detailed panel data set that covers a period of decentralization which facilitates identification of the effects of decentralization. Consistent with theoretical predictions we find that wages are more dispersed under firm-level bargaining compared to more centralized wage-setting systems. However, the differences across wage-setting systems are reduced substantially when controlling for unobserved individual level heterogeneity.

Keywords: Wage bargaining, decentralization, panel data, quantile regression

JEL Classification: J31, J51, C23.

---

*Financial support from the Danish Social Science Research Council is gratefully acknowledged. We thank Bo Honoré, Lars Haagen Pedersen, Rune Vejlin and participants at CAM Research Workshop 2007 and Labour Market Dynamic Growth Conference, Sandbjerg 2007 for helpful comments.

†Address: School of Economics and Management, University of Aarhus, Building 1322, DK- 8000 Aarhus C, Tel: +45 89421559, E-mail: cdahl@econ.au.dk.

‡Address: Department of Economics, University of Copenhagen, Studiestræde 6, DK - 1455 Copenhagen K, Tel.: +45 35323063, Fax: +45 35323064, E-mail: Christian.Daniel.le.Maire@econ.ku.dk.

§Address: Department of Economics, University of Copenhagen, Studiestræde 6, DK - 1455 Copenhagen K, Tel.: +45 35323019, Fax: +45 35323000, E-mail: Jakob.Roland.Munch@econ.ku.dk.

1 Introduction

Several advanced countries have undergone a process towards more decentralized wage bargaining in the labor market during the past decades. Comparing the 1970's to the 1990's not a single OECD country moved towards centralization, whereas a considerable number moved towards greater decentralization according to OECD (2004). This movement has in many countries been accompanied with a steady decline in union densities, while the extent of bargaining coverage has typically been unchanged. Decentralization of collective bargaining may have important implications for wage formation and wage dispersion in particular, but only scarce microeconometric evidence exist to document such effects.

The principal aim of this paper is to empirically examine the movement of decentralization in wage bargaining in terms of its impact on wage dispersion. From a theoretical standpoint decentralization may lead to increased wage dispersion because firm- and individual-specific characteristics are more likely to enter the wage contracts, while under centralized bargaining egalitarian union preferences are easier to accomplish.[1] Obviously, changes in wage dispersion may have important direct welfare implications through increased income inequality, but there may also be more indirect consequences. First, a movement away from a standard wage rate applying for all workers means that wages are more in accordance with individual productivity and local conditions, which tends to reduce misallocation, inefficiencies and unemployment in the labor market. In contrast to this view, Moene and Wallerstein (1997) argue that centralized bargaining tends to bolster expanding progressive industries and hamper declining ones, while local bargaining allows less productive plants to reduce wages and remain in operation. Also when risk-averse individuals face uncertainty about their position in the income distribution, unions may improve welfare by compressing the wage structure, see Agell and Lommerud (1992). In any case, it is clear that the link between bargaining level and wage dispersion is important for welfare, and a first step should be to empirically assess the extent to

[1] See e.g. Farber (1978) and Booth (1984) for theoretical models explicitly handling the role of wage dispersion in union preferences.

which decentralization increases wage dispersion.

Another aspect of decentralization is its impact on wage *levels*. A number of different explanations for higher mean wages under firm level bargaining may be put forth. First, higher wages at the local level may be due to rent sharing, see e.g. Blanchflower, Oswald and Sanfey (1996). Second, firms with local bargaining may encourage workers to work harder by offering higher wages through efficiency wage considerations, see e.g. Akerlof and Yellen (1988). Third, it may be argued that decentralization of collective bargaining makes it less likely that unions internalize externalities of many different types, see Calmfors (1993). For example decentralized wage increases may lead to higher product prices, thus increasing the cost of inputs for other firms. Such externalities may be taken into account in more centralized bargaining settings and may induce unions to restrain their wage demands. However, Calmfors and Driffill (1988) argue that the relationship between centralization and wage outcomes is hump shaped. At the national level unions internalize externalities and moderate their wage demands, but at firm level they also restrain wage demands because higher wages lead to higher product prices and lower demand for the goods produced by the firm, thereby reducing employment in the firm. At the industry level neither of these mechanisms are present to the same extent and so unions negotiate for higher wages at this level. For open economies Danthine and Hunt (1994) show that the hump shaped relationship between wages and centralization level flattens out as product market competition increases and so the room left open for diverging wage policies narrows. Thus the predictions concerning the impact of decentralization on wage levels are less clear-cut and is ultimately an empirical question.

We have access to a very rich longitudinal data set for private sector workers in the Danish labor market. The Danish labor market is interesting to study because four different wage setting systems, representing three different levels of centralization, coexist, and so their influence on wage formation may readily be compared. First, in one segment of the labor market wages are negotiated at industry level for all workers – this is the so-called standard-rate system. Clearly the scope for wages to reflect individual productivity is limited under this system. Second, a considerable part of the labor market has bargaining between unions and employers at the industry level over a contractual wage, which is accompanied by local bargaining at the firm level over an individual wage supplement (the minimum-pay and minimum-wage systems). In this case wages may better be in accordance with individual qualifications due to the local level bargaining. Third, a segment of the labor market has no centrally negotiated contractual wage, and wages are entirely determined at the firm level. Importantly, our data set covers a period where many labor market segments changed wage setting system towards bargaining at more decentralized levels. In particular, the importance of the segment with only firm-level bargaining has increased during our sample window.

The longitudinal dimension of the data is crucial for two main reasons. First, identification of the effects of decentralization on wage dispersion is greatly facilitated by the change of wage setting system over time for many workers. The wage setting system for the individual worker may change because the labor market segment changed its system due to the decentralization process or because the worker changed job. Second, in contrast to all the existing empirical evidence, longitudinal data allows us to control for unobserved heterogeneity. Our econometric approach is quantile regression, since this, in a very transparent way, illustrates the impact of wage setting systems in different quantiles of the wage distribution. However, it is only recently that quantile regression methods have been developed to better exploit the advantages of longitudinal data, see Koenker (2004) and Abrevaya and Dahl (2007). We apply the correlated random effects approach suggested by Abrevaya and Dahl (2007).

We find that decentralization of wage bargaining increases wage dispersion, i.e., wages are most dispersed under the most decentralized system – firm level bargaining. However, by using the panel data quantile regression approach we also find that the differences in wage dispersion between the wage setting systems are reduced substantially when unobserved individual heterogeneity is controlled for. With respect to the impact on mean wages we do not find important differences across bargaining systems.

The paper is organized as follows. Section 2 briefly reviews the existing empirical literature on unions and the dispersion of wages. Section 3 describes the institutional framework for wage bargaining in Denmark. This section also summarizes the aggregate development towards more decentralized wage bargaining in Denmark in the 1990's. Section 4 describes the data set, section 5 outlines the empirical framework, and the results are presented in section 6. Section 7 concludes.

2 Unions and the dispersion of wages

The impact of unions on wage formation and wage dispersion is a subject that has long attracted the attention of economists. There exists a large literature assessing the wage differential between union and non-union workers and the impact of unions on wage inequality (see e.g. Freeman (1980) for an early exposition and Card, Lemieux and Riddell (2004) for a recent review). This is an interesting issue in Anglo-Saxon countries where it makes sense to focus on union membership of the individual worker. However, in most continental European countries the relevant measure is the centralization level of bargaining, because even in countries with low union densities, bargaining agreements are typically extended to the majority of the workforce. In this section we briefly review the existing microeconometric evidence of the impact of the bargaining level on wage formation.

One of the first studies of the subject is Dell'Aringa and Lucifora (1994), who investigated the Italian metal-mechanical industry with establishment survey data from 1990. They found a positive wage differential in firms where unions are recognized for local bargaining as compared to firms where only the nationally bargaining wages apply. In addition, they find that firm-level bargaining raises wages more for white collar workers than for blue collar workers.

These results are consistent with a more recent paper by Card and de la Rica (2006) who study the effect of firm level contracting relative to regional or national contracts in Spain. They use the European Structure of Earnings Survey (ESES) from 1995, which is a matched worker-firm data set with information on whether the worker belongs to a multi-employer bargaining regime or a regime with single-employer bargaining (firm-level bargaining). They show that there is a positive wage premium of 5-10 per cent associated with single-employer bargaining. Interestingly, they also find that the premium is higher for more highly-paid workers using a weighted least squares approach. They take this as weak evidence for a more flexible wage structure under firm-level bargaining.

Two other recent contributions use the ESES data set for 1995 to examine the effect on the wage dispersion. Dell'Aringa and Pagani (2007) perform a variance decomposition of the ESES data for Italy, Belgium and Spain. In Italy and Belgium there is no clear effect of single employer bargaining on wage dispersion, while for Spain, consistently with Card and de la Rica (2006), they find a small positive effect. In addition to the variance decomposition, Dell'Aringa and Pagani separately for each wage setting system estimate a quantile regression model in order to compute wage inequality measures conditional on the different explanatory variables. Thus, when taking observable heterogeneity into account, they find that, if anything, single employer bargaining tends to decrease wage dispersion in Italy and Belgium, while the opposite is true for Spain.

Plasman et al. (2007) also perform a variance decomposition exercise and find for Belgium, Denmark and Spain that decentralized bargaining increases the mean wage. Furthermore, single-employer bargaining increases the dispersion of wages in Denmark and Belgium while it decreases the wage dispersion in Spain which is in contrast to the findings of Card and de la Rica (2006) and Dell'Aringa and Pagani (2007).

Using a cross section data set for 1991 Hartog, Leuven and Teulings (2002) investigate the impact of bargaining regime on wages in the Netherlands, and they find that mean wages under firm-specific and industry-level contracting are very similar. They also observe workers in firms with no collective bargaining and in firms with mandatory extensions of an industry agreement, and wage differentials between regimes were found to be no larger than 4 per cent. Also in terms of wage dispersion modest differences are found between the four regimes, but firm specific bargaining yields the greatest residual variation of wages.

Comparing contractual wages and actual wages Cardoso and Portugal (2005) find for Portugal a substantial wage cushion with industry averages of 20-50 per cent of the contractual wages. From tobit regressions it is found that the effects of worker and firm characteristics on contractual wage and the wage drift have the same sign, so that wage drift stretches the wage distribution. A measure for the degree of union bargaining power is constructed as the concentration of bargaining and Cardoso and Portugal find that the higher concentration the higher contractual wage rate and – by interacting this bargaining power measure with worker attributes – the lower returns to these attributes. Interestingly, the higher contractual wage rate is off-set by a smaller wage drift.

To sum up, most results indicate that wages are higher when they are negotiated at the firm level as compared to the industry level. However this result is refuted by the evidence from the Dutch labor market. With regards to the effects on wage dispersion the evidence is more mixed although most results suggest that local bargaining leads to higher wage dispersion than industry level bargaining.

A distinguishing feature is that all the mentioned studies use cross section data, and a caveat applying here is that there may be unobserved skill differences between workers covered by centrally and locally negotiated wage contracts. For example it may be argued that if firms with local bargaining reward observed skills such as education more generous, they will likely also reward unobserved skills better. Besides this, if local bargaining is known to imply more dispersed wages, the Roy (1951) model would suggest that high ability workers sort into decentralized bargaining segments. Hence, we expect a positive correlation between local bargaining and unobserved ability and that appropriately controlling for unobserved heterogeneity should imply smaller estimated effects of different bargaining arrangements. With access to longitudinal data covering a period of decentralization we are in position to take account of unobserved heterogeneity and we may more reliably identify the effects of decentralization since it seems reasonable to take the decentralization process to be exogenous to the individual worker.

3 The Danish wage setting system

Whereas job protection is low in Denmark and, thus, provides much flexibility, the wage setting has been rather inflexible – Denmark has been one of the OECD countries with the most compressed wage structures – which in part is due to a combination of three factors. First, the benefit system is generous with a high benefit level for low income groups and a long benefit period of up to four years. Second, the Danish labor market is highly organized on both employer and worker sides: In 2000 74 per cent of the workers were members of a union and more than 80 per cent were covered by a collective agreement cf. OECD (2004). Third, wage bargaining has historically been centralized, but, as explained below, this has changed during the 1980's and 1990's. According to Boeri et al. (2001) the centralization/coordination index of the bargaining system (which lies between 0 and 1) has for Denmark dropped from 0.64 for the period 1973-1977, to 0.47 for 1983-1987 and 0.34 for 1993-1997.

Wage bargaining at the industry and firm levels depends on the wage setting system used in the industry collective agreement. In Denmark there are four different systems: First, under the *standard-rate system* ("normallønssystemet") actual wages of workers are set by the industry collective agreement and the wages are not modified at the firm level. Second, under the *minimum-wage system* ("minimallønssystemet") the wage rates set at the industry level represent a floor and are intended to be used only for very inexperienced workers. Hence, for other workers this wage rate is supplemented by a personal pay supplement. In practice, the personal pay supplements are often negotiated collectively with the cooperation of the workplace union members' representative. Third, a somewhat similar *minimum-pay system* ("mindstebetalingssystemet") exists. Rather than operating with a personal pay supplement on top of the industry-level negotiated wage rate, the minimum pay system uses a personal wage. The wage rate negotiated at the industry level can be thought of as a safety net in the form of a minimum hourly rate that must be paid under all circumstances. Finally, under *firm-level bargaining* the collective agreements state that wages are negotiated at the plant or firm level without any centrally bargained wage rates ("uden lønsats").

Table 1 shows the development in the use of these four wage setting systems in the private sector labor market covered by the two bargaining parties at national level; The Danish Confederation of Trade Unions (LO) and The Confederation of Danish Employers (DA). There has been a trend towards more decentralized and flexible wage setting, where the per centage with a standard wage rate was reduced by 50 per cent. Since 1993 the most decentralized segment (i.e. the firm-level bargaining segment) has grown from a coverage of 4 per cent to 22 per cent in 2004. For the two remaining decentralized wage systems, that is the minimum-wage and the minimum-pay systems, we also see considerable variation over time. As an example of an important bargaining segment making the transition to firm-level bargaining the area covering office clerks can be mentioned. In the empirical analysis below we use data for 1994-1999, so we capture the increased importance of firm-level bargaining in particular.

Insert Table 1 about here

4 Data and descriptive statistics

We have access to information about individual characteristics for the full population of workers aged 18-65 years in the Danish labor market for the years 1994-1999.[2] These characteristics are extracted from the Integrated Database of labor Market Research (IDA) and the Income Registers in Statistics Denmark – see Abowd and Kramarz (1999) for a brief description. The hourly wage rate is obviously an important individual level variable in the analysis, and this wage rate is calculated as the sum of total labor income and mandatory pension fund payments divided by the total number of hours worked in any given year. The measure for total labor income as such is highly reliable since it comes from the tax authorities, and the pension fund payments are also available in the registers. These payments were introduced in the early 1990s, and have been rising throughout the

[2] To reduce estimation time we extract a 40 per cent random sample as a starting point. This section reports summary statistics for this subsample.

sample period, but not in a uniform manner across collective bargaining segments of the labor market and they are therefore important to account for.

We use very detailed industry and occupation variables to determine the collective agreement to which the individual belongs. The industry code follows the NACE industry classification, and the occupation variable is based on the so-called DISCO code, which is the Danish version of the ISCO-88 classification. We use the most disaggregated definition of the industry- and occupation codes, i.e., the six digit NACE code and the four digit DISCO code. By using these industry and occupation variables to define bargaining segments of the labor market we follow the two bargaining parties at national level, LO and DA, since they use the codes to assess the economic implications of proposals for the workers and employers they represent. That is, we determine the bargaining segments in the same way as DA and LO, when the parties evaluate the bargaining outcome. However, the construction of such bargaining segments is not completely flawless. For example, a firm may wish to stay outside its industry's collective agreement and we will not be able to see this in the data. Nevertheless we are confident that our allocation of workers into bargaining segments is fairly accurate since we end up with a distribution of workers across wage setting systems that resembles Table 1 quite closely (more on this below). We have identified 31 bargaining segments within the DA/LO segment which correspond to roughly 50 per cent of workers in the organized part of the private labor market in Denmark. Coupled with information about the bargaining system each segment operates under in each year, it was straightforward to partition all workers into the four bargaining systems under consideration.

A long list of individual socio economic characteristics are used as control variables in the analysis. We use dummies for gender, the presence of children, marriage, immigrant status, city size ('Large city', and 'Rural' with 'Copenhagen' as the omitted category), education ('Unskilled', 'Short term higher education', 'Long term higher education' with 'Vocational education' as the omitted category)[3] and experience (measured as actual labor

[3]The classification of education groups rely on a Danish education code that corresponds to the International Standard Classification of Education (ISCED). 'Higher education' basically corresponds to the two highest categories (5 and 6) in the International Standard Classification of Education (ISCED),

market experience since 1964). There are also dummies for the size of the workplace measured in terms of the workforce. Furthermore, different industries may face different degrees of competition from abroad, which may well be reflected in both the wage level and the wage dispersion within a given industry. To avoid that wage setting dummies pick up differences in business conditions between industries we include industry dummies.

In Table 2 we show some summary statistics for each of the four wage setting systems in 1997. With respect to the average wage level the unconditional evidence is mixed since the most decentralized segment, firm-level bargaining, has the highest wage level while the standard-rate system, which is the least decentralized wage setting system, has the third highest wage level. To assess the extent of wage dispersion we have also computed the unconditional 90th/10th, 90th/50th and the 50th/10th per centile ratios for each of the wage setting systems. The wage dispersion is much higher for the workers belonging to the minimum-wage system which is particularly true for the lower end of the wage distribution. Wage dispersion under firm-level bargaining appear to be only slightly higher than the remaining two wage setting systems – the standard-rate system and the minimum-pay system. It should be noted that, since the standard-rate system typically applies for unskilled workers while many skilled workers belong to the minimum-wage system, we should not put too much emphasis on the unconditional evidence in Table 2.

Insert Table 2 about here

With our longitudinal data set identification of the impact of wage setting system on wages rests on the existence of workers who change wage setting system. This can happen for two reasons; the bargaining segment may change its system as a part of the decentralization process or the worker may change job. Table 3 tracks the persons in our sample that change wage setting system in each year. The second column shows the total number of workers changing wage setting system, and it is seen that there is a transition rate of

i.e., the individual has a tertiary education. 'Vocational education' is defined as the final stage of secondary education encompassing programmes that prepare students for direct entry into the labor market. Thus persons with just high school or equivalent or less than that are classified as 'Unskilled'.

around 5-10 per cent each year. Column 3-6 decompose the total annual changes further. First, the entire bargaining segment can change wage setting system due to the decentralization process (column 3), which contributes with the majority of transitions. Second, a worker can change occupation and/or industry and, thereby, perhaps also bargaining segment and wage setting system (column 4-6).[4]

Insert Table 3 about here

Since the wage setting system variable is constructed based on the industry and occupation codes we know that measurement error may arise – in particular the occupation code is known to be unstable within job spells in some years, and this may bias our estimates. In relation to panel data estimations of a union membership effect on wages Freeman (1984) argues that measurement error in the union membership variable will lead to a downward-biased estimate of the effect. However, when entire bargaining segments change wage setting system as in our data, measurement error is less of a problem compared to the situation where we only rely on people changing jobs and, thereby, wage setting system. The data still include job changers though (see columns 4-6 in Table 3), so in the empirical analysis below we restrict the sample further to reduce potential problems with measurement error. Specifically we throw away all workers that change wage setting system because of a shift in the occupation code (column 5) unless they also change employer. This reduces the number of wage setting system changes due to occupation changes by approximately 90 per cent.

The sample version of Table 1 is Table 4. Even though we only distinguish between 31 bargaining segments and, thus, leave out part of the DA/LO segment, the development in Table 4 resembles that of Table 1 quite closely. As described above, much of the

[4]It should be noticed that it is only the year when a collective agreement is initiated that the wage setting system changes. For most bargaining segments this happened every second year in the early 1990s, i.e., the years 1991 and 1993. However, some collective agreements in 1995 and 1997 had a duration of three years. The 710 persons whose bargaining segment seems to have changed wage setting system because of decentralization in 1996 are persons that are in the given segment in 1994 and 1996, but out of the sample in 1995. This is also the case for the 1,858 persons in 1998 and the 976 persons in 1999.

decentralization of the bargaining level in Denmark took place before 1993, but this is not essential to the analysis as long as we still have considerable time variation in the data.

Insert Table 4 about here

5 Empirical framework

To assess the impact of decentralization on wage dispersion we use quantile regression. Quantile regression techniques for panel data have only recently been developed, and this section outlines the approach we follow.

The standard (cross section) quantile regression model of Koenker and Bassett (1978) is given by

$$y_i = x_i\beta + u_i \quad \text{with} \quad Q_\tau\left(y_i|x_i\right) = x_i\beta_\tau, \tag{1}$$

where $i = 1, ..., N$ is indexing the individuals, y_i is the log of the individual hourly wage rate, β_τ is a $k \times 1$ vector and x_i is a $1 \times k$ vector of explanatory variables. $Q_\tau\left(y_i|x_i\right)$ denotes the τth conditional quantile of y given x, $\tau \in (0, 1)$.

In the linear model the solution to endogeneity problems in presence of panel data is typically the fixed-effects estimation. Unfortunately, the usual differencing strategy does not apply here since the conditional quantiles are not linear operators, that is

$$Q_\tau(y_{it} - y_{is}|x_i) \neq Q_\tau(y_{it}|x_i) - Q_\tau(y_{is}|x_i), \tag{2}$$

where time periods $t \neq s$ and where $x_i \equiv (x_{i1}, ..., x_{iT})$.

Abrevaya and Dahl (2007) suggest to estimate a Chamberlainian correlated random effects quantile regression model to take account of unobserved heterogeneity. Their estimator is most easily understood if we begin by considering the standard linear panel data model

$$y_{it} = x_{it}\beta + c_i + u_{it}, \tag{3}$$

where $t = 1, 2, ..., T$, c_i is the individual specific term and u_{it} the error term.

As in Chamberlain (1982, 1984) assume that the unobservable term c_i is a linear projection onto the observables plus a disturbance v_i, that is

$$c_i = \psi + x_{i1}\lambda_1 + ... + x_{iT}\lambda_T + v_i. \tag{4}$$

Plugging this into equation (3) gives

$$y_{it} = x_{it}\beta + \psi + x_{i1}\lambda_1 + ... + x_{iT}\lambda_T + v_i + u_{it}. \tag{5}$$

We need to make two assumptions in order to estimate the model in equation (5):

$$(A1) \quad v_i \text{ independent of } x_i$$

and

$$(A2) \quad Q_\tau\left(u_{it}|x_i, v_i\right) = Q_\tau\left(u_{it}|x_{it}\right).$$

Assumption $(A1)$ is also needed in the traditional random-effects probit model (see for example Wooldridge (2002)), but is stronger than the conditional mean independence needed in the linear Chamberlainian random-effects model. By assumption $(A2)$ we assume strict exogeneity, which effectively rules out feedback-effects from current wages, y_{it}, on future values of x_{it}.

The partial derivative of the conditional quantile with respect to x_{it} is

$$\frac{\partial Q_\tau\left(y_{it}|x_i\right)}{\partial x_{it}} = \beta + \lambda_t + \frac{\partial Q_\tau\left(u_{it}|x_{it}\right)}{\partial x_{it}}, \tag{6}$$

and

$$\frac{\partial Q_\tau\left(y_{is}|x_i\right)}{\partial x_{it}} = \lambda_t \tag{7}$$

for $t \neq s$. Following Abrevaya and Dahl we measure the effect of x_{it} as

$$\frac{\partial Q_\tau\left(y_{it}|x_i\right)}{\partial x_{it}} - \frac{\partial Q_\tau\left(y_{is}|x_i\right)}{\partial x_{it}} = \beta + \frac{\partial Q_\tau\left(u_{it}|x_{it}\right)}{\partial x_{it}}, \tag{8}$$

which essentially reflects the desirable feature of quantile regression that x_{it} is allowed to have different effects on y_{it} in different quantiles. In other words, this is a general result of quantile regression (cf. Koenker and Bassett (1982)) and does not only pertain to the Abrevaya and Dahl estimator. For illustrative purposes Abrevaya and Dahl assumes that $u_{it}|x_i, c_i \sim N\left(0, (\gamma_t x_{it})^2\right)$. This panel data version of the linear-scale model implies that equation (8) becomes

$$\frac{\partial Q_\tau\left(y_{it}|x_i\right)}{\partial x_{it}} - \frac{\partial Q_\tau\left(y_{is}|x_i\right)}{\partial x_{it}} = \beta + \gamma_t Q_\tau\left(\epsilon_{it}\right), \tag{9}$$

where ϵ_{it} is a standard normal random variable. It is apparent that for $\gamma_t > 0$ the effect of x_{it} is larger in the upper part of the y_{it} distribution.

A drawback is that the approach only works for balanced panels. Therefore, we estimate the model on different (sub-) samples. First, we construct a sample with individuals we observe twice or more and randomly select two observations for each individual. Second, we extract a sample with persons observed at least four times and randomly select four observations per individual. Third, we use the Mundlak (1978) version of the correlated random-effects model where the unobserved part includes averages rather than the values of each period. In this case the unobservable term becomes

$$c_i = \psi + \bar{x}_i \lambda + v_i. \tag{10}$$

This allows us to use all observations, but at the expense of restricting the linear projection in equation (4).

6 Results

This section first presents results for the impact of wage setting systems on mean wages using a standard Mincer wage equation approach. This is followed by results for the impact on wage dispersion using the panel data quantile regression approach outlined in the previous section. Finally we present some robustness checks of our preferred specification.

6.1 Wage levels and wage-setting systems

While our focus is on the the impact of decentralization on wage dispersion it is instructive to first study how mean wages differ across wage setting systems controlling for individual heterogeneity. Table 5 reports estimation results from a pooled OLS as well as linear random-effects and fixed-effects models. It is first seen that we obtain the usual signs of the human capital and socio-demographic variables.

With respect to the wage setting systems the most clean comparison is between the standard-rate system (where wages are negotiated at sector or industry level) and firm level bargaining since these systems represent the most centralized and the most decentralized systems respectively. As described in section 3 the minimum payment and minimum wage systems are intermediate cases since they both have elements of a centrally negotiated wage and locally negotiated wages. In the following we use as the base category the standard-rate system. For the OLS regression we find that wages are 5.2 per cent higher under firm level bargaining than under the standard-rate system. However, this quite substantial wage differential vanishes if unobserved individual heterogeneity is controlled for through random effects, and if the fixed effects estimator is used the effect even changes sign such that wages are 1.3 per cent lower under firm level bargaining. This clearly suggests that it is important to control for unobserved heterogenety and that failure to do so leads to an upward bias in the coefficient, i.e., unobserved ability may be better rewarded under local bargaining. With respect to the two intermediate systems there is also a negative effect of minimum pay once unobserved heterogeneity is accounted for while there is no wage differential between the standard rate system and the minimum wage system.

Insert Table 5 about here

To sum up, we find evidence of lower mean wages under the more decentralized bargaining systems, and this seems to be at odds with what is expected from simple rent-sharing or efficiency-wage considerations while it is more consistent with the externality explanation of Calmfors and Driffil (1988) as argued in the introduction. Also, it is important to be able to control for unobserved heterogeneity as otherwise the wage differentials between wage-setting systems are greatly exaggerated. One important aspect which cannot be studied using the simple mean regressions is the fact that the decentralization process may have very uneven effects across the wage distribution – an issue to which we now turn.

6.2 Wage dispersion and wage-setting systems

As a first step we will start out with a simple quantile regression without exploiting the longitudinal nature of our data. Table 6 displays the results from pooled quantile regression models for the quantiles 0.10, 0.25, 0.50, 0.75, and 0.90. In general the coefficients on the individual level variables are fairly constant across the different quantiles, but there are also some notable exceptions. For example women and immigrants have a higher wage penalty in the top end of the wage distribution, which suggests the existence of a glass ceiling for these groups in the labor market (this is consistent with the results of Albrecht, Björklund and Vroman (2003) and Pendakur and Woodcock (2008)). Also, unskilled workers have relatively lower wages than workers with vocational education in the bottom of the wage distribution. Of particular interest is the effect of the wage system dummies, and it is found that the coefficient on the variables for the three decentralized systems increase between almost all of the reported quantiles, so that the effects at the 90th quantile are substantially higher compared to the effects at the 10th quantile. For example the effect of working under firm-level bargaining compared to the standard-rate system more than triples from the 10th to the 90th quantile (from 3.2 per cent to 10.7

per cent). Thus these results support the prediction that decentralization leads to increased wage dispersion for example because firm- and individual-specific characteristics are more likely to enter wage contracts, or because egalitarian union preferences become more difficult to accomplish. However, we suspect that the coefficients on the wage system dummies are biased upwards because unobserved heterogeneity is not controlled for.

Insert Table 6 about here

Therefore the next step is to apply the panel data quantile regression techniques outlined in section 5. Table 7 shows results for estimation of the Abrevaya and Dahl (2007) correlated random-effects quantile regression model for the case where we balance the panel by randomly selecting only two observations for each individual for the reasons explained above. It is first noted that the effects of individual level variables are only changed slightly. Some variables like age, experience and workplace size appear to have somewhat stronger effects now, but otherwise the results are robust. However, for the wage system dummies the picture changes in important ways. For firm level bargaining we again find that wage dispersion is higher than under the standard-rate system, but the coefficients are in accordance with the fact that there is roughly no mean effects, cf. Table 5. That is, we find negative coefficients in the lower quantiles and positive coefficients in the higher quantiles such that workers under firm-level bargaining earn 2.6 per cent lower wages at the 10th quantile and 3.9 per cent higher wages at the 90th quantile compared to workers under the standard-rate system. There appears to be no significant differences between the minimum-pay system and the standard-rate system, while the minimum-wage system increases wages in the upper part of the wage distribution.

Insert Table 7 about here

To study how these results depend on the sampling scheme we also estimate the correlated random-effects quantile regression model where we randomly select four instead of two observations for each worker. This effectively corresponds to selecting stable workers in the sense that they enter the original sample at least four out of the six years in our sample window. With respect to the firm-level bargaining system the results are qualitatively similar but the effects are slightly stronger such that workers now earn 4.6 per cent lower wages at the 10th quantile and 4.3 per cent higher wages at the 90th quantile compared to workers under the standard-rate system, see Table 8. The minimum-pay system now has a negative effect in the 10th quantile but there are no changes otherwise, and the effects of the minimum-wage system are also not changed in any important way.

Insert Table 8 about here

To cast further light on the importance of the sampling scheme we also estimate a version of the correlated random effects quantile regression model where we approximate the unobservable part with the individual means of the explanatory variables as in Mundlak (1978), see equation (10). This has the advantage that we can use all observations in our original sample and thus circumvent the requirement of a balanced sample, but it comes at the expense of a more restrictive functional form for the unobservables. The results are displayed in Table 9, and it is seen that they are very much in accordance with the two previous sets of results. In fact, the effects in Table 9 typically lie in between the effects found for the balanced sample with two observations per worker and the balanced sample with four observation per worker. For example, workers under firm-level bargaining now earn 3.0 per cent lower wages at the 10th quantile and 3.9 per cent higher wages at the 90th quantile compared to workers under the standard-rate system.

Insert Table 9 about here

To sum up, the three different versions of the correlated-random effects quantile regression model yield fairly robust results showing that decentralization of wage bargaining increases wage dispersion. Under the most clear cut comparison, i.e. the effect of working under firm-level bargaining (where wages are set entirely at the firm level) compared to the standard-rate system (where wages are set entirely at the sector level) negative effects are found in the lower part of the wage distribution and positive effects are found in the upper part. The two intermediate bargaining systems – the minimum-pay and the minimum-wage systems – wages are only slightly more dispersed than under the standard-rate system.

In the following we will take the Mundlak version of the empirical model as our main specification as it yields very similar results to the more flexible models while still being based on the full sample. For illustrative purposes we have used this model to compute the coefficients for every two per centiles and plot them with 5 per cent confidence bands – see Figure 1. This shows very clearly that wage dispersion is higher under firm level bargaining.

Insert Figure 1 about here

6.3 Robustness

A major advantage of our analysis vis-à-vis the existing literature is that we exploit time variation in the wage system of the individual worker, but this also raises the question about whether wage system changes are exogenous. We argue that if the wage system change because of the decentralization process, i.e. a whole bargaining segment changes wage system, then this can safely be taken to be exogenous to the worker. The wage system may also change because workers change jobs from one bargaining segment of the labor market to another, and in this case endogeneity may be an issue as e.g. high paid workers in the standard-rate system may be inclined to change to jobs under firm-level bargaining to receive a higher wage. In traditional Mincer human capital wage equations this issue may be approached by also estimating a selection equation for the choice of wage system (see e.g. Vella and Verbeek (1998) for an application to union wage premia). However, corresponding techniques are not yet developed for the panel data quantile regression case, and in any case this approach also requires proper instruments which is not immediately available in our data. Therefore, to proceed we have to settle for more indirect evidence for exogeneity of the wage system variables.

Table 3 showed that most wage system changes are due to the decentralization process, so a straightforward sensitivity test would be to simply leave out all wage system changes that can be ascribed to job changes. However, our reference wage system is the standard-rate system because it represents the most clear-cut example of a wage system with wage determined solely at the sector level, but no bargaining segments changed to or from the standard-rate system during our sample window (see Tables 1 and 4), so we have to rely on job movers. In the following we study whether these job changes are plagued by endogeneity.

The first step is to provide some further descriptive statistics for the wage system changes. There are 53,012 wage system changes in the data, and two thirds of these are due to the decentralization process and the rest is job mobility. Among job changes involving the standard rate system the most frequent type is between the standard rate system and the minimum pay system – more than one third of all job changes are in this category. In fact only very few workers change job from the standard-rate system to a job under firm-level bargaining or vice versa, but once we can identify the wage effects of changing between the standard-rate system and minimum-pay we have also identified the effects of firm-level bargaining because of sufficiently many exogenous 'decentralization' transitions between minimum-pay and firm-level bargaining.

Since most of the mobility in and out of the standard-rate system is to/from the minimum-pay system and since identification therefore relies on this transition in particular we will now study potential endogeneity of this transition only. One way to do this is to include two additional dummy variables for a changes between the two wage systems, i.e. the variable 'Change standard-rate to minimum-pay' in Table 10 takes the value 1 if the worker has experienced this transition as the latest transition and 0 otherwise. If mobility is endogenous we would expect that these variables enter the model with significant effects. Wages should rise in the top end of the wage distribution if workers change to the more decentral minimum-pay system and they should fall in the bottom end. Likewise wages should fall in the top end of the wage distribution if workers change to the centralized standard-rate system and they should rise in the bottom end. However, we find no evidence for such effects – all coefficients on the change variables are insignificant. At the same time the direct effects of the wage system dummies are not dramatically changed.

Insert Table 10 about here

The next question is whether we would get similar results to the main results of Table 9 if in addition to the exogenous 'decentralization' transitions we rely only on job movers between the standard-rate system and the minimum-wage system, i.e. if we remove all observations involving other types of job mobility. A Comparison between the results of Table 9 and 11 shows only small changes in the coefficients on the wage system variables, so mobility between standard-rate and minimum-wage systems is sufficient to get the main results and these job changes appear not to be driven by wage concerns. We take this as evidence for our main results not being seriously plagued by endogenity bias through job mobility.

Insert Table 11 about here

7 Conclusion

Many European labor markets have undergone a process towards more decentralized wage bargaining during recent decades. Such changes may have important welfare implications both in terms of efficiency and equity. When wages are negotiated locally at the firm level as opposed to more centralized bargaining, wages are more likely to reflect individual productivity and firm specific conditions. This should lead to increased wage dispersion.

We use a unique register-based panel data set covering a period of decentralization in the Danish labor market, and to the best of our knowledge we are the first to study these questions using longitudinal data. This is a crucial element because the time variation allows us to identify the effects of decentralization as many workers have seen their wage setting system change as a result of the decentralization process. In contrast, the existing literature has relied on cross section data. Also, in contrast to previous studies, the longitudinal dimension allows us to control for unobserved individual heterogeneity. This is important because by doing so the wage structure differences across wage-setting systems are substantially narrowed down.

We find empirical evidence in support of the predictions from theory, i.e., wage dispersion is higher under the more decentralized wage setting systems. In our main specification workers under firm-level bargaining, where wages are set entirely at the firm level, earn 3.0 per cent lower wages at the 10th quantile and 3.9 per cent higher wages at the 90th quantile compared to workers under the standard-rate system, where wages are entirely set at the sector level.

References

[1] Abowd, J. and F. Kramarz (1999): The Analysis of Labor Markets using Matched Employer-Employee Data, in: Orley Ashenfelter and David Card (eds.), *Handbook of Labor Economics*, Vol. 3B, Elsevier Science, Amsterdam.

[2] Abrevaya, J. and C.M. Dahl. (2007): The Effects of Smoking and Prenatal Care on Birth Outcomes: Evidence from Quantile Estimation on Panel Data, forthcoming in *Journal of Business and Economic Statistics.*

[3] Agell, J. and K. E. Lommerud (1992): Union Egalitarianism as Income Insurance, *Economica* **59**, pp. 295-310.

[4] Akerlof, G. and J. Yellen (1988): Fairness and Unemployment, *American Economic Review*, **78**, pp 44-49.

[5] Albrecht, J., A. Björklund and S. Vroman (2003): Is there a Glass Ceiling in Sweden?, *Journal of Labor Economics* **21**, pp. 145177.

[6] Andersen, T.M (2003): Changes in Danish Labor Market Bargaining. The Prototypical Case of Organized Decentralization? in E.J. Dølvik and F. Engelstad (eds.): *National Regimes of Collective Bargaining in Transformation: Nordic Trends in a Comparative Perspective*, Makt og demokratiutredningens rapportserie, Rapport 54.

[7] Blanchflower, D. G., A. J. Oswald and P. Sanfey (1996): Wages, Profits, and Rent Sharing, *Quarterly Journal of Economics* **111**, pp. 227-251.

[8] Boeri, T., A. Brugiavini and L. Calmfors (2001): *The Role of Unions in the Twenty-First Century*, Oxford University Press, Oxford.

[9] Booth, A. (1994): A Public Choice Model of Trade Union Behaviour and Membership, *Economic Journal* **94**, pp. 883-898.

[10] Calmfors, L. (1993): Centralization of Wage Bargaining and Macroeconomic Performance, *OECD Economic Studies* **21**, pp. 161-191.

[11] Calmfors, L. and J. Driffill (1988): Bargaining Structure, Corporatism and Macroeconomic Performance, *Economic Policy* **6**, pp. 13-61.

[12] Card, D. and S. de la Rica (2006): Firm-Level Contracting and the Structure of Wages in Spain, forthcoming in *Industrial and Labor Relations Review.*

[13] Card, D., T. Lemieux and W. C. Riddell (2004): Unions and Wage Inequality, *Journal of Labor Research* **25**, pp. 519-562.

[14] Cardoso, A. and P. Portugal (2005): Bargained Wages and the Wage Cushion under Different Bargaining Settings, *Journal of Labor Economics* **23**, pp. 875-902.

[15] Chamberlain, G. (1982): Multivariate Regression Models for Panel Data, *Journal of Econometrics* **18** , pp. 5-46.

[16] Chamberlain, G. (1984): Panel Data, in *Handbook of Econometrics, Vol. 2*, eds. Z Griliches and M.D. Intriligator, Amsterdam, North-Holland.

[17] Danthine, J. P. and J. Hunt (1994): Wage Bargaining Structure, Employment and Economic Integration, *Economic Journal* **104,** pp. 528-541.

[18] Dell'Aringa, C. and C. Lucifora (1994): Collective Bargaining and Relative Earnings in Italy, *European Journal of Political Economy* **10**, pp. 727-747.

[19] Dell'Aringa, C. and L. Pagani (2007): Collective Bargaining and Wage Dispersion in Europe, *British Journal of Industrial Relations* **45**, pp. 29-45.

[20] Farber, H. (1978): Individual Preferences and Union Wage Determination: The Case of the United Mine Workers, *Journal of Political Economy* **86**, pp. 923-942.

[21] Freeman, R.B. (1980): Unionism and the Dispersion of Wages, *Industrial and Labor Relations Review* **34**, pp. 3-23.

[22] Freeman, R.B. (1984): Longitudinal Analyses of the Effects of Trade Unions, *Journal of Labor Economics* **2**, pp. 1-26.

[23] Hartog, J., E. Leuven and C. Teulings (2002): Wages and the Bargaining Regime in a Corporatist Setting: the Netherlands, *European Journal of Political Economy* **18**, pp. 317-331.

[24] Koenker, R. (2004): Quantile Regression for Longitudinal Data, *Journal of Multivariate Analysis* **91**, pp. 74-89.

[25] Koenker, R. and G.W. Bassett (1978): Regression Quantiles, *Econometrica* **46**, pp. 33-50.

[26] Koenker, R. and G.W. Bassett (1982): Rubust Tests for Heteroscedasticity Based on Regression Quantiles, *Econometrica* **50**, pp. 43-62.

[27] Moene, K. O. and M. Wallerstein (1997): Pay Inequality, *Journal of Labor Economics* **15**, pp. 403-430.

[28] Mundlak (1978), On the Pooling of Time Series and Cross Section Data, *Econometrica* **46**, pp. 69–85.

[29] OECD (2004): *Employment Outlook*, Paris.

[30] Pendakur, K. and S. Woodcock (2008), Glass Ceilings or Glass Doors? Wage Disparity Within and Between Firms, forthcoming in *Journal of Business and Economic Statistics*.

[31] Plasman, R., M. Rusinek, F. Rycx (2007): How Do Company Collective Agreements Affect Wages? Evidence from Four Corporatist Countries, *European Journal of Industrial Relations* **13-2**, pp. 161-180.

[32] Roy, A.D. (1951): Some Thoughts on the Distribution of Earnings, *Oxford Economic Papers* **3**, pp. 135-146.

[33] Vella, F. and M. Verbeek (1998): Whose Wages do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men, *Journal of Applied Econometrics* **13**, pp. 163-183.

Table 1: Private sector wage setting systems 1989-2004

	1989	1991	1993	1995	1997	2000	2004
Standard-rate	34	19	16	16	16	15	16
Minimum-wage	32	37	13	12	21	23	27
Minimum-pay	30	40	67	61	46	42	35
Firm level	4	4	4	11	17	20	22
Total	100	100	100	100	100	100	100

Source: Danish Employers' Federation (DA).

Table 2: Wage dispersion in 1997 by type of wage setting system

	No. of obs	Mean	90th/10th	50th/10th	90th/50th
Standard-rate	21,354	150.75	1.89	1.37	1.39
Minimum wage	37,665	139.04	2.87	1.89	1.52
Minimum payment	85,976	153.39	1.89	1.33	1.42
Firm level bargaining	30,840	155.89	2.06	1.39	1.48

Table 3: Transitions between wage setting systems 1995-1999

	No. of obs.	All changes	Decentralization	Change in occ. and industry	Change in occupation	Change in industry
1994	165,190	.	.	.	.	.
1995	166,365	16,988	14,333	1,191	934	530
1996	169,023	4,975	710	2,123	1,367	775
1997	175,835	22,539	17,095	3,206	1,307	931
1998	180,297	7,816	1,858	3,741	1,285	932
1999	183,885	7,985	976	4,248	1,842	919
Total no. of obs.	1,040,595	60,303	34,972	14,509	6,735	4,087

Table 4: Private sector wage setting systems 1994-1999, data

	1994	1995	1996	1997	1998	1999
Standard-rate	12.6	11.6	11.9	12.1	12.1	12.3
Minimum wage	13.5	13.8	13.7	21.4	21.2	21.6
Minimum pay	73.3	62.2	62.2	48.9	49.2	48.6
Firm level bargaining	0.6	12.4	12.2	17.5	17.5	17.4
Total no. of obs.	165,190	166,365	169,023	175,835	180,297	183,885

Figure 1: Abrevaya-Dahl Mundlak quantile regression

Table 5: Linear panel data models

	OLS	Random effects	Fixed effects
Age	0.048	0.053	0.111
	(0.000)**	(0.000)**	(0.001)**
Age squared	-0.001	-0.001	-0.001
	(0.000)**	(0.000)**	(0.000)**
Woman	-0.154	-0.127	
	(0.001)**	(0.001)**	
Children aged 0-6 years	0.043	0.010	0.000
	(0.001)**	(0.001)**	(0.001)
Non-western immigrant	-0.020	-0.002	
	(0.002)**	(0.003)	
Large city	-0.069	-0.058	-0.020
	(0.001)**	(0.001)**	(0.003)**
Rural	-0.076	-0.066	-0.050
	(0.001)**	(0.001)**	(0.002)**
Experience	0.017	0.015	0.025
	(0.000)**	(0.000)**	(0.001)**
Experience squared	0.000	0.000	-0.001
	(0.000)**	(0.000)**	(0.000)**
Unskilled	-0.112	-0.276	
	(0.001)**	(0.001)**	
Short term higher education	0.057	-0.008	
	(0.001)**	(0.002)**	
Long term education	0.180	0.115	
	(0.002)**	(0.003)**	
Workplace size, 10-50 workers	0.040	0.033	0.027
	(0.001)**	(0.001)**	(0.001)**
Workplace size, 100-200 workers	0.074	0.067	0.052
	(0.001)**	(0.001)**	(0.001)**
Workplace size, 200+ workers	0.109	0.095	0.071
	(0.001)**	(0.001)**	(0.001)**
Firm-level bargaining	0.051	-0.002	-0.012
	(0.001)**	(0.002)	(0.003)**
Minimum pay	-0.001	-0.014	-0.021
	(0.001)	(0.001)**	(0.002)**
Minimum wage	0.040	0.002	0.004
	(0.001)**	(0.002)	(0.003)
Observations	1,016,389	1,016,389	1,016,389
R-squared	0.42		0.33
Number of pnr	320,561	320,561	320,561

Robust standard errors in parentheses. * Significant at 5%; ** Significant at 1%.

Table 6: Pooled quantile regression

	10	25	50	75	90	OLS
Age	0.067	0.048	0.030	0.027	0.028	0.046
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Age squared	-0.001	-0.001	0.000	0.000	0.000	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.00)**	(0.000)**
Woman	-0.122	-0.144	-0.164	-0.179	-0.196	-0.158
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Children aged 0-6 years	0.062	0.045	0.028	0.026	0.027	0.041
	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**
Non-western immigrant	-0.002	-0.022	-0.034	-0.047	-0.063	-0.030
	(0.006)	(0.005)**	(0.004)**	(0.005)**	(0.006)**	(0.004)**
Large city	-0.063	-0.074	-0.075	-0.076	-0.065	-0.070
	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.004)**	(0.002)**
Rural	-0.069	-0.082	-0.084	-0.082	-0.077	-0.076
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Experience	0.026	0.020	0.014	0.011	0.009	0.018
	(0.001)**	(0.000)**	(0.000)**	(0.001)**	(0.001)**	(0.000)**
Experience squared	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(.0000)**	(0.000)**	(0.000)**
Unskilled	-0.116	-0.109	-0.078	-0.067	-0.062	-0.111
	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Short term higher education	0.073	0.069	0.066	0.066	0.061	0.058
	(0.005)**	(0.004)**	(0.003)**	(0.004)**	(0.005)**	(0.003)**
Long term higher education	0.111	0.148	0.181	0.199	0.220	0.167
	(0.008)**	(0.006)**	(0.006)**	(0.006)**	(0.008)**	(0.005)**
Workplace size, 10-50 workers	0.033	0.036	0.035	0.038	0.036	0.039
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Workplace size, 100-200 workers	0.067	0.065	0.059	0.062	0.060	0.071
	(0.003)**	(0.003)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Workplace size, 200+ workers	0.104	0.104	0.098	0.096	0.083	0.106
	(0.003)**	(0.003)**	(0.002)**	(0.003)**	(0.003)**	(0.002)**
Firm-level bargaining	0.032	0.042	0.055	0.084	0.107	0.050
	(0.005)**	(0.004)**	(0.003)**	(0.004)**	(0.005)**	(0.003)**
Minimum pay	-0.003	-0.005	-0.003	0.010	0.028	-0.002
	(0.003)	(0.003)*	(0.002)	(0.003)**	(0.004)**	(0.002)**
Minimum wage	0.018	0.020	0.042	0.074	0.105	0.043
	(0.004)**	(0.003)**	(0.003)**	(0.004)**	(0.005)**	(0.003)**

Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 10,000 persons and 1,000 iterations. * Significant at 5%; ** Significant at 1%.

Table 7: Abrevaya-Dahl quantile regression, 2 observations

	10	25	50	75	90	OLS
Age	0.089	0.096	0.059	0.052	0.048	0.082
	(0.004)**	(0.003)**	(0.002)**	(0.003)**	(0.003)**	(0.002)**
Age squared	-0.001	-0.001	-0.001	0.000	0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Woman	-0.116	-0.133	-0.161	-0.181	-0.199	-0.153
	(0.002)**	(0.002)**	(0.001)**	(0.002)**	(0.002)**	(0.001)**
Children aged 0-6 years	0.063	0.053	0.032	0.026	0.028	0.042
	(0.002)**	(0.001)**	(0.001)**	(0.002)**	(0.002)**	(0.001)**
Non-western immigrant	0.008	-0.007	-0.023	-0.037	-0.050	-0.016
	(0.005)	(0.004)	(0.003)**	(0.004)**	(0.005)**	(0.003)**
Large city	-0.062	-0.072	-0.074	-0.072	-0.063	-0.068
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Rural	-0.068	-0.081	-0.080	-0.076	-0.071	-0.073
	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**
Experience	0.079	0.041	0.020	0.011	0.008	0.029
	(0.004)**	(0.003)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Experience squared	-0.001	-0.001	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Unskilled	-0.119	-0.128	-0.092	-0.073	-0.064	-0.122
	(0.002)**	(0.001)**	(0.002)**	(0.002)**	(0.002)**	(0.001)**
Short term higher education	0.063	0.063	0.059	0.058	0.056	0.052
	(0.004)**	(0.003)**	(0.003)**	(0.003)**	(0.005)**	(0.003)**
Long term higher education	0.129	0.166	0.182	0.205	0.226	0.175
	(0.006)**	(0.004)**	(0.004)**	(0.004)**	(0.007)**	(0.004)**
Workplace size, 0-50 workers	0.029	0.032	0.026	0.024	0.030	0.030
	(0.004)**	(0.003)**	(0.003)**	(0.003)**	(0.004)**	(0.002)**
Workplace size, 100-200 workers	0.050	0.055	0.049	0.048	0.062	0.056
	(0.004)**	(0.004)**	(0.003)**	(0.003)**	(0.004)**	(0.003)**
Workplace size, 200+ workers	0.077	0.075	0.071	0.071	0.081	0.078
	(0.005)**	(0.004)**	(0.004)**	(0.004)**	(0.005)**	(0.003)**
Firm-level bargaining	-0.026	-0.020	-0.002	0.017	0.039	0.007
	(0.009)**	(0.007)**	(0.006)	(0.007)**	(0.010)**	(0.005)
Minimum pay	-0.010	-0.007	-0.009	-0.002	0.013	0.000
	(0.008)	(0.007)	(0.006)	(0.006)	(0.009)	(0.005)
Minimum wage	-0.007	0.020	0.022	0.016	0.031	0.022
	(0.009)	(0.008)**	(0.006)**	(0.007)*	(0.010)**	(0.006)**

Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 10,000 persons and 1,000 iterations. * Significant at 5%; ** Significant at 1%.

Table 8: Abrevaya-Dahl quantile regression, 4 observations

	10	25	50	75	90	OLS
Age	0.108	0.075	0.051	0.045	0.046	0.078
	(0.008)**	(0.005)**	(0.004)**	(0.005)**	(0.005)**	(0.004)**
Age squared	-0.001	-0.001	0.000	0.000	0.000	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Woman	-0.131	-0.153	-0.174	-0.189	-0.204	-0.169
	(0.004)**	(0.002)**	(0.002)**	(0.003)**	(0.004)**	(0.002)**
Children aged 0-6 years	0.043	0.027	0.020	0.020	0.023	0.031
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Non-western immigrant	-0.001	-0.024	-0.032	-0.040	-0.057	-0.032
	(0.008)	(0.006)**	(0.006)**	(0.007)**	(0.007)**	(0.006)**
Large city	-0.065	-0.072	-0.071	-0.070	-0.057	-0.066
	(0.004)**	(0.003)**	(0.003)**	(0.003)**	(0.005)**	(0.003)**
Rural	-0.074	-0.082	-0.083	-0.080	-0.073	-0.076
	(0.003)**	(0.003)**	(0.002)**	(0.002)**	(0.004)**	(0.002)**
Experience	0.055	0.020	0.007	0.005	0.001	0.017
	(0.008)**	(0.004)**	(0.003)*	(0.003)	(0.004)	(0.003)**
Experience squared	-0.001	-0.001	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Unskilled	-0.095	-0.078	-0.059	-0.054	-0.051	-0.088
	(0.002)**	(0.002)**	(0.003)**	(0.004)**	(0.004)**	(0.001)**
Short term higher education	0.065	0.065	0.060	0.061	0.053	0.055
	(0.005)**	(0.004)**	(0.004)**	(0.005)**	(0.006)**	(0.004)**
Long term higher education	0.106	0.147	0.186	0.205	0.230	0.178
	(0.011)**	(0.007)**	(0.007)**	(0.007)**	(0.013)**	(0.007)**
Workplace size, 0-50 workers	0.024	0.025	0.024	0.025	0.022	0.025
	(0.004)**	(0.003)**	(0.002)**	(0.003)**	(0.004)**	(0.002)**
Workplace size, 100-200 workers	0.047	0.041	0.041	0.045	0.040	0.046
	(0.004)**	(0.003)**	(0.003)**	(0.003)**	(0.005)**	(0.003)**
Workplace size, 200+ workers	0.063	0.061	0.061	0.062	0.047	0.062
	(0.005)**	(0.004)**	(0.003)**	(0.004)**	(0.006)**	(0.003)**
Firm-level bargaining	-0.046	-0.026	-0.004	0.011	0.043	0.002
	(0.010)**	(0.008)**	(0.008)	(0.009)	(0.013)**	(0.007)
Minimum pay	-0.024	-0.021	-0.014	-0.005	0.024	-0.006
	(0.009)*	(0.008)**	(0.008)	(0.009)	(0.013)	(0.007)
Minimum wage	-0.020	0.015	0.012	0.016	0.040	0.015
	(0.012)	(0.009)	(0.008)	(0.010)	(0.014)*	(0.008)

Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 10,000 persons and 1,000 iterations. * Significant at 5%; ** Significant at 1%.

Table 9: Abrevaya-Dahl Mundlak quantile regression

	10	25	50	75	90	OLS
Age	0.112	0.098	0.060	0.056	0.057	0.086
	(0.005)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**
Age squared	-0.001	-0.001	0.000	0.000	0.000	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Woman	-0.125	-0.146	-0.167	-0.183	-0.200	-0.161
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.004)**	(0.002)**
Children aged 0-6 years	0.061	0.044	0.028	0.026	0.027	0.040
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**
Non-western immigrant	-0.002	-0.020	-0.031	-0.043	-0.060	-0.029
	(0.007)	(0.005)**	(0.004)**	(0.005)**	(0.006)**	(0.004)**
Large city	-0.063	-0.072	-0.073	-0.074	-0.063	-0.068
	(0.004)**	(0.003)**	(0.002)**	(0.003)**	(0.004)**	(0.003)**
Rural	-0.068	-0.080	-0.081	-0.080	-0.073	-0.074
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Experience	0.054	0.027	0.014	0.009	0.004	0.023
	(0.004)**	(0.002)**	(0.002)**	(0.003)**	(0.003)	(0.002)**
Experience squared	-0.001	-0.001	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Unskilled	-0.114	-0.108	-0.075	-0.065	-0.060	-0.109
	(0.002)**	(0.002)**	(0.003)**	(0.003)**	(0.004)**	(0.001)**
Short term higher education	0.065	0.064	0.059	0.058	0.055	0.053
	(0.006)**	(0.004)**	(0.004)**	(0.004)**	(0.006)**	(0.004)**
Long term higher education	0.109	0.143	0.173	0.188	0.212	0.163
	(0.008)**	(0.006)**	(0.005)**	(0.006)**	(0.009)**	(0.005)**
Workplace size, 0-50 workers	0.020	0.023	0.022	0.022	0.022	0.023
	(0.003)**	(0.002)**	(0.002)**	(0.003)**	(0.004)**	(0.002)**
Workplace size, 100-200 workers	0.045	0.043	0.042	0.042	0.043	0.046
	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.004)**	(0.002)**
Workplace size, 200+ workers	0.064	0.061	0.062	0.060	0.055	0.065
	(0.004)**	(0.003)**	(0.003)**	(0.003)**	(0.005)**	(0.003)**
Firm-level bargaining	-0.030	-0.022	0.000	0.015	0.039	0.004
	(0.009)**	(0.008)**	(0.006)	(0.007)*	(0.010)**	(0.006)
Minimum pay	-0.018	-0.018	-0.010	0.000	0.017	-0.007
	(0.008)*	(0.007)*	(0.006)	(0.007)	(0.009)	(0.005)
Minimum wage	-0.017	0.008	0.016	0.018	0.047	0.017
	(0.009)*	(0.008)	(0.006)*	(0.008)*	(0.010)**	(0.006)**

Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 10,000 persons and 1,000 iterations. * Significant at 5%; ** Significant at 1%.

Table 10: Abrevaya-Dahl Mundlak quantile regression with change dummies

	10	25	50	75	90	OLS
Age	0.112	0.098	0.060	0.056	0.057	0.085
	(0.005)**	(0.003)**	(0.003)**	(0.003)**	(0.004)**	(0.003)**
Age squared	-0.001	-0.001	0.000	0.000	0.000	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Woman	-0.126	-0.146	-0.167	-0.184	-0.200	-0.161
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.004)**	(0.002)**
Children 0-6 years	0.061	0.044	0.028	0.026	0.027	0.040
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Non-western immigrant	-0.003	-0.021	-0.031	-0.044	-0.060	-0.029
	(0.007)	(0.005)**	(0.004)**	(0.004)**	(0.006)**	(0.004)**
Large city	-0.063	-0.072	-0.073	-0.074	-0.063	-0.068
	(0.004)**	(0.003)**	(0.003)**	(0.003)**	(0.004)**	(0.003)**
Rural	-0.067	-0.080	-0.081	-0.080	-0.073	-0.074
	(0.003)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Experience	0.054	0.027	0.014	0.009	0.004	0.023
	(0.004)**	(0.003)**	(0.002)**	(0.003)**	(0.003)**	(0.002)**
Experience squared	-0.001	-0.001	-0.000	-0.000	-0.000	-0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Unskilled	-0.114	-0.108	-0.075	-0.065	-0.060	-0.109
	(0.002)**	(0.002)**	(0.003)**	(0.003)**	(0.004)**	(0.001)**
Short term higher education	0.065	0.064	0.059	0.058	0.055	0.053
	(0.005)**	(0.004)**	(0.004)**	(0.004)**	(0.006)**	(0.004)**
Long term higher education	0.109	0.143	0.174	0.188	0.212	0.163
	(0.008)**	(0.006)**	(0.005)**	(0.006)**	(0.009)**	(0.005)**
Workplace size, 0-50 workers	0.020	0.023	0.021	0.023	0.022	0.023
	(0.003)**	(0.002)**	(0.002)**	(0.003)**	(0.004)**	(0.002)**
Workplace size, 50-200 workers	0.045	0.043	0.042	0.042	0.043	0.046
	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.004)**	(0.002)**
Workplace size, 200+ workers	0.064	0.061	0.062	0.060	0.055	0.064
	(0.005)**	(0.003)**	(0.003)**	(0.004)**	(0.005)**	(0.003)**
Firm-level bargaining	-0.048	-0.023	0.000	0.005	0.024	-0.002
	(0.012)**	(0.010)*	(0.009)	(0.009)	(0.012)*	(0.008)
Minimum pay	-0.036	-0.018	-0.011	-0.011	0.002	-0.013
	(0.011)**	(0.010)	(0.009)	(0.009)	(0.012)	(0.008)
Minimum wage	-0.035	0.007	0.015	0.007	0.033	0.010
	(0.011)**	(0.011)	(0.009)	(0.009)	(0.013)**	(0.008)
Change standard rate to minimum pay	0.014	-0.003	-0.006	0.011	0.017	0.010
	(0.010)	(0.009)	(0.008)	(0.008)	(0.010)	(0.007)
Change minimum pay to standard rate	-0.022	-0.004	-0.009	-0.016	-0.022	-0.005
	(0.012)	(0.010)	(0.009)	(0.009)	(0.014)	(0.009)

Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 10,000 persons and 1,000 iterations. * Significant at 5%; ** Significant at 1%.

Table 11: Abrevaya-Dahl Mundlak quantile regression, sample with exogenous changes

	10	25	50	75	90	OLS
Age	0.113	0.085	0.057	0.053	0.054	0.082
	(0.006)**	(0.003)**	(0.003)**	(0.003)**	(0.004)**	(0.003)**
Age squared	-0.001	-0.001	0.000	0.000	0.000	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Woman	-0.124	-0.148	-0.169	-0.186	-0.202	-0.163
	(0.003)**	(0.002)**	(0.002)**	(0.003)**	(0.003)**	(0.002)**
Children 0-6 years	0.059	0.037	0.025	0.025	0.026	0.037
	(0.003)**	(0.002)**	(0.002)**	(0.003)**	(0.003)**	(0.002)**
Non-western immigrant	-0.008	-0.026	-0.036	-0.049	-0.057	-0.033
	(0.007)	(0.006)**	(0.006)**	(0.007)**	(0.008)**	(0.006)**
Large city	-0.065	-0.076	-0.075	-0.074	-0.063	-0.071
	(0.005)**	(0.003)**	(0.003)**	(0.004)**	(0.005)**	(0.003)**
Rural	-0.070	-0.085	-0.084	-0.081	-0.073	-0.076
	(0.003)**	(0.002)**	(0.002)**	(0.003)**	(0.004)**	(0.002)**
Experience	0.046	0.021	0.009	0.006	0.003	0.017
	(0.005)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.002)**
Experience squared	-0.001	-0.001	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Unskilled	-0.115	-0.101	-0.071	-0.062	-0.054	-0.104
	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.002)**
Short term higher education	0.063	0.062	0.059	0.060	0.057	0.053
	(0.006)**	(0.005)**	(0.004)**	(0.005)**	(0.007)**	(0.004)**
Long term higher education	0.110	0.139	0.169	0.185	0.205	0.159
	(0.010)**	(0.008)**	(0.006)**	(0.007)**	(0.010)**	(0.007)**
Workplace size, 0-50 workers	0.023	0.022	0.023	0.024	0.021	0.023
	(0.004)**	(0.004)**	(0.003)**	(0.003)**	(0.004)**	(0.003)**
Workplace size, 100-200 workers	0.051	0.040	0.044	0.047	0.048	0.048
	(0.006)**	(0.004)**	(0.003)**	(0.003)**	(0.004)**	(0.003)**
Workplace size, 200+ workers	0.066	0.057	0.066	0.060	0.051	0.062
	(0.006)**	(0.004)**	(0.004)**	(0.004)**	(0.006)**	(0.004)**
Firm-level bargaining	-0.039	-0.009	0.007	0.024	0.039	0.013
	(0.011)**	(0.008)	(0.008)	(0.009)*	(0.013)**	(0.007)
Minimum pay	-0.032	-0.014	-0.010	00	0.012	-0.005
	(0.011)**	(0.008)	(0.007)	(0.009)	(0.013)	(0.006)
Minimum wage	-0.034	0.023	0.019	0.025	0.050	0.025
	(0.012)**	(0.009)*	(0.008)*	(0.011)*	(0.015)**	(0.008)**

Bootstrapped standard errors are given in parentheses. The bootstrap was done using a sample size of 10,000 persons and 1,000 iterations. * Significant at 5%; ** Significant at 1%.

Job Sampling and Sorting*

Daniel le Maire
University of Copenhagen and CAM

Christian Scheuer
Copenhagen Business School and CEBR

November 6, 2008

Abstract

We propose an equilibrium search model with on-the-job search, where a continuum of heterogenous workers and continuum of heterogenous firms match. Workers are allowed to select a sample size of vacancies to apply to, and more productive workers sample more firms. The possibility of sampling job offers imply that we only need the production function to be strictly supermodular for assortative matching to arise. The model delivers a log-linear wage equation with additively separable worker and firm effects. Therefore, we can simulate the theoretical model and estimate the two-way fixed-effects model as Abowd, Kramarz and Margolis (1999). By adding a match-effect we obtain an estimated negative correlation between the worker and firm effects even though there is assortative matching (positive correlation) in the theoretical model.

Keywords: Assortative matching, labor market search, linked employer-employee data, person and firm effects.

JEL Classification: C23, C78, J31, J62, J64.

*We are very grateful for helpful comments from Dale Mortensen. Furthermore, we also thank Karsten Albæk, Martin Browning, Fane Groes, Bo Honore, Peter Sørensen and Chris Taber for valuable comments. All remaining errors are ours.

1 Introduction

In this paper we propose a labor market search model with two new features. First, the model has three types of continuous heterogeneity, which are worker productivity, firm productivity, and a random match effect. Each of these three productivity terms enters the production function. Second, assortative matching arises even though we only use a strictly supermodular production function.

We solve our search model and simulate data from the model. Using the two-way-fixed effect estimator invented by Abowd, Kramarz and Margolis (1999) (henceforth, AKM) on this simulated data, we show that the presence of a match effect can imply an estimated negative correlation between the worker and firm effects even if there is true positive assortative matching in the model economy. Therefore, this paper aims at explaining the puzzling negative correlation between worker and firm effects found empirically.[1]

Existing models of assignment along the lines of Becker's (1973) marriage model imply positive assortative matching. In fact, the ranking is such that a more productive worker always will have a better job than a less productive one. This result holds in a frictionless economy as long as the production function is strictly supermodular implying that the worker's human capital and firm's capital are complements in production. However, Shimer and Smith (2000) show that for assortative matching to arise in a search model setting it is no longer sufficient that the production function is strictly supermodular. Instead, Shimer and Smith find that a sufficient condition is that the production function is strictly log supermodular. As noted by Atakan (2006) the problem is that the higher gains to more search for more productive workers are offset by higher costs of rejecting an offer. Using this, Atakan shows that if unmatched agents have constant flow costs independent of their type and if the discount rate is zero, the production function only needs to be strictly supermodular in order to imply assortative matching.

We obtain assortative matching with the strictly supermodular Cobb-Douglas production function by letting workers choose how many job offers they wish to sample as in the seminal contribution by Stigler (1961). Assortative matching arises since more productive workers sample more jobs and, hence, they will, on average, end up in better jobs. In order to isolate the effect of the workers' job sampling we - similarly to Acemoglu and Shimer (2000) - assume that each firm that receives more than one application randomly chooses one of its applicants above a reservation threshold.

Previous studies on assortative models have examined the sets of matches that are acceptable

[1] See Abowd, Creecy and Kramarz (2002) and Abowd, Kramarz, Lengermann and Perez-Durate (2004) for results for both the US and France, Gruetter and Lalive (2004) for results for Austria, Piekkola (2005) for Finland, Andrews, Gill, Schank and Upward (2007) for results for Germany, and Barth and Dale-Olsen (2003) for results for Norway.

by both workers and firms. Most of the studies (e.g. Atakan (2006), Becker (1973), Burdett and Coles (1999), Chade (2001) and Smith (2006)) have found perfect segregation, which implies that the labor market is segmented into multiple non-overlapping markets. Such an approach to study assortative matching is not useful with on-the-job search where the same search technology is available as unemployed and employed, since workers accept any offer greater or equal to the benefit level and subsequently climb the wage ladder by searching on-the-job. Instead of the matching sets approach, we compare the worker distribution over different firm and match productivities conditional on worker type. The latter approach is also used in Lentz (2008), where assortative matching also arises with a strictly supermodular production function since workers are allowed to choose their own search intensity.

When heterogeneous workers and firms match, it seems obvious that both worker and firm productivity should affect the wage paid. However, there are no a priori reasons to believe that these two effects should capture all variation in wages. It is very likely that complementarities between specific workers and firms could exist, or that human capital is accumulated according to the quality of the match between the worker and the firm involved. Both suggest a role for a heterogeneous match specific component. Moreover, several contributions have argued that the quality of the match between workers and firms in itself influences the earnings variation. A prominent example is the search model in Jovanovic (1979), where the flow production is a match quality plus a stochastic term. However, in principle there need not to be a match productivity in order to have a match component of the wage. The match effect could be due to differences in bargaining strength, for instance, as a consequence of the labor market tightness at the time of contract negotiation.

In the proposed model framework we let the match effect be a random productivity component, which enters the production function together with the worker and firm productivity. To our knowledge the model we propose is the first to offer a theoretical framework encompassing the idea of heterogeneity in both firm, worker and match effects.

A useful feature of our theoretical model is that it also implies a log linear wage equation with additively separable worker, firm and match heterogeneity. This is only possible because we can relax the requirement for assortative matching to be strictly supermodularity of the production function. The log linear wage equation with additively separable effects enables us to link the theoretical model directly to the AKM model as also done by Abowd, Kramarz, Lengermann and Perez-Duarte (2004). However, since the log linear wage equation is built into an equilibrium search model with assortative matching it gives us the opportunity to simulate our theoretical model and perform the AKM estimation on this data. When we simulate our model without the match effect and perform the AKM estimation and apply the bias correction in Andrews, Gill, Schank and Upward (2007) the estimated correlation equals the true correlation. In contrast, simulating our model with the match effect and again estimating the AKM model and bias-correcting we obtain the very interesting result, that even though data is generated from a model with positive assortative matching the estimated correlation is negative.

The intuition behind our results is as follows. The firm effects are identified by workers making job-to-job transitions, and those with relative low realizations of the match effect will ceteris paribus be more likely to change jobs. However, since the econometrician does not estimate the match effect, the firm effect will be under-valued, which in turn will imply an over-valued worker effect, and the spurious negative correlation is established.

These results are in line with Woodcock (2007) who suggests an estimation procedure for an empirical model with person, firm and match effects. Using this estimation procedure he finds that the estimated negative correlation of the AKM model on US-data, in fact, is positive when taking the match heterogeneity into account. We apply his estimator both to our simulated and empirical data. From the simulations we learn, that the model performs very well in attributing the different parts of variation in the dependent variable to the firm, person and match effects, but also this estimator consequently underestimates the true correlation between workers and firms.

We apply both set of estimators to a panel of Danish employer-employee data. We estimate that the match effect indeed has empirical relevance since it accounts for 15 per cent of the variation in log wages. Furthermore, we estimate (what consequently is a downward biased) a correlation of worker and firm effects of 12 per cent, which suggests that the Danish labor market is characterized by positive assortative matching.

The paper is organized as follows. In section 2, we present our search model where more productive workers sort themselves into more productive matches. In section 3 we briefly consider the AKM model and the recent alternative estimation procedure outlined in Woodcock (2007). In section 4, we simulate our theoretical model and estimate both the AKM and Woodcock models on this model-generated data, whereas we in section 5 perform estimations on Danish register data. In section 6, we conclude.

2 Theoretical Model

Consider an economy with a continuum of firms and a continuum of workers that participate in the labor market. Both the measure of workers and firms are fixed and normalized to 1. All agents are rational, forward-looking, risk-neutral and infinitely lived. Workers have the opportunity of searching both while they are unemployed and employed.

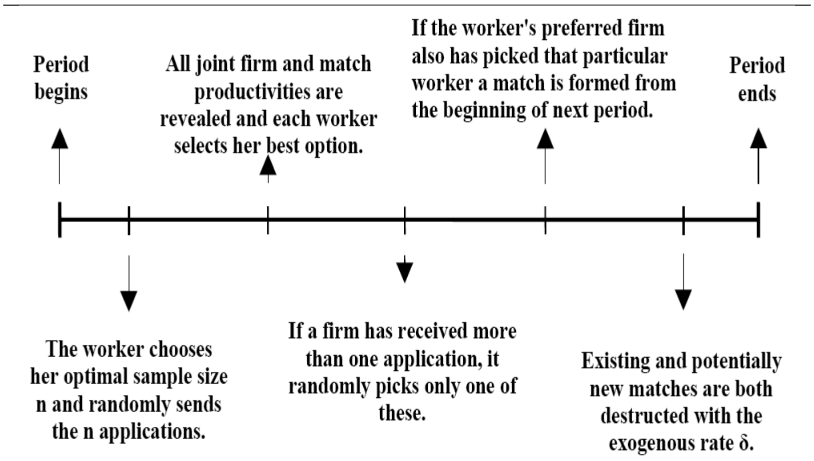
Heterogeneity exists on both sides of the market as well as in the match between a worker and a firm. Workers differ in respect to their productivity such that more productive workers in all types of jobs are more productive. Denote the worker productivity by p_w, the firm productivity by p_f and the match effect by p_m. Both the worker and the firm know their own productivity term as well as the decomposition in a given match and, hence, we rule out uncertainty and learning about any of the productivity terms.

We assume that we have a Cobb-Douglas production function $f(p_w, p_f, p_m) = p_w^{\alpha_1} p_f^{\alpha_2} p_m^{\alpha_3}$. The worker searches for the highest possible combination of firm and match productivity. Therefore, it is useful for us to work with the joint firm and match productivity $p_{fm} = p_f^{\frac{\alpha_2}{\alpha_2+\alpha_3}} p_m^{\frac{\alpha_3}{\alpha_2+\alpha_3}}$, allowing us to express the production function as $f(p_w, p_f, p_m) = f(p_w, p_{fm}) = p_w^{\alpha_1} p_{fm}^{\alpha_2+\alpha_3}$. The Cobb-Douglas function has two important properties. First, it is strictly supermodular, that is, $\frac{\partial^2 f(p_w, p_{fm})}{\partial p_w \partial p_{fm}} > 0$ for $\alpha_1 > 0$ and $\alpha_2 + \alpha_3 > 0$. Second, it is multiplicatively separable. We need not to assume the Cobb-Douglas form, and all implications of the model are true for all functions admitting strictly supermodularity and multiplicative separability. We use the assumption of multiplicative separability to ease the derivation of the worker's reservation productivity below, but note also that this assumption rules out comparative advantages, since the ratio of output of two firms is independent of the worker's productivity, cf. Sattinger (1975).

The distribution of worker productivity p_w is given by the cumulative distribution function $H(p_w) = \int_{\underline{p}_w}^{p_w} h(p'_w)\, dp'_w$. The distribution of the firm and match productivity p_{fm} is given by $\Gamma(p_{fm}) = \int_{\underline{p}_{fm}}^{\bar{p}_{fm}} \gamma(p'_{fm})\, dp'_{fm}$.

The search environment is closest to Acemoglu and Shimer (2000) although we allow for on-the-job search, but do not consider the firm's investment decision. The model is set in discrete time and the timing of events is illustrated in Figure 1.

FIGURE 1: ONE PERIOD SEQUENCE OF EVENTS



The job searcher, whether unemployed or employed, is given the opportunity to apply for n different jobs in each time period. Searching is costly and the more jobs applied to, the larger costs. We assume that this flow cost function is strictly convex in the number of jobs applied to, $c'(n) > 0$ and $c''(n) > 0$ and that $c(0) = c'(0) = 0$. The more jobs applied to, the more likely the worker is to draw a high joint firm and match productivity. In order for the worker to realize the joint firm and match productivity, the worker needs to apply for a job in the particular firm. Hence, the worker applies to all jobs sampled, although he/ she is only willing to match with the job which turns out to be most productive (if above a reservation productivity level).

Depending on the number of job applications and firms, it is likely that a firm gets more than one application in each discrete time interval. Each firm is only capable of hiring one worker and among the applicants, one application is chosen randomly. One could think of the firms' random choice of worker as a framework, where the firm is only able to use sequential search in continuous time, and where it is random, whose application is the first to arrive.[2]

The implication of this search environment is that the expected number of job applications received is the same for all firms and that the acceptance rate workers face is independent of the worker's productivity. If both the worker and firm choose each other a match will be formed from the beginning of next period. Finally, in the end of each period a fraction δ of existing and newly formed matches are exogenously destructed.

There is no traditional job arrival rate in this search context, but by abusing the traditional notation the worker's chance of getting his/ her chosen firm is denoted λ. That is

$$\lambda = (1-\delta)\frac{\#\text{firms}}{\#\text{ job applications}}$$

where the term $(1-\delta)$ takes account of that δ of the new matches are destroyed before they come to existence.

2.1 The Worker Side

As unemployed, the worker receives benefits which depend on her own productivity $bp_w^{\alpha 1}$. This takes account of the fact that benefits typically are dependent on previous income as employed. Furthermore, the exact specification used here will simplify the analysis.[3] With the usual notation we denote the value of unemployment $U(\cdot)$ and the value of employment $W(\cdot,\cdot)$. Letting r denote the discount factor the Bellman equation for an unemployed worker is

$$(1+r)U(p_w) = bp_w^{\alpha_1} - c(n) + \lambda\int_{\underline{p}_{fm}}^{\bar{p}_{fm}} \max\left\{W\left(p_w, p'_{fm}\right), U(p_w)\right\} dQ\left(p'_{fm}|n\right) + (1-\lambda)U(p_w) \tag{1}$$

[2] Restricting the firms to consider one worker at a time, makes the model tractable. If also allowing firms to sample workers it is no longer necessarily the case that all workers will prefer the highest joint firm and match productivity. Rather, less productive workers may have a higher chance of becoming employed in less productive firms and, thereby, prefer to join these. Hence, allowing for endogenous sampling on both sides of the market complicates the math, but will only tend to increase the degree of assortative matching.

[3] A similar assumption is used in Postel-Vinay and Robin (2002).

where $\bar{p}_{fm}$ and $\underline{p}_{fm}$ denote respectively the upper and lower points of the joint firm and match productivity distribution p_{fm}, and $Q(p_{fm}|n) = \Gamma(p_{fm})^n$ denote the distribution conditional on the number of sampled jobs n. Letting p^r_{fm} be the worker's reservation productivity and using integration by parts we can express this Bellman equation as

$$rU(p_w) = bp_w^{\alpha_1} - c(n) + \lambda \int_{p^r_{fm}}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} \tag{2}$$

As employed the worker earns the wage $w(p_w, p_{fm})$, and as wage setting rule we employ the simplest: linear output sharing or piece-rate contract as in for example Bagger, Fontaine, Postel-Vinay and Robin (2007): $w(p_w, p_{fm}) = \beta p_w^{\alpha_1} p_{fm}^{\alpha_2+\alpha_3}$ where β is the worker's share of the flow production. With probability δ the job is destructed and the worker returns to unemployment. The Bellman equation for an employed worker is

$$(r+\delta) W(p_w, p_{fm}) = \beta p_w^{\alpha_1} p_{fm}^{\alpha_2+\alpha_3} - c(n) + \lambda \int_{p_{fm}}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} + \delta U(p_w) \tag{3}$$

Both workers and firms have minimum values of productivities that they are willing to match with. On the worker side this reservation value is the firm and match productivity for which the worker is indifferent between being employed compared to staying unemployed. This can be derived as

$$\begin{gathered}
\beta p_w^{\alpha_1} \left(p^r_{fm}\right)^{\alpha_2+\alpha_3} - c(n) + \lambda \int_{p^r_{fm}}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} \\
= bp_w^{\alpha_1} - c(n) + \lambda \int_{p^r_{fm}}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} \\
\Leftrightarrow \\
p^r_{fm} = \left(\frac{b}{\beta}\right)^{\frac{1}{\alpha_2+\alpha_3}}
\end{gathered} \tag{4}$$

Thus, the firm reservation productivity p^r_{fm} is identical for all workers. This would not have been the case if all workers received the same amount of unemployment benefits. In that case more productive workers would have a lower joint firm and match reservation productivity than less productive workers, since more productive workers have higher opportunity costs of staying unemployed and, therefore, would be more eager to get a job.[4] With identical p^r_{fm} across worker type, the reservation wage is increasing in the worker productivity p_w.

Differentiating the Bellman equation for an unemployed worker (2) and substituting out $W'_{p_f}(p_w, p_f)$ from (3) gives us the first-order condition for the sample size n for an unemployed worker

[4] A similar result appears in the general model of Burdett and Coles (1999).

$$c'(n) = \lambda \int_{\left(\frac{b}{\beta}\right)^{\frac{1}{\alpha_2+\alpha_3}}}^{\bar{p}_{fm}} \frac{\beta(\alpha_2+\alpha_3) p_w^{\alpha_1} \left(p'_{fm}\right)^{\alpha_2+\alpha_3-1} \left[-\Gamma\left(p'_{fm}\right)^n \ln\left(\Gamma\left(p'_{fm}\right)\right)\right]}{r+\delta+\lambda\left(1-\Gamma\left(p'_{fm}\right)^n\right)} dp'_{fm} \tag{5}$$

Since $[-\Gamma(p_{fm})^n \ln(\Gamma(p_{fm}))] \geq 0$ for all p_{fm}, the right hand side is increasing in p_w for $\alpha_1, (\alpha_2+\alpha_3) > 0$. Since $c''(n) > 0$ and the r.h.s. is decreasing in n for $\Gamma(p_{fm}) \in]0,1[$, more productive workers sample more jobs, that is $n'_{p_w}(p_w, p_{fm}) > 0$. The restriction that $\alpha_1, (\alpha_2+\alpha_3) > 0$ corresponds to the production function having a positive cross derivative, $f''_{p_w,p_{fm}}(p_w, p_{fm}) > 0$, which again exactly is the requirement of complementarity (or supermodularity) in the production function for Becker's (1973) model implying assortative matching.

n only take on integer values and the n maximizing (5) is most likely not an integer. However, since the l.h.s. is increasing in n, while r.h.s. is decreasing in n, the optimal integer value of n is one of the two integers adjacent to n unless the optimal value of (5) itself is an integer.

Employed workers maximize the right hand side of equation (3) with respect to n

$$c'(n) = \lambda \int_{p_{fm}}^{\bar{p}_{fm}} \frac{\beta(\alpha_2+\alpha_3) p_w^{\alpha_1} \left(p'_{fm}\right)^{\alpha_2+\alpha_3-1} \left[-\Gamma\left(p'_{fm}\right)^n \ln\left(\Gamma\left(p'_{fm}\right)\right)\right]}{r+\delta+\lambda\left(1-\Gamma\left(p'_{fm}\right)^n\right)} dp'_{fm} \tag{6}$$

For a given joint match and firm productivity, more productive workers sample more jobs than less productive workers as long as $\alpha_1, (\alpha_2+\alpha_3) > 0$ corresponding to the production function being strictly supermodular.

We can summarize the findings above in the following proposition:

Proposition 1 *When the production function $f(p_w, p_{fm})$ admits supermodularity such that $\alpha_1, (\alpha_2+\alpha_3) > 0$, more productive workers sample more jobs conditional on their current joint firm and match productivity $n'_{p_w}(p_w, p_{fm}) > 0$.*

Proof. See the text above. ■

Since workers are employed at $p_{fm} \geq \left(\frac{b}{\beta}\right)^{\frac{1}{\alpha_2+\alpha_3}}$ they have smaller expected gains of searching than unemployed workers with the same productivity and consequently they search less. Therefore, it is not necessarily the case that more productive employed workers search more than less productive employed workers since, on average, they will be employed in more productive matches already in the first match after being unemployed.

2.2 The Firm-Side

Unemployed workers choose the firm with the highest productivity among the sampled productivities given $p_{fm} \geq p^r_{fm}$, while employed workers only accept to join a firm of productivity p_{fm} if their current productivity is lower and if p_{fm} is the highest productivity sampled. The firm picks the worker at random given that the worker's productivity together with the drawn match effect is above a reservation threshold. We will express the reservation threshold in terms of the match effect, but make it depend on the worker productivity, that is $p_m^r(p_w)$. The firm's reservation productivity is decreasing in p_w, such that the highest reservation match effect is at the lower support of the worker distribution. To make the model tractable we assume that workers are accepted at any firm. More formally,

Assumption 1 $p_m^r\left(\underline{p}_w\right) \leq \underline{p}_m$ for all $p_f \in \left[\underline{p}_f, \bar{p}_f\right]$, where $\underline{p}_m$ is the lowest possible match effect.

Rather than matching with a worker of a given p_w, a firm may actually prefer to match with workers of lower productivity since workers with higher productivity sample more jobs and, hence, are more likely to draw a better firm and leave. In other words the expected value of a match may not be an increasing function of worker productivity. Nevertheless, since a firm selects its workers randomly it will be willing to hire a high productivity worker, since this is better than not hiring anyone.

2.3 Steady State

Since firms choose workers randomly, we can express the inflow to and outflow from unemployment as in usual search models. In steady-state the inflow and outflow must balance

$$\delta(1-u) = \lambda u \tag{7}$$

The mass $G(p_w, p_{fm})$ is the share of the employed workers with a worker productivity less or equal to p_w working at a joint firm and match productivity less or equal to p_{fm}. The flow into $G(p_w, p_{fm})$ must equal the outflow in steady-state. The outflow, on the l.h.s. below, consists of two terms, the exogenous destruction which happens at the rate δ and the endogenous job quits. When considering outflow from $G(p_w, p_{fm})$ in the form of quits, we - by definition - only consider workers with productivity p_w or less, who leaves a job with productivity equal or below p_{fm} and gets a job with productivity above p_{fm}. Since workers only change to more productive matches, inflow into $G(p_w, p_{fm})$ only comes from unemployment, where the number of jobs sampled is $n\left(p_{w,}p_{fm}^r\right)$. The steady-state condition is given by

$$\begin{aligned} & \delta(1-u)G(p_w, p_{fm}) + (1-u)\lambda \int_{\underline{p}_w}^{p_w}\int_{\underline{p}_{fm}}^{p_{fm}} \left(\Gamma(\bar{p}_{fm})^{n\left(p'_w, p'_{fm}\right)} - \Gamma(p_{fm})^{n\left(p'_w, p'_{fm}\right)}\right) g\left(p'_w, p'_{fm}\right) dp'_{fm} dp'_w \\ = \; & u\lambda \int_{\underline{p}_w}^{p_w} \left(\Gamma(p_{fm})^{n\left(p'_w, p^r_{fm}\right)} - \Gamma\left(\underline{p}_{fm}\right)^{n\left(p'_w, p^r_{fm}\right)}\right) h\left(p'_w\right) dp'_w \end{aligned} \tag{8}$$

Rearranging and using the equilibrium equation for unemployment gives us

$$\int_{\underline{p}_w}^{p_w}\int_{\underline{p}_{fm}}^{p_{fm}}\left[\delta+\lambda\left(1-\Gamma\left(p_{fm}\right)^{n\left(p'_w,p'_{fm}\right)}\right)\right]g\left(p'_w,p'_{fm}\right)dp'_{fm}dp'_w=\delta\int_{\underline{p}_w}^{p_w}\Gamma\left(p_{fm}\right)^{n\left(p'_w,p^r_{fm}\right)}h\left(p'_w\right)dp'_w$$

Differentiating this with respect to p_w gives

$$\int_{\underline{p}_{fm}}^{p_{fm}}\left[\delta+\lambda\left(1-\Gamma\left(p_{fm}\right)^{n\left(p_w,p'_{fm}\right)}\right)\right]g\left(p_w,p'_{fm}\right)dp'_{fm}=\delta\Gamma\left(p_{fm}\right)^{n\left(p_w,p^r_{fm}\right)}h\left(p_w\right) \tag{9}$$

Evaluating this expression in $\bar{p}_{fm}$ obviously gives $\int_{\underline{p}_{fm}}^{\bar{p}_{fm}}g\left(p_w,p'_{fm}\right)dp'_{fm}=h\left(p_w\right)$, and when combining this with equation (9), we obtain

$$\int_{\underline{p}_{fm}}^{p_{fm}}\tilde{g}\left(p_w,p'_{fm}\right)dp'_{fm}=\frac{\delta}{\delta+\lambda}\Gamma\left(p_{fm}\right)^{n\left(p_w,p^r_{fm}\right)}+\frac{\lambda}{\delta+\lambda}\int_{\underline{p}_{fm}}^{p_{fm}}\Gamma\left(p_{fm}\right)^{n\left(p_w,p'_{fm}\right)}\tilde{g}\left(p_w,p'_{fm}\right)dp'_{fm} \tag{10}$$

where $\int_{\underline{p}_{fm}}^{p_{fm}}\tilde{g}\left(p_w,p'_{fm}\right)dp'_{fm}\equiv\int_{\underline{p}_{fm}}^{p_{fm}}g\left(p_w,p'_{fm}\right)dp_{fm}\left(\int_{\underline{p}_{fm}}^{\bar{p}_{fm}}g\left(p_w,p'_{fm}\right)dp'_{fm}\right)^{-1}$ is the cumulative distribution function of firm productivities conditional on worker skill.

We want to examine whether the allocation of workers in the economy reflects positive sorting. While previous studies on assortative models have studied the sets of matches that are acceptable by both workers and firms, this approach is not useful with on-the-job search. Instead, we compare the worker distribution over different firms conditional on worker type similarly to Lentz (2008). For this purpose we use the following definition:

Definition 1 *Consider two workers A and B, where $p_w^A>p_w^B$. Assortative matching implies that $G^B\left(p_{fm}\right)\equiv\int_{\underline{p}_{fm}}^{p_{fm}}\tilde{g}\left(p_w^B,p'_{fm}\right)dp'_{fm}\geq\int_{\underline{p}_{fm}}^{p_{fm}}\tilde{g}\left(p_w^A,p'_{fm}\right)dp'_{fm}\equiv G^A\left(p_{fm}\right)$ for $p_{fm}\in\left]\underline{p}_{fm};p_{fm}\right[$.*

In the present model set-up, it is not immediately clear on the grounds of equation (10) whether or not the model framework implies assortative matching. Before we more formally prove this, we can gain some intuition by assuming that the number of jobs sampled is independent of the current firm and match productivity, that is $n\left(p_w,p_{fm}\right)=n\left(p_w\right)$. With this assumption it is clear that we have

$$\int_{\underline{p}_{fm}}^{p_{fm}}\tilde{g}\left(p_w,p_{fm}|n\left(p_w\right)=n\left(p_w,p_{fm}\right)\right)dp'_{fm}=\frac{\delta\Gamma\left(p_{fm}\right)^{n\left(p_w\right)}}{\delta+\lambda\left(1-\Gamma\left(p_{fm}\right)^{n\left(p_w\right)}\right)} \tag{11}$$

which for the number of jobs sampled being constant across worker type, that is $n\left(p_w\right)=1$ gives us the usual distribution of realized matches (see e.g. Postel-Vinay and Robin (2002)). If equation (11) is differentiated with respect to p_w, we see that the r.h.s. becomes negative implying assortative matching. However, assuming that $n\left(p_w,p_{fm}\right)=n\left(p_w\right)$ will overvalue the degree of assortative matching since it does not take into account that workers in better matches searches less.

A formal proposition for assortative matching is below. The proof uses a discretized version of equation (10) and is in Appendix A.[5]

Proposition 2 *When the production function $f(p_w, p_{fm})$ admits supermodularity such that $\alpha_1, (\alpha_2 + \alpha_3) > 0$, and the workers are allowed to sample $n = 0, 1, 2$ jobs the model features assortative matching.*

Proof. See appendix A. ■

Furthermore, the steady-state equilibrium is a tuple $\left\{p^r_{fm}, n(p_w, p_{fm}), n\left(p_w, p^r_{fm}\right), G(p_w, p_{fm}), u\right\}$ satisfying equations (4), (5), (6), (7) and (8), which leads us to the following proposition.

Proposition 3 *There exists a unique steady-state.*

Proof. See appendix B. ■

3 Econometric Methodology

3.1 The Two-Way Fixed Effects Model

Our theoretical model implies a log-linear wage equation where the log of worker productivity and the log of firm productivity are additively separable. This property is very convenient, since it aligns perfectly with the leading empirical model for estimating the degree of assortative matching in the labor market developed by Abowd, Kramarz and Margolis (1999) (henceforth AKM). Before we exploit this direct link between the theoretical and empirical model, we will give a brief introduction to the AKM model which takes the following form

$$y_{it} = x_{it}\beta + \theta_i + \psi_{J(i,t)} + \varepsilon_{it} \qquad i = 1, ..., N; \quad t = 1, .., T_i \tag{12}$$

where y_{it} is the log of the hourly wage rate for individual i in period t, x_{it} is a $1 \times k$ vector of time-varying explanatory variables that may both relate to the individual and the firm, β is the parameter vector, θ_i is the unobserved person effect, $\psi_{J(i,t)}$ is the unobserved firm effect, and ε_{it} is the error term with $E\left(\varepsilon_{it}|x_{it}, \theta_i, \psi_{J(i,t)}\right) = 0$. The function $J(i,t)$ associates an employer, indexed by J, with an individual i at time t.

Since most empirical applications relate to data with more individuals than firms $(N > J)$, we begin by making a within-individual transformation that sweeps away θ_i. Expressing the

[5] This proposition only encompass $n = 0, 1, 2$ since allowing for higher n makes the math very cumbersome. By computer simulation it is fairly straight-forward to show that this property of assortative matching holds as we let the maximum number of jobs sampled increase far above 2. These programs are available from the authors upon request.

model in matrix form, using $\sim$ to denote within transformed variables and letting D be the matrix of firm dummies gives us the following model to estimate

$$\tilde{Y} = \tilde{X}\beta + \tilde{D}\Psi + \tilde{\varepsilon} \tag{13}$$

where Ψ is a vector of firm effects for every firm in the sample.[6] The parameter estimates can be found by solving

$$\begin{bmatrix} \hat{\beta} \\ \hat{\Psi} \end{bmatrix} = \begin{bmatrix} \tilde{X}'\tilde{X} & \tilde{X}'\tilde{D} \\ \tilde{D}'\tilde{X} & \tilde{D}'\tilde{D} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{X}'\tilde{Y} \\ \tilde{D}'\tilde{Y} \end{bmatrix} \tag{14}$$

The problem of estimating the resulting model is that the cross-product matrix is potentially very high-dimensional due to $\tilde{D}'\tilde{D}$ containing a dummy for each firm. However, with use of sparse matrix algebra we can estimate the $\left[\hat{\beta}, \hat{\Psi}\right]$ and afterwards recover $\hat{\theta}$.

One key aspect of this estimator worth noticing is that the firm coefficient is only identified by workers moving between firms in the sample period. Hence, looking only at a worker employed in the same firm in all years there is no way for the econometrician to disentangle the person and firm fixed element of log wages.

As in all fixed effects models the variances of the fixed effects have a positive bias. Furthermore, due to the additive structure the covariance of the estimated worker and firm effects will be biased, since an over-estimate of the one fixed effect will lead to an under-estimate of the other. Andrews, Gill, Schank and Upward (2007) develop the formulae to correct for these biases. We use their method of correcting the estimates under the assumption that the explanatory variables are uncorrelated with the worker and firm effects.[7]

3.2 Woodcock's Hybrid Mixed Effects Estimation

Woodcock (2007) argues that the presence of a match effect will bias the correlation between worker and firm effects, but his agenda is broader than this, since omitted match effects will bias all estimates unless the match effect is completely orthogonal. Estimating an AKM model which also includes a match effect is impossible and, therefore, Woodcock suggest using the mixed effects model. Instead of estimating all the individual effects, Woodcock's estimation procedure implies estimating the variances of the worker, firm, match effects and error term and subsequently predicting all the individual effects. Woodcock needs to assume that the three

[6] Identification of firm effects is in principle only possible within a group, where a group is defined by the movement of workers between firms. For a thorough discussion see Abowd, Creecy and Kramarz (2002). For expositional simplicity we assume that we already have identified the groups, dropped one firm dummy for each group, and normalized the mean in each group to zero to allow for cross-group comparison. All while redefined $\tilde{D}$ accordingly.

[7] This assumption is made since without it we need too invert a $N \times N$ matrix to solve for the biases, and this is not feasible with our current computational power. Note also that unlike in the empirical estimation we have no explanatory variables when we simulate data from our theoretical model. Hence, in that case the assumption is met by definition.

random effects are uncorrelated with each other, when estimating the variances. However, there is no such restriction on the predicted individual effects. Woodcock's model is

$$y_{it} = x_{it}\beta + \theta_i + \psi_j + \phi_{ij} + \varepsilon_{it} \qquad i = 1, ..., N;\ t = 1, .., T_i;\ j = 1, ..., J$$

where ϕ_{ij} is the match effect. It is easy to estimate β even if the worker effect θ_i, the firm effect ψ_j, and the match effect ϕ_{ij} are fixed effects, since β is the within-match estimator. However, separately identifying the worker, firm, and match effects in a fixed effects context is impossible since we from, say, M matches cannot estimate $M + N + J$ effects. However, if we are willing to assume that the worker, firm and match effects are orthogonal random effects, we can estimate the model. Obviously, the orthogonality assumption is a strong assumption - especially since we are mainly interested in the correlation between the worker and firm effects.

Woodcock suggests the following 3-step estimation procedure: First, estimate $\hat{\beta}$ as the within-match estimator in the first stage and compute $\left(y_{it} - x_{it}\hat{\beta}\right)$. In the second stage, the variance of the random effects $\left(\sigma_\theta^2, \sigma_\psi^2, \sigma_\phi^2\right)$ and the error variance σ_ε^2 are estimated by Restricted Maximum Likelihood (REML) on $\left(y_{it} - x_{it}\hat{\beta}\right)$. These REML estimates are computed with the use of the average information algorithm of Gilmour, Thompson and Cullis (1995), which exploits the sparsity of the matrix design. In the third and final stage, Woodcock makes the Best Linear Unbiased Predictor (BLUP) of the random effects and estimate the correlations between the various terms.

4 Simulating of the Search Model

Since our theoretical model delivers a log-linear wage equation it seems as a natural starting point to solve and simulate data from the theoretical model and estimate the AKM model as well as Woodcock's mixed effect model on this simulated data.

To simulate our model we need a number of functional assumptions. In the following we assume that worker productivity p_w, firm productivity p_f and match productivity p_m all are log-normal distributed; $p_w \sim LN\left(\mu_w, \sigma_w^2\right)$, $p_f \sim LN\left(\mu_f, \sigma_f^2\right)$, $p_m \sim LN\left(\mu_m, \sigma_m^2\right)$ and specify the search cost function as $c\left(n\right) = c_1 n^{c_2}$.

Before we can simulate data from our model, we need to approximate the function $n\left(p_w, p_{fm}\right)$. To do this, we solve equations (5) and (6) for n using quadrature methods to approximate the integrals. The function $n\left(p_w, p_{fm}\right)$ is stepwise increasing in p_w and decreasing in p_{fm}, as illustrated in Figure 2, so we just need to know precisely in what values of $\left(p_w, p_{fm}\right)$ that n changes. We use an algorithm that determines the number of points where n changes in the $\left(p_w, p_{fm}\right)$ space with a precision such that the maximum deviation in p_{fm} and p_w from their true values is

0.01.[8] Next, we use cubic splines to approximate a curve for each change in n in the (p_w, p_{fm}) space. These cubic splines are then use to draw $n(p_w, p_{fm})$ in the actual simulation.

Notice that since the production function is given by $f(p_w, p_{fm}) = p_w^{\alpha_1} p_f^{\alpha_2} p_m^{\alpha_3}$ our assumption of log-normality in all inputs implies that the joint firm and match productivity $p_{fm} = p_f^{\frac{\alpha_2}{\alpha_2+\alpha_3}} p_m^{\frac{\alpha_3}{\alpha_2+\alpha_3}}$ is also log-normally distributed $p_{fm} \sim LN\left(\mu_{fm}, \sigma_{fm}^2\right)$ with $\mu_{fm} = \frac{\alpha_2}{\alpha_2+\alpha_3}\mu_f + \frac{\alpha_3}{\alpha_2+\alpha_3}\mu_m$ and $\sigma_{fm} = \sqrt{(\frac{\alpha_2}{\alpha_2+\alpha_3}\sigma_f)^2 + (\frac{\alpha_3}{\alpha_2+\alpha_3}\sigma_m)^2}$. This again implies that when we want to simulate from a model without a match effect, we can simply let $\alpha_3 \equiv 0$ which implies that $\mu_{fm} = \mu_f$ and $\sigma_{fm} = \sigma_f$ such that $p_{fm} = p_f$ and $f(p_w, p_{fm}) = f(p_w, p_f)$.

To facilitate comparison between simulations with and without match effect below in the simulations with match effect we always choose μ_m and σ_m such that we have $\mu_{fm} = \mu_w$ and $\sigma_{fm} = \sigma_w$, and in the simulations without the match effect we simply choose $\mu_f = \mu_w$ and $\sigma_f = \sigma_w$.

We wish to simulate an economy inhabited with firms whose size is log-normal distributed. Given μ_f and σ_f we define the minimum and maximum p_f as the values corresponding to the 0.05 percentile and the 99.5 percentile. In between the minimum and maximum values, we let each firm productivity be equally spaced, whereas each firm's share of the job offers is the log-normal density. Hence, the firm productivity is only approximately log-normal, but will as the number of firms increases, converge to the log-normal distribution. Given our initial guess of a correlation between worker and the joint firm and match effect, the model runs for 30 periods before the worker allocation is completely stable and, hence, we discard the first 29 simulated periods.

We also need to choose values for our parameters. The parameters describing flows in labor market are fixed at $\lambda = 0.9$ and $\delta = 0.1$, which implies an equilibrium unemployment of 10 per cent. We have set the parameters of the log normal distributions such that $\alpha_3 = 0.25$ implying that the match effect constitutes roughly 25 per cent of the explained variation in log wages. Obviously, the influence of the match effect that we find below will be smaller (larger) if the match effect constitutes a smaller (larger) share of the wage than the assumed 25 per cent. The rest of the parameters in the two simulations are given in Table 3.

[8] First, we set $p_{fm} = 0$ and determine for which values of p_w n changes. Then for $\underline{p}_w$, each of these points in p_w, $\bar{p}_w$, and two additional values evenly distributed among each of these points we determine all values of p_{fm} where n changes.

Table 1: Parameter values

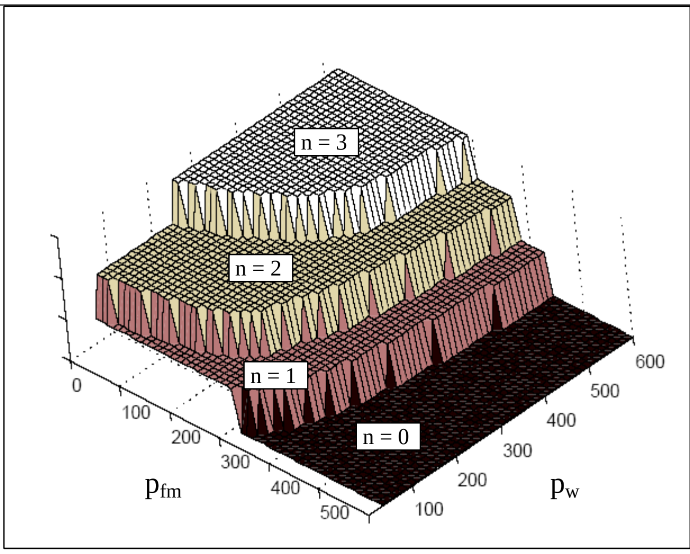
	α_1	α_2	α_3	$\mu_w = \mu_{fm}$	$\sigma_w = \sigma_{fm}$	μ_f	σ_f	μ_m	σ_m	β	c_1	c_2	r
Match Effect Included:	0.40	0.35	0.25	5.00	0.60	5.00	0.84	5.00	0.84	0.50	5.50	1.20	0.05
Match Effect Not Included:	0.40	0.60	0.00	5.00	0.60	5.00	0.60	.	.	0.50	5.50	1.20	0.05

To complete the link between the theoretical model and empirical model we add a normal distributed error term to the log linear wage equation. We set the variance of the error term, such that it contributes with 5-10 per cent of the total wage variation.

With the assumed parameter values, the maximum number of jobs that any worker samples is 3. In Figure 2 a surface plot of the n-function is shown. Unemployed workers with productivity above approximately 225, corresponding to 25 per cent of the workers, sample 3 jobs. Since almost no unemployed workers sample just one job, almost 75 per cent sample 2 jobs. The figure also shows that as workers climb the firm and match productivity ladder they reduce the number of jobs they sample and when they reach sufficiently far, they stop searching.

Since the workers in the economy without a match effect and in the economy with a match effect draw productivities from the same distribution the n-function in the two scenarios is identical.

Figure 2: The optimal job sample size



In Table 2 we show the results from the AKM estimation on the simulated data. There are only small differences in the overall labor market between the two scenarios. The differences in the dispersion of wages arise, since whereas we assume a discrete distribution of firm productivities which is cut off, the match effect is allowed to be continuous. Therefore, the joint firm and match distribution will have a larger dispersion than the firm productivity in the case without match effect. Furthermore, in the economy with a match effect workers can get a better job within a firm by drawing an offer from the exact same firm, but with a higher match effect. This might have a small positive effect on both the mean wage and the dispersion of wages.

Table 2: Monte-carlo estimations

	Match effect included			Match effect not included		
No. of replications		100			100	
No. time periods		6			6	
No. of observations		66,760			66,760	
No. of persons		12,500			12,500	
No. of firms		495			473	
Average of wages		124.7			124.4	
Variance of wages		47.1			41.2	
Ave. no. of obs per firm		135			141	
	Min	Max	Mean	Min	Max	Mean
AKM, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	-0.086	-0.035	-0.064	0.107	0.140	0.125
AKM, Corrected $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	-0.076	-0.023	-0.052	0.112	0.146	0.131
MIXED, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.014	0.042	0.029	.	.	.
True $cor(\alpha_1 * ln(p_w), \alpha_2 * ln(p_f))$	0.079	0.115	0.095	0.112	0.147	0.131
True $cor(\alpha_1 * ln(p_w), (\alpha_2 + \alpha_3) * ln(p_{fm}))$	0.093	0.135	0.112	0.112	0.147	0.131
Mean influence on $var(Y)$:						
	True	AKM	Mixed	True	AKM	
Firm effect	0.208	0.194	0.193	0.392	0.392	
Person effect	0.442	0.678	0.444	0.543	0.555	
Match effect	0.272	.	0.285	0.000	.	
Error term	0.077	0.128	0.078	0.065	0.052	

In the right panel of the table we show the results for the case without a match effect. The estimated correlation is very close to the true correlation. With the bias correction we obtain the true correlation.

In the left panel of the table we include the match effect. Here, workers change jobs on the grounds of the composite firm and match productivity rather than just the firm productivity. This implies that the true correlation between the worker and firm productivities, on average, is 0.095. However, with the presence of the match effect both the estimated and corrected correlations from the fixed effects estimation are severely biased and even negative with mean values of -0.064 and -0.052 respectively. Therefore, existing studies finding a negative correlation between worker and firm effects may actually reflect a labor market with positive assortative matching. Moreover, the bias from the match effect seems to be more important than the statistical bias. Woodcock's mixed effects model does considerably better with a positive correlation of 0.029 although it is also biased downwards.

The assumed wage setting in our theoretical model is the piece-rate contract. Cahuc, Postel-Vinay and Robin (2006) and Lentz (2008) use an alternative bargaining mechanism, in which employers are allowed to respond to an outside bid for its workers. Besides the worker's own productivity and her current firm's productivity, the worker's current wage also depends on the productivity of the previous firm. Thereby, this wage-setting mechanism would compared to the piece-rate contract ceteris paribus imply a lower correlation between the worker and firm productivities. To some extent the match effect plays a similar role here, and it is also apparent that the correlation between the worker productivity and the firm productivity is lower when there is a match component in the wages.

Woodcock (2007) argues that a match effect, which is positively correlated with the worker effect and negatively correlated with the firm effect, will imply that the AKM model overestimates the proportion of variation attributable to the worker effect and underestimates the proportion attributable to the firm effect. Our simulation results show that this is certainly true; while the true worker effect amounts 44 per cent of the variation, the estimated is 68 per cent. In this perspective Woodcock's model performs much better since it attributes exactly 44 per cent of variation to the worker effect.

Andrews, Gill, Schank and Upward (2007) argue that the negative correlation between worker and firm effects to a large extent is due to limited mobility bias. Perhaps, limited mobility bias and the bias due to match effects are the same. In other words, will the match effect bias disappear if we consider firms with more worker mobility? Due to the lognormal firm size distribution firms with the highest firm productivity advertise the fewest jobs. Thus, if we only estimate the AKM model using a 6 period sample of firms with more than 100 workers, we are essentially restricting the range of firm productivities and the estimated correlation will be biased due to the selection on an endogenous variable. Hence, for the simulated data the sample selection bias makes such exercise useless.

Instead of estimating on a sample of large firms, as we will do for the empirical estimation in the next section, we extend the sample length. Table 3 shows that more estimation periods imply that the gap between the true and the corrected estimate narrows since as the number of observations per firm increases, the firm effects become more precisely estimated. With 20 periods the estimated correlation from the AKM estimation turns positive, but even with 50 periods the bias still amounts to approximately 20 per cent. Nevertheless, we can conclude that the match effect bias is a cause of the limited mobility bias. Furthermore, Table 3 reveals that the correlation from the mixed effects estimation in all simulations lies in between the true correlation and the correlation from the AKM estimation.

Table 3: Limited mobility bias

No.of timeperiods	No. of obs.	Ave. obs. per firm	True correlation	AKM corrected	MIXED
6	66,760	135	0.095	−0.052	0.029
10	111,281	224	0.096	−0.013	0.033
20	222,523	447	0.095	0.026	0.050
30	333,746	670	0.095	0.045	0.063
50	555,296	1115	0.096	0.062	0.075

Note: The complete set of results for the simulations can be found in Table 2 and Tables 8-11 in Appendix C.

5 Empirical Results

We have access to a population register data set from Statistics Denmark and we restrict attention to all workers aged 25-59 years and employed in the private sector. The data set is an unbalanced panel data set for 1999-2004 and the variables originate from three databases and include detailed information on a wide range of variables. Due to the nature of a fixed effect estimator, which precludes the use of time-invariant explanatory variables, we only include a small number of explanatory variables. The dependent variable is the log of hourly wages taken from the IDA database. We use two groups of explanatory variables. Firstly, we include a second order polynomial in the actual labor market experience and dummies for whether the individual lives in a rural area or in a large city. These variables are also drawn from the IDA database. Secondly, firm variables including value added per employee, capital stock per employee, the ratio of males in the firm's workforce and log of number of employees drawn from the IDA and FIDA databases. Furthermore, we include time dummies in the regression. All nominal variables are deflated with the GDP deflator.

Since firm effects are only identified by movers we have chosen to restrict our estimation to firms with more than 50 worker observations in the 6-year sample window. The resulting data set has 4 million observations on $925,000$ persons employed in approximately $15,000$ firms.

Table 4: AKM parameter estimates

	AKM Estimator	MIXED Estimator
Experience	0.0187	0.0200
	(80.20)	(237.22)
(Experience)^2	-0.0062	-0.0041
	−(233.40)	−(190.91)
Large cities$^{(*)}$	-0.0022	-0.0273
	−(3.66)	−(52.22)
Rural area$^{(*)}$	0.0027	-0.0269
	(7.65)	−(83.70)
Value added per employee (mill. DKK)	0.0085	0.0189
	(33.55)	(85.74)
Capital stock per employee (thousands DKK)	0.0049	0.0000
	(1.89)	(6.42)
Male ratio	-0.0175	0.1379
	−(7.74)	(116.44)
ln(employees)	0.0194	0.0127
	(54.02)	(90.53)
R^2	0.8939	0.8874
No. of observations		4,003,929
No. of persons		925,011
No. of firms		15,455
No. of worker-firm spells		1,350,944
No. of groups		95

Note: $^{(*)}$ The omitted category is the greater Copenhagen area. t-values are in parenthesis.

Table 4 presents the parameter estimates. All coefficients are significant besides the coefficient for the capital stock per employee [9], and all have the expected signs. As our main focus is assortative matching we proceed to look at the estimated correlation between worker and firm fixed effects.

[9] In Appendix D it is shown that the coefficient to capital stock per employee is significantly positive when restricting the sample to firms with more than 300 observations, where the firm effect is better identified.

Table 5: AKM Estimator, Standard Deviations and Correlation Matrix (obs per firm >50)

	Standard deviation	Correlations								
		Y	$X_w\hat{\beta}_w$	$X_f\hat{\beta}_f$	$\hat{T}_{dum}$	$\hat{\psi}_{J(i,t)}$	$\hat{\theta}_i$	$X_f\hat{\beta}_f+$ $\hat{\psi}_{J(i,t)}$	$X_w\hat{\beta}_w+$ $\hat{\theta}_i$	$\hat{\varepsilon_{it}}$
Y	0.351	1.000	−0.049	0.044	0.010	0.281	0.886	0.294	0.905	0.326
$X_w\hat{\beta}_w$	0.068	−0.049	1.000	−0.018	−0.042	0.003	−0.256	−0.004	−0.051	0.000
$X_f\hat{\beta}_f$	0.037	0.044	−0.018	1.000	0.009	−0.165	−0.018	0.235	−0.022	0.000
$\hat{T}_{dum}$	0.017	0.010	−0.042	0.009	1.000	−0.001	−0.032	0.003	−0.042	0.000
$\hat{\psi}_{J(i,t)}$	0.009 (0.008)	0.281	0.003	−0.165	−0.001	1.000	0.036 (0.064)	0.920	0.037	0.000
$\hat{\theta}_i$	0.107 (0.102)	0.886	−0.256	−0.018	−0.032	0.036 (0.064)	1.000	0.028	0.978	0.000
$X_f\hat{\beta}_f+\hat{\psi}_{J(i,t)}$	0.094	0.294	−0.004	0.235	0.003	0.920	0.028	1.000	0.028 (0.056)	0.000
$X_w\hat{\beta}_w+\hat{\theta}_i$	0.316	0.905	−0.051	−0.022	−0.042	0.037	0.978	0.028 (0.056)	1.000	0.000
$\hat{\varepsilon_{it}}$	0.114	0.326	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Note: Lower case parenthesis denotes the correlation corrected for statistical bias. X_w includes experience, (experience)^2, large cities and rural areas. X_f includes value added per employee, capital stock per employee, male ratio and ln(employees).

The resulting correlation matrix from the AKM model is shown in the Table 5. The estimated correlation between the unobserved worker and firm effects is 0.036, while the bias corrected correlation, shown in parenthesis, is 0.064. Taking account of both observed and unobserved effects, we estimate the correlation between the overall worker effect and the overall firm effect to be 0.028, while the corrected is 0.056.[10] With the two-way fixed effects model we explain 89 per cent of the variation of the log wages, and approximately 3/4 of this variation is due to observed and unobserved worker characteristics.

The positive correlation is in contrast to existing international studies which have almost all found negative correlations. Hence, the evidence from the AKM estimation implies that there is positive assortative matching in the Danish labor market. It turns out that the positive correlation is not robust to inclusion of persons aged 18-24 years. In this case the estimated correlation is -0.0047 while the corrected correlation is 0.0155. This result suggests that the process of sorting takes time.

In addition to the statistical bias, both Abowd, Kramarz, Lengermann and Perez-Duarte (2004) and Andrews, Gill, Schank and Upward (2007) argue that there might also exist a

[10] To compute the corrected correlation we simply subtract the bias in the $cov\left(\hat{\theta}_i, \hat{\psi}_{J(i,t)}\right)$ from $cov\left(X_w\hat{\beta}_w+\hat{\theta}_i, X_f\hat{\beta}_f+\hat{\psi}_{J(i,t)}\right)$, and subtract the biases in $var\left(\hat{\theta}_i\right)$ and $var\left(\hat{\psi}_{J(i,t)}\right)$ from respectively $var\left(X_w\hat{\beta}_w+\hat{\theta}_i\right)$ and $var\left(X_f\hat{\beta}_f+\hat{\psi}_{J(i,t)}\right)$.

negative bias due to poorly identified firm effects, when there is limited mobility. We can quantify this bias by only estimating the AKM model on larger firms.[11] As Andrews, Gill, Schank and Upward we also find that this bias is more important than the statistical bias, and as shown in Appendix D the corrected correlation between the unobserved effects becomes 0.092 and 0.111 when we estimate on samples, where the minimum number of observations per firm are respectively 100 and 300. The correlation between the overall worker and firm effects are respectively 0.085 and 0.113.

TABLE 6: MIXED ESTIMATOR, STANDARD DEVIATIONS AND CORRELATION MATRIX (OBS PER FIRM >50)

		Correlations									
	Standard deviation	Y	$X_w\hat{\beta}_w$	$X_f\hat{\beta}_f$	$\hat{T}_{dum}$	$\hat{\psi}_{J(i,t)}$	$\hat{\theta}_i$	$\hat{\phi}_{ij}$	$X_f\hat{\beta}_f+$ $\hat{\psi}_{J(i,t)}$	$X_w\hat{\beta}_w+$ $\hat{\theta}_i$	$\hat{\varepsilon_{it}}$
Y	0.351	1.000	0.145	0.155	0.027	0.477	0.837	0.706	0.492	0.851	0.398
$X_w\hat{\beta}_w$	0.055	0.145	1.000	0.066	−0.039	−0.009	−0.019	−0.011	0.008	0.228	0.001
$X_f\hat{\beta}_f$	0.037	0.155	0.066	1.000	−0.011	0.069	0.025	0.001	0.320	0.041	−0.007
$\hat{T}_{dum}$	0.012	0.027	−0.039	−0.011	1.000	0.003	−0.002	0.001	0.000	−0.012	0.000
$\hat{\psi}_{J(i,t)}$	0.137	0.477	−0.009	0.069	0.003	1.000	0.119	0.038	0.967	0.114	0.004
$\hat{\theta}_i$	0.217	0.837	−0.019	0.025	−0.002	0.119	1.000	0.690	0.119	0.969	0.114
$\hat{\phi}_{ij}$	0.072	0.706	−0.011	0.001	0.001	0.038	0.690	1.000	0.036	0.670	0.216
$X_f\hat{\beta}_f+\hat{\psi}_{J(i,t)}$	0.144	0.492	0.008	0.320	0.000	0.967	0.119	0.036	1.000	0.118	0.002
$X_w\hat{\beta}_w+\hat{\theta}_i$	0.223	0.851	0.228	0.041	−0.012	0.114	0.969	0.670	0.118	1.000	0.111
$\hat{\varepsilon_{it}}$	0.099	0.398	0.001	−0.007	0.000	0.004	0.114	0.216	0.002	0.111	1.000

Note: Lower case parenthesis denotes the correlation corrected for statistical bias. X_w includes experience, (experience)^2, large cities and rural areas. X_f includes value added per employee, capital stock per employee, male ratio and ln(employees).

Table 6 presents the correlation matrix when estimating using Woodcock's hybrid mixed effects model. We obtain a correlation of 0.119 on the sample where the minimum number of observations per firm is 50. In Table 7 we compute the mean influence of each term on the variance of the log of the wage rate. While the AKM estimation suggests that roughly 80 per cent of the variance in the wages can be attributed to the individual heterogeneity, the mixed effects estimation only suggests about 50 per cent are due to worker heterogeneity. This repeats the finding from the simulated data and is due to the high positive correlation of 0.69 between the match effect and the worker effect. From the mixed effects estimation we find that the match effect explains 14.5 per cent of the variation in the log wages, which is of the same magnitude as Woodcock finds using US data.

[11] Andrews, Gill, Schank and Upward examine this by estimating on a sample of plants with 30 or more workers moving during the sample period.

Table 7: Mean Influence On $var(Y)$		
	AKM Estimator	MIXED Estimator
$X_w\hat{\beta}_w$	(0.0097)	0.0226
$X_f\hat{\beta}_f$	0.0049	0.0161
$\hat{T}_{dum}$	0.0008	0.0009
$\hat{\psi}_{J(i,t)}$	0.0745	0.1854
$\hat{\theta}_i$	0.8235	0.5173
$\hat{\phi}_{ij}$	.	0.1450
$\hat{\varepsilon_{it}}$	0.1061	0.1126

The correlation from Woodcock's mixed effects estimation is higher than the estimated correlation from the AKM model. Still our simulation results from section 4 combined with the sizeable match effect found using the Woodcock estimator show that the true correlation in the Danish labor market will be above the estimated 11.9 per cent.

6 Conclusion

In this paper we argue that thepresence of a match effect in the wages will imply that the estimated correlation in an AKM estimation will be negatively biased. In order to analyze this we develop a theoretical search model with continuous heterogeneity on both worker and firm sides. Like the empirical model, the theoretical model implies a log-linear wage equation which is additively separable in the worker effects and the firm effects. Besides this, the theoretical model also provides the opportunity of having an additively separable match effect. Importantly for our agenda, the model implies assortative matching even though we only use a strictly supermodular production function. We achieve this by letting the workers choose how many jobs they want to sample. Compared to Shimer and Smith (2000) we can as Lentz (2008) relax the assumption of log supermodularity and re-instate Becker's frictionless result.

Our model shares similarities with Lentz (2008), but has the desirable feature that more productive workers, on average, leave unemployment or any given employment level to find better matches than their less productive colleagues. However, compared to Lentz' model this is at the expense of no differences in unemployment durations. A model that could combine the differences in expected unemployment duration and that more productive workers leave unemployment to find more productive matches than less productive workers would be most appealing. It seems to be the case that such a model can be achieved by allowing not only workers to choose the sample, but also firms to choose the optimal sample of job candidates.

Thus, for future research on assortative matching the discrete time sampling approach seems to be more relevant than its continuous approximation.

In the case of Denmark the correlation between the worker and firm effects is estimated to be 0.12. We argue that it is most likely still a downward biased estimate, since from simulations of the search model we show that we can obtain the standard negative correlation estimate from an economy where there is indeed positive assortative matching.

In conclusion, although researches have found negative correlations from estimating the AKM model for a series of countries, labor markets might after all be characterized by positive assortative matching.

References

[1] Abowd, J. M., Creecy, R. & Kramarz, F. (2002), Computing person and firm effects using linked longitudinal employer-employee data, Technical Paper 2002-06, U.S. Census Bureau.

[2] Abowd, J. M., F. Kramarz and D.N. Margolis (1999): High Wage Workers and High Wage Firms, *Econometrica* **67-2**, pp. 251-333.Abowd, J. M., F. Kramarz and D.N. Margolis (1999): High Wage Workers and High Wage Firms, *Econometrica* **67-2**, pp. 251-333.

[3] Abowd, J. M., F. Kramarz, P. Lengermann and S. Perez-Duarte (2004): Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Matching Model for France and the United States, Mimeo, Cornell University.

[4] Acemoglu, D. and R. Shimer (2000): Wage and Technology Dispersion, *Review of Economic Studies* **67**, pp. 585-607.

[5] Andrews, M., L. Gill, T. Schank and R. Upward (2007): High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?, *Journal of the Royal Statistical Society (forthcoming).*

[6] Atakan, A.E. (2006): Assortative Matching with Explicit Search Costs, *Econometrica* **74-3**, pp. 667-680.

[7] Bagger J., F. Fontaine, F. Postel-Vinay and J.-M. Robin (2007): A Tractable Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation, working paper.

[8] Barth, E. and H. Dale-Olsen (2003): Assortative Matching in the Labour Market? Stylised facts about workers and plants, working paper.

[9] Becker, G.S. (1973): A Theory of Marriage: Part I, *Journal of Political Economy* **81**, pp. 813-846.

[10] Burdett, K. and M.G. Coles (1999): Long Term Partnership Formation: Marriage and Employment, *Economic Journal* **109**, pp. F307-F334.

[11] Cahuc, P. F. Postel-Vinay and J.-M. Robin (2006): Wage Bargaining with On-the-Job Search: Theory and Evidence, *Econometrica* **74-2**, pp. 323-364.

[12] Chade, H. (2001): Two-Sided Search and Perfect Segregation with Fixed Search Costs, *Mathematical Social Sciences* **42**, pp. 31-51.

[13] Gilmour, A. R., R. Thompson, and B. R. Cullis (1995): Average information REML: An Efficient Algorithm for Variance Parameter Estimation in Linear Mixed Models, *Biometrics* **51**, pp. 1440–1450.

[14] Gruetter, M. and R. Lalive (2004): The Importance of Firms in Wage Determination, IZA Discussion Paper No. 1367.

[15] Jovanovic, B (1979): Job Matching and the Theory of Turnover, *Journal of Political Economy* **87-5**, pp. 972-990.

[16] Lentz, R. (2008): An Equilibrium Model of Wage Dispersion with Sorting, working paper, University of Wisconsin-Madison.

[17] Mortensen, D.T. (2003): *Wage Dispersion: Why are Similar Workers Paid Differently?*, Zeuthen Lectures Bookseries, MIT Press.

[18] Piekkola, H. (2005): Knowledge Capital as the Source of Growth, working paper.

[19] Postel-Vinay, F. and J.-M. Robin (2002): Equilibrium Wage Dispersion with Worker and Employer Heterogeneity, *Econometrica* **70-6**, pp. 2295-2350.

[20] Sattinger, M. (1975): Comparative Advantages and the Distribution of Earnings and Abilities, *Econometrica* **43-3**, pp. 455-468.

[21] Shimer, R. (2005): The Assignment of Workers to Jobs in an Economy with Coordination Frictions, *Journal of Political Economy* **113-5**, pp. 996-1025.

[22] Shimer, R. and L. Smith (2000): Assortative Matching and Search, *Econometrica* **68-2**, pp. 343-369.

[23] Smith, L. (2006): The Marriage Model with Search Frictions, *Journal of Political Economy* **114-6**, pp. 1124-1144.

[24] Stigler, G.J. (1961): The Economics of Information, *Journal of Political Economy* **69-3**, pp. 213-225.

[25] Woodcock, S.D. (2007): Match Effects, working paper, Simon Fraser University.

Appendix

Appendix A: Proof of Proposition 2

Proof. The number of jobs sampled is necessarily discrete and recognizing this, it is useful for us to work with segments of p_{fm} for which a given worker with productivity p_w does not change her number of jobs sampled $n(p_w, p_{fm})$. Defining $\hat{g}\left(p_w, p_{fm}^j\right) \equiv \int_{p_{fm}^{j-1}}^{p_{fm}^j} \tilde{g}\left(p_w, p_{fm}'\right) dp_{fm}'$ and $p_{fm}^0 \equiv p_{fm}^r$ we can express equation (10) evaluated in p_{fm}^i as

$$\sum_{j=1}^{i} \hat{g}\left(p_w, p_{fm}^j\right) = \frac{\delta}{\delta+\lambda} \Gamma\left(p_{fm}^i\right)^{n\left(p_w, p_{fm}^0\right)} + \frac{\lambda}{\delta+\lambda} \sum_{j=1}^{i} \Gamma\left(p_{fm}^i\right)^{n\left(p_w, p_{fm}^{j-1}\right)} \hat{g}\left(p_w, p_{fm}^j\right) \tag{15}$$

Rearranging to solve for $\hat{g}\left(p_w, p_f^i\right)$ we get

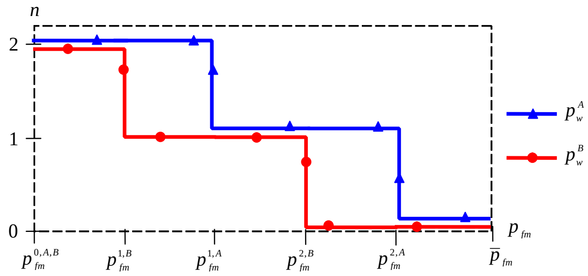
$$\hat{g}\left(p_w, p_{fm}^i\right) = \frac{\frac{\delta}{\delta+\lambda}\Gamma\left(p_{fm}^i\right)^{n\left(p_w, p_{fm}^0\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^i\right)^{n\left(p_w, p_{fm}^{i-1}\right)}} - \{i>1\} \sum_{j=1}^{i-1} \frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^i\right)^{n\left(p_w, p_{fm}^{j-1}\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^i\right)^{n\left(p_w, p_{fm}^{i-1}\right)}} \hat{g}\left(p_w, p_{fm}^j\right) \tag{16}$$

Next, for some workers of productivity p_w^X we have that

$$\hat{G}^X\left(p_{fm}^I\right) = \sum_{i=1}^{I} \frac{\frac{\delta}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)^{n\left(p_w^X, p_{fm}^{0,X}\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)^{n\left(p_w^X, p_{fm}^{I-1,X}\right)}} - \{I>1\} \sum_{i=2}^{I}\sum_{j=1}^{i-1} \left(\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{i,X}\right)^{n\left(p_w^X, p_{fm}^{j-1,X}\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{i,X}\right)^{n\left(p_w^X, p_{fm}^{i-1,X}\right)}} \cdot \hat{g}\left(p_w^X, p_{fm}^{j,X}\right) \right) \tag{17}$$

Consider two workers A and B, where $p_w^A > p_w^B$, so that we by proposition 1 have that $n\left(p_w^A, p_{fm}\right) \geq n\left(p_w^B, p_{fm}\right)$. Furthermore, assume that $n = 0, 1, 2$. Using Definition 1, we wish to compare the allocation of the two worker types to show that $G^B\left(p_{fm}^I\right) \geq G^A\left(p_{fm}^I\right)$. The proof is in two parts. In part A we show that for the particular example depicted in Figure 3 $G^B\left(p_{fm}^I\right) \geq G^A\left(p_{fm}^I\right)$ for any $p_{fm}^I \in [0,1]$. In part B of the proof, we show that all other possible scenarios are special cases of the example in part A.

Figure 3



Part A: The discreteness of the number of jobs sampled implies that n in the interval $[0,2]$ is a decreasing step-wise function of p_{fm} with five segments. As illustrated in Figure 3, the considered example implies that p_w^A and p_w^B are sufficiently close so that workers A and B are sampling the same number of jobs for some ranges of p_{fm}. Next, we will use that if $n\left(p_w^A, p_{fm}\right) = n\left(p_w^B, p_{fm}\right)$ it must be the case that person B will decrease her n at some firm productivity p_{fm} lower than the firm productivity where A will decrease her n.

By the use of equation (16) and (17) we have that $\hat{G}^B\left(p_{fm}^I\right) - \hat{G}^A\left(p_{fm}^I\right) \geq 0$ no matter which segment p_{fm}^I is situated on in Figure 3. The solutions for each segment are:

$$\left(\hat{G}^B\left(p_{fm}^I\right) - \hat{G}^A\left(p_{fm}^I\right)\right)|_{0 \leq p_{fm}^I < p_{fm}^{1,B}} = 0 \tag{18}$$

$$\left(\hat{G}^B\left(p_{fm}^I\right) - \hat{G}^A\left(p_{fm}^I\right)\right)|_{p_{fm}^{1,B} \leq p_{fm}^I < p_{fm}^{1,A}} =$$
$$\frac{\lambda\delta}{(\delta+\lambda)^2}\Gamma\left(p_{fm}^I\right)\left(\frac{\Gamma\left(p_{fm}^I\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)^2} - \frac{\Gamma\left(p_{fm}^{1,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,B}\right)^2}\right)\frac{1-\Gamma\left(p_{fm}^I\right)}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)} > 0 \tag{19}$$

$$\left(\hat{G}^B\left(p_{fm}^I\right) - \hat{G}^A\left(p_{fm}^I\right)\right)_{p_{fm}^{1,A} \leq p_{fm}^I < p_{fm}^{2,B}} =$$
$$\frac{\lambda\delta}{(\delta+\lambda)^2}\Gamma\left(p_{fm}^I\right)\left(\frac{\Gamma\left(p_{fm}^{1,A}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,A}\right)^2} - \frac{\left(p_{fm}^{1,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,B}\right)^2}\right)\frac{1-\Gamma\left(p_{fm}^I\right)}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)} > 0 \tag{20}$$

$$\left(\hat{G}^{B}\left(p_{fm}^{I}\right)-\hat{G}^{A}\left(p_{fm}^{I}\right)\right)|_{p_{fm}^{2,B}\leq p_{fm}^{I}<p_{fm}^{2,A}} =$$

$$\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p_{fm}^{I}\right)}{1-\frac{\lambda}{\delta+\lambda}}\frac{1}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,A}\right)^2}\left(\frac{\Gamma\left(p_{fm}^{I}\right)^2-\Gamma\left(p_{fm}^{1,A}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{I}\right)}-\frac{\Gamma\left(p_{fm}^{2,B}\right)^2-\Gamma\left(p_{fm}^{1,A}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)}\right)$$

$$+\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p_{fm}^{I}\right)}{1-\frac{\lambda}{\delta+\lambda}}\left(\frac{\Gamma\left(p_{fm}^{1,A}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,A}\right)^2}-\frac{\Gamma\left(p_{fm}^{1,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,B}\right)^2}\right)\left(\frac{1-\Gamma\left(p_{fm}^{I}\right)^2}{1-\Gamma\left(p_{fm}^{I}\right)}-\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)}\right)$$

$$>0 \tag{21}$$

$$\left(\hat{G}^{B}\left(p_{fm}^{I}\right)-\hat{G}^{A}\left(p_{fm}^{I}\right)\right)|_{p_{fm}^{2,A}\leq p_{fm}^{I}<1} =$$

$$\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p_{fm}^{I}\right)}{1-\frac{\lambda}{\delta+\lambda}}\cdot\frac{1}{1-\frac{\lambda}{\delta+\lambda}\left(p_{fm}^{1,A}\right)^2}\left(\frac{\Gamma\left(p_{fm}^{2,A}\right)^2-\Gamma\left(p_{fm}^{1,A}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,A}\right)}-\frac{\Gamma\left(p_{fm}^{2,B}\right)^2-\Gamma\left(p_{fm}^{1,A}\right)^2}{\left(1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)\right)}\right)$$

$$+\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p_{fm}^{I}\right)}{1-\frac{\lambda}{\delta+\lambda}}\left(\frac{\Gamma\left(p_{fm}^{2,A}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,A}\right)^2}-\frac{\Gamma\left(p_{fm}^{1,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,B}\right)^2}\right)\left(\frac{1-\Gamma\left(p_{fm}^{I}\right)^2}{1-\Gamma\left(p_{fm}^{I}\right)}-\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)}\right)$$

$$>0 \tag{22}$$

where the only term not immediately seen to be positive is

$$\left(\frac{1-\Gamma\left(p_{fm}^{I}\right)^2}{1-\Gamma\left(p_{fm}^{I}\right)}-\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{2,B}\right)}\right)$$

But since $\frac{\partial\frac{1-x^2}{1-x}}{\partial x}=1>0$ this term as well is clearly positive. Hence, we conclude that

$$\hat{G}^{B}\left(p_{fm}^{I}\right)\geq\hat{G}^{A}\left(p_{fm}^{I}\right)\quad\forall p_{fm}^{I}\in[0,1]$$

.

Part B: Since we know that

$$p_w^A>p_w^B,\ n'_{p_w}\left(p_w,p_{fm}\right)>0\quad\Longrightarrow\quad n\left(p_w^A,p_{fm}\right)\geq n\left(p_w^B,p_{fm}\right) \tag{23}$$

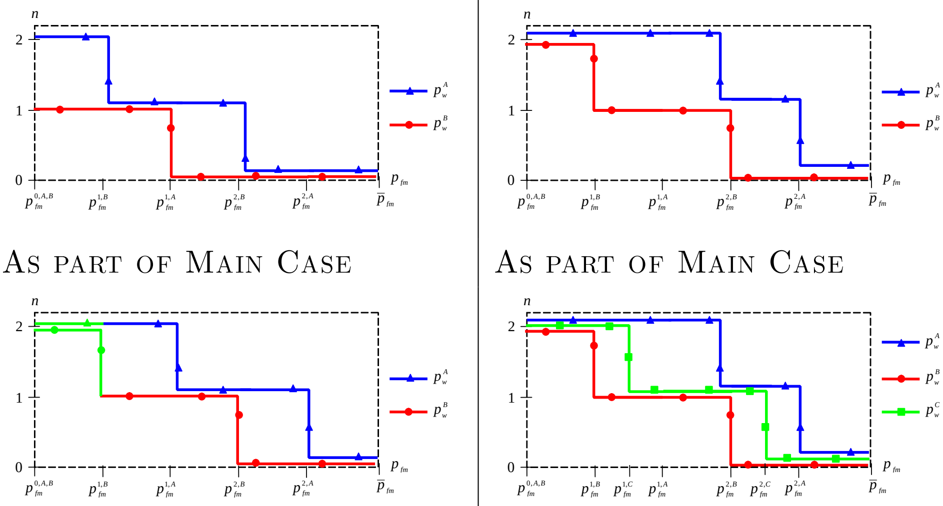
it is straightforward to show that every possible scenario satisfying (23) can be included either directly as part of the main case above or be re-stated as combinations of a number of sub-cases, all part of the main case above.

To illustrate consider the two examples in Figure 4.

FIGURE 4

CASE I) | CASE II)

AS PART OF MAIN CASE | AS PART OF MAIN CASE



Case I) just parallels the last four segments of the main case from part A since from (18) $\hat{G}^{B}\left(p_{fm}^{I}\right) = \hat{G}^{A}\left(p_{fm}^{I}\right)$ for $0 \leq p_{fm}^{I} < p_{fm}^{1}$. In case II) one can include a worker C with $p_{w}^{A} > p_{w}^{C} > p_{w}^{B}$ (as illustrated with the middle line). By the proof in part A we know that $\hat{G}^{B}\left(p_{fm}^{I}\right) \geq \hat{G}^{C}\left(p_{fm}^{I}\right)$ and $\hat{G}^{C}\left(p_{fm}^{I}\right) \geq \hat{G}^{A}\left(p_{fm}^{I}\right)$, whereby we also have that $\hat{G}^{B}\left(p_{fm}^{I}\right) \geq \hat{G}^{A}\left(p_{fm}^{I}\right)$. By the same line of argument, all cases in the $n \in [0, 1, 2]$ space can be showed to have assortative matching by using only part A of this proof. ■

Appendix B: Proof of Proposition 3

Proof. Existence of a unique reservation productivity p_{w}^{r} and a unique unemployment rate u follows trivially from equations (4) and (7). The l.h.s. of equations (5) and (6) is increasing in n, while the r.h.s. is decreasing in n, which implies that there is a unique solution for each of the equations. If this value is not an integer, it is possible that the two integers adjacent to the solution of first-order condition imply the same value of the value function. In this case, we assume that the worker chooses the lowest n. Under this assumption there exist a unique $n\left(p_{w}, b\right)$ and $n\left(p_{w}, p_{fm}\right)$ for any $p_{w} \in \left[\underline{p}_{w}, \bar{p}_{w}\right]$ and any $p_{fm} \in \left[\underline{p}_{f}, \bar{p}_{fm}\right]$. Differentiating (10) with respect to p_{fm} gives

$$\begin{aligned}
\hat{g}(p_w, p_{fm}) &= \frac{\delta}{\delta+\lambda} n(p_w, b)\, \Gamma(p_{fm})^{n(p_w,b)-1} + \frac{\lambda}{\delta+\lambda} \Gamma(p_{fm})^{n(p_w,p_{fm})}\, \hat{g}(p_w, p_{fm}) \\
&\quad + \frac{\lambda}{\delta+\lambda} \int_{\underline{p}_{fm}}^{p_{fm}} n\left(p_w, p'_{fm}\right) \Gamma(p_{fm})^{n\left(p_w,p'_{fm}\right)-1}\, \hat{g}\left(p_w, p'_{fm}\right) dp'_{fm} \\
&\Leftrightarrow \\
\hat{g}(p_w, p_{fm}) &= \frac{\frac{\delta}{\delta+\lambda} n(p_w, b)\, \Gamma(p_{fm})^{n(p_w,b)-1}}{1 - \frac{\lambda}{\delta+\lambda} \Gamma(p_{fm})^{n(p_w,p_{fm})}} + \frac{\frac{\lambda}{\delta+\lambda} \int_{\underline{p}_{fm}}^{p_{fm}} n\left(p_w, p'_{fm}\right) \Gamma(p_{fm})^{n\left(p_w,p'_{fm}\right)-1}\, \hat{g}\left(p_w, p'_{fm}\right) dp'_{fm}}{1 - \frac{\lambda}{\delta+\lambda} \Gamma(p_{fm})^{n(p_w,p_{fm})}}
\end{aligned}$$

which is an inhomogenous Volterra equation of second kind with an everywhere continuous and uniformly bounded integral kernel. Given uniqueness of $n(p_w, b)$ and $n(p_w, p_{fm})$ we also have a unique solution for $\hat{g}(p_w, p_{fm})$ for any $p_w \in \left[\underline{p}_w, \bar{p}_w\right]$ and any $p_{fm} \in \left[\underline{p}_{fm}, \bar{p}_{fm}\right]$. $\blacksquare$

Appendix C: Tables from Simulation

Table 8: Monte-carlo estimations with 10 periods

	Match effect included			Match effect not included		
No. of replications		100			100	
No. time periods		10			10	
No. of observations		111,281			111,281	
No. of persons		12,500			12,500	
No. of firms		496			477	
Average of wages		124.7			124.4	
Variance of wages		47.2			41.3	
Ave. no. of obs per firm		224			233	
	Min	Max	Mean	Min	Max	Mean
AKM, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	-0.035	0.004	-0.018	0.114	0.145	0.130
AKM, Corrected $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	-0.031	0.010	-0.013	0.116	0.147	0.132
MIXED, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.019	0.049	0.033	.	.	.
True $cor(\alpha_1 * ln(p_w), \alpha_2 * ln(p_f))$	0.082	0.110	0.096	0.116	0.146	0.132
True $cor(\alpha_1 * ln(p_w), (\alpha_2 + \alpha_3) * ln(p_{fm}))$	0.097	0.128	0.113	0.116	0.146	0.132
Mean influence on $var(Y)$:						
	True	AKM	Mixed	True	AKM	
Firm effect	0.208	0.189	0.191	0.392	0.392	
Person effect	0.443	0.642	0.448	0.543	0.551	
Match effect	0.272	.	0.284	0.000	.	
Error term	0.078	0.169	0.078	0.065	0.057	

Table 9: Monte-carlo estimations with 20 periods

	Match effect included			Match effect not included		
No. of replications		100			100	
No. time periods		20			20	
No. of observations		222,523			222,523	
No. of persons		12,500			12,500	
No. of firms		498			481	
Average of wages		124.7			124.4	
Variance of wages		47.2			41.4	
Ave. no. of obs per firm		447			462	
	Min	Max	Mean	Min	Max	Mean
AKM, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.007	0.038	0.024	0.118	0.143	0.131
AKM, Corrected $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.009	0.040	0.026	0.118	0.144	0.131
MIXED, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.035	0.061	0.050	.	.	.
True $cor(\alpha_1 * ln(p_w), \alpha_2 * ln(p_f))$	0.080	0.109	0.095	0.118	0.144	0.131
True $cor(\alpha_1 * ln(p_w), (\alpha_2 + \alpha_3) * ln(p_{fm}))$	0.099	0.124	0.113	0.118	0.144	0.131
Mean influence on $var(Y)$:						
	True	AKM	Mixed	True	AKM	
Firm effect	0.208	0.181	0.187	0.392	0.392	
Person effect	0.442	0.592	0.453	0.544	0.547	
Match effect	0.272	.	0.282	0.000	.	
Error term	0.078	0.227	0.078	0.065	0.061	

Table 10: Monte-carlo estimations with 30 periods

	Match effect included			Match effect not included		
No. of replications		100			100	
No. time periods		30			30	
No. of observations		333,746			333,746	
No. of persons		12,500			12,500	
No. of firms		498			484	
Average of wages		124.8			124.4	
Variance of wages		47.3			41.4	
Ave. no. of obs per firm		670			689	
	Min	Max	Mean	Min	Max	Mean
AKM, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.032	0.060	0.043	0.120	0.143	0.131
AKM, Corrected $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.034	0.062	0.045	0.120	0.143	0.132
MIXED, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.052	0.076	0.063	.	.	.
True $cor(\alpha_1 * ln(p_w), \alpha_2 * ln(p_f))$	0.084	0.108	0.095	0.119	0.143	0.132
True $cor(\alpha_1 * ln(p_w), (\alpha_2 + \alpha_3) * ln(p_{fm}))$	0.101	0.124	0.112	0.119	0.143	0.132
Mean influence on $var(Y)$:						
	True	AKM	Mixed	True	AKM	
Firm effect	0.208	0.177	0.185	0.391	0.391	
Person effect	0.443	0.564	0.458	0.544	0.546	
Match effect	0.271	.	0.279	0.000	.	
Error term	0.078	0.259	0.078	0.065	0.063	

Table 11: Monte-carlo estimations with 50 periods

	Match effect included			Match effect not included		
No. of replications		100			100	
No. time periods		50			50	
No. of observations		556,296			556,296	
No. of persons		12,500			12,500	
No. of firms		499			487	
Average of wages		124.8			124.5	
Variance of wages		47.3			41.3	
Ave. no. of obs per firm		1115			1142	
	Min	Max	Mean	Min	Max	Mean
AKM, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.051	0.070	0.061	0.124	0.140	0.132
AKM, Corrected $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.052	0.070	0.062	0.124	0.140	0.132
MIXED, Estimated $cor(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.066	0.084	0.075	.	.	.
True $cor(\alpha_1 * ln(p_w), \alpha_2 * ln(p_f))$	0.086	0.104	0.096	0.124	0.141	0.132
True $cor(\alpha_1 * ln(p_w), (\alpha_2 + \alpha_3) * ln(p_{fm}))$	0.102	0.123	0.113	0.124	0.141	0.132
Mean influence on $var(Y)$:						
	True	AKM	Mixed	True	AKM	
Firm effect	0.208	0.173	0.183	0.392	0.392	
Person effect	0.442	0.534	0.462	0.543	0.545	
Match effect	0.272	.	0.276	0.000	.	
Error term	0.078	0.294	0.078	0.065	0.063	

Appendix D: Tables from Empirical Estimation

TABLE 12: AKM PARAMETER ESTIMATES

	(obs per firm >100)	(obs per firm >300)
Experience	0.0168	0.0137
	(67.34)	(47.76)
$(\text{Experience})^2$	-0.0061	-0.0060
	$-(219.63)$	$-(191.50)$
Large cities$^{(*)}$	-0.0018	-0.0022
	$-(2.92)$	$-(3.18)$
Rural area$^{(*)}$	0.0024	0.0014
	(6.41)	(3.25)
Value added per employee (mill. DKK)	0.0082	0.0067
	(31.23)	(19.12)
Capital stock per employee (thousands DKK)	0.0047	0.4968
	(1.82)	(18.63)
Male ratio	-0.0190	-0.0260
	$-(6.87)$	$-(6.89)$
ln(employees)	0.0182	0.0211
	(47.64)	(45.37)
R^2	0.8986	0.9045
No. of observations	3,458,066	2,596,705
No. of persons	820,883	645,302
No. of firms	7,585	2,293
No. of groups	11	1

Note: $^{(*)}$ The omitted category is the greater Copenhagen area.

Table 13: AKM Estimator, Standard Deviations and Correlation Matrix (obs per firm >100)

	Standard deviation	Correlations: Y	$X_w\hat{\beta}_w$	$X_f\hat{\beta}_f$	$\hat{T}_{dum}$	$\hat{\psi}_{J(i,t)}$	$\hat{\theta}_i$	$X_f\hat{\beta}_f+$ $\hat{\psi}_{J(i,t)}$	$X_w\hat{\beta}_w+$ $\hat{\theta}_i$	$\hat{\varepsilon_{it}}$
Y	0.351	1.000	−0.067	0.020	0.009	0.286	0.891	0.296	0.916	0.318
$X_w\hat{\beta}_w$	0.079	−0.067	1.000	−0.009	−0.048	−0.001	−0.303	−0.004	−0.070	0.000
$X_f\hat{\beta}_f$	0.033	0.020	−0.009	1.000	0.013	−0.210	−0.024	0.183	−0.027	0.000
$\hat{T}_{dum}$	0.020	0.009	−0.048	0.013	1.000	0.001	−0.038	0.007	−0.052	0.000
$\hat{\psi}_{J(i,t)}$	0.007 (0.007)	0.286	−0.001	−0.210	0.001	1.000	0.071 (0.092)	0.923	0.075	0.000
$\hat{\theta}_i$	0.110 (0.106)	0.891	−0.303	−0.024	−0.038	0.071 (0.092)	1.000	0.063	0.972	0.000
$X_f\hat{\beta}_f + \hat{\psi}_{J(i,t)}$	0.083	0.296	−0.004	0.183	0.007	0.923	0.063	1.000	0.064 (0.085)	0.000
$X_w\hat{\beta}_w + \hat{\theta}_i$	0.317	0.916	−0.070	−0.027	−0.052	0.075	0.972	0.064 (0.085)	1.000	0.000
$\hat{\varepsilon_{it}}$	0.112	0.318	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Note: Lower case paranthesis denotes the correlation corrected for statistical bias. X_w includes experience, (experience)^2, large cities and rural areas. X_f includes value added per employee, capital stock per employee, male ratio and ln(employees).

Table 14: AKM Estimator, Standard Deviations and Correlation Matrix (obs per firm >300)

	Standard deviation	Correlations: Y	$X_w\hat{\beta}_w$	$X_f\hat{\beta}_f$	$\hat{T}_{dum}$	$\hat{\psi}_{J(i,t)}$	$\hat{\theta}_i$	$X_f\hat{\beta}_f+$ $\hat{\psi}_{J(i,t)}$	$X_w\hat{\beta}_w+$ $\hat{\theta}_i$	$\hat{\varepsilon_{it}}$
Y	0.350	1.000	−0.081	−0.005	0.007	0.282	0.887	0.299	0.926	0.309
$X_w\hat{\beta}_w$	0.099	−0.081	1.000	0.000	−0.051	−0.008	−0.367	−0.009	−0.083	0.000
$X_f\hat{\beta}_f$	0.033	−0.005	0.000	1.000	0.016	−0.356	−0.023	0.079	−0.025	0.000
$\hat{T}_{dum}$	0.022	0.007	−0.051	0.016	1.000	0.004	−0.048	0.011	−0.067	0.000
$\hat{\psi}_{J(i,t)}$	0.006 (0.006)	0.282	−0.008	−0.356	0.004	1.000	0.099 (0.111)	0.904	0.104	0.000
$\hat{\theta}_i$	0.116 (0.112)	0.887	−0.367	−0.023	−0.048	0.099 (0.111)	1.000	0.095	0.957	0.000
$X_f\hat{\beta}_f + \hat{\psi}_{J(i,t)}$	0.073	0.299	−0.009	0.079	0.011	0.904	0.095	1.000	0.099 (0.113)	0.000
$X_w\hat{\beta}_w + \hat{\theta}_i$	0.318	0.926	−0.083	−0.025	−0.067	0.104	0.957	0.099 (0.113)	1.000	0.000
$\hat{\varepsilon_{it}}$	0.108	0.309	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Note: Lower case paranthesis denotes the correlation corrected for statistical bias. X_w includes experience, (experience)^2, large cities and rural areas. X_f includes value added per employee, capital stock per employee, male ratio and ln(employees).

Two-Sided Sampling and Sorting

Daniel le Maire*

University of Copenhagen and CAM

December 12, 2008

Abstract

This paper presents a search model of marriage with heterogenous agents. All agents are allowed to sample two potential partners in a one-shot game. In this search envirronment the production function needs only to be supermodular for assortative matching to arise. Hence, we are able to re-instate Becker's (1973) result in a search context without relying on agents' having either constant costs as in Atakan (2006) or different search efforts as in Lentz (2008) and le Maire and Scheuer (2008). Furthermore, in contrast to the latter two models we achieve symmetric assortative matching.

1 Introduction

This paper considers two-sided matching between heterogeneous agents. When two agents meet and match, they generate an output, which can be shared among the two. Such matching situation arises in many contexts although the search and matching literature most often has considered matching between workers and firms in the labor market or matching between men and women in the marriage market. In this paper we consider matching in the marriage market.

When considering heterogeneous agents we are especially interested in when it is the case that high productivity agents match with other high productivity agents, which is termed positive assortative matching. In a seminal contribution Becker (1973) argued that when there is complementarity in the production function we can expect positive assortative matching. In the frictionless model considered by Becker, the sufficient mathematical property of the production function is supermodularity, which implies that the cross derivative of the production function is positive.

Shimer and Smith (2000) show that for assortative matching to arise in a search model setting it is no longer sufficient that the production function is strictly supermodular. Instead,

*I thank Karsten Albæk, Dale Mortensen and Louise Rathsach Skouby for very helpful comments. All mistakes are mine.

Shimer and Smith find that a stronger version of complementarity is needed, that is log supermodularity. As noted by Atakan (2006) the problem is that the higher gains to more search for more productive workers are offset by higher costs of rejecting an offer. Using this, Atakan shows that if unmatched agents have constant flow costs independent of their type and if the discount rate is zero, the production function only needs to be strictly supermodular in order to imply assortative matching.

Recently, Lentz (2008) and le Maire and Scheuer (2008) have shown that when workers can choose their search effort the sufficient condition for assortative matching is just strictly supermodularity of the production function. Lentz casts his model in continuous time and in his model more productive workers' higher search intensity implies that they have a higher job arrival rate. Hence, after some time productive workers will, on average, have changed jobs more times and, therefore, be in more productive firms. Thus, more productive workers do not leave unemployment to a more productive first job and without on-the-job search there is not assortative matching with a strictly supermodular production function.

Le Maire and Scheuer set their model in discrete time and allow workers to optimally choose a sample of either $0, 1$, or 2 jobs. When more productive workers sample more jobs, they will not match more frequently, but instead each time select among a larger sample of firms and, thereby, on average select a more productive firm. Therefore, in the model of le Maire and Scheuer assortative matching does not break-down in the absence of on-the-job search.

In this paper we want to abstract from the ingredients that re-instate assortative matching with a strictly supermodular production function in a search context, that is, constant costs of search and endogenous search effort together with on-the-job search are assumed away. Instead, we allow both sides to sample partners simultaneously in a one-shot matching game.

When only workers are able to use sampling and firms just fill their vacancy randomly, it is optimal for workers to apply to the most productive firm. When both sides use sampling it will, in general, not be optimal for low productivity agents to select the best alternative. To take account of this, we use a framework as in Gale and Shapley (1962) and allow agents to apply for all sampled partners and use a downward-recursively matching strategy, that is, first try to match with the best alternative, and if rejected try to match with the second best alternative, etc.

The paper is organized as follows. In the next section the theoretical model is outlined, and in section 3 we conclude.

2 Theoretical Model

Consider a market with N women and N men. All agents are rational, forward-looking, risk-neutral and infinitely lived. Agents look for a partner of the opposite sex to match with. We allow agents to sample $1 \leq S \leq N$ potential partners in a one-shot matching game. The S potential partners are randomly selected among all agents on the other side of the market. To keep things simple we assume that the matching behavior is controlled by a marriage agency. This agency makes sure that each agent sample exactly S agents from the other side.

When a man and a woman meet, they can generate a non-negative output according to the production function $f(p_m, p_f)$ where $p_m \in \left[\underline{p}_m, \bar{p}_m\right]$ is the productivity of the man and $p_f \in \left[\underline{p}_f, \bar{p}_f\right]$ is the productivity of the woman. Agents on both sides of the market are heterogenous. The distribution of men's productivity is given by $H(p_m)$ while for women it is given by $Q(p_f)$. We assume that the utility or production is transferable and can be shared among the two, such that the man gets the share β of the production, while the woman get the remainder. All possible matches have positive pay-off and, hence, all types of match have a probability strictly greater than zero.

When $f'_{p_f}(p_m, p_f) > 0$ men prefer to match with the woman of highest p_f among their S potential partners, and similarly when $f'_{p_m}(p_m, p_f) > 0$ women prefer to match with the most productive man among their S potential partners. Both conditions are satisfied under the following assumption

Assumption 1 *The production function $f(p_m, p_f)$ is non-negative, increasing in its inputs $f'_{p_m}(p_m, p_f) > 0$, $f'_{p_f}(p_m, p_f) > 0$, and is supermodular, that is, if $p_m^1 > p_m^0$ and $p_f^1 > p_f^0$ we have that $f\left(p_m^1, p_f^1\right) + f\left(p_m^0, p_f^0\right) \geq f\left(p_m^1, p_f^0\right) + f\left(p_m^0, p_f^1\right)$.*

The fact that matching with more productive women yields a higher output, that is $f'_{p_f}(p_m, p_f) > 0$, suggests a possibility of different matches not carrying the same probability. When a man receives his S draws from the marriage agency, he sorts the draws according to productivity and downward-recursively he tries to match with each of the women until he is accepted. For $S = N$ this is the "deferred-acceptance" procedure of Gale and Shapley (1962) with the additional assumption that all men rank women in the same way and all women rank men in the same way. To keep things simple, but still deriving new results, we assume that $S = 2$ such that the marriage agency allows each person to consider exactly two potential partners from the other side. If agents on the one side all randomly sample two partners on the other side, not all agents on the other side would be sampled twice. To overcome this problem we assume that there exists a marriage agency, which randomly allocates two potential partners from the other side of the market to each agent.

Let $\lambda_1\left(p_m, p_f^1\right)$ denote the probability that a man with productivity p_m will match with the preferred woman among his two draws, and let $\lambda_0\left(p_m, p_f^0\right)$ denote the probability that the man matches with the other draw. The probability that the best of the two women has a productivity below p_f^1 is given by $Q\left(p_f^1\right)^2$. Conditional on the best of the draws is p_f^1, the probability that the other draw is below some $p_f^0 < p_f^1$ is $q\left(p_f^0\right) / Q\left(p_f^1\right)$.

If we let the value of being unmatched be denoted by $U_i(\cdot)$ where $i \in \{m, f\}$. The Bellman equation for an unmatched man of productivity p_m is

$$rU_m(p_m) = \int_{\underline{p}_f}^{\bar{p}_f} \left(\begin{array}{c} \lambda_1\left(p_m, p_f^1\right)\beta f\left(p_m, p_f^1\right) \\ +\frac{1-\lambda_1\left(p_m,p_f^1\right)}{Q\left(p_f^1\right)} \int_{\underline{p}_f}^{p_f^1} \lambda_0\left(p_m, p_f^0\right)\beta f\left(p_m, p_f^0\right) dQ\left(p_f^0\right) \end{array} \right) d\left(Q\left(p_f^1\right)^2\right) \tag{1}$$

where r is the discount factor. The corresponding value functions for a woman of productivity p_f are

$$rU_f(p_f) = \int_{\underline{p}_m}^{\bar{p}_m} \left\{ \begin{array}{c} \eta_1\left(p_f, p_m^1\right)(1-\beta) f\left(p_m^1, p_f\right) \\ +\frac{1-\eta_1\left(p_f,p_m^1\right)}{H\left(p_m^1\right)} \int_{\underline{p}_m}^{p_m^1} \eta_0\left(p_f, p_m^0\right)(1-\beta) f\left(p_m^1, p_f\right) Q\left(dp_m^0\right) \end{array} \right\} d\left(H\left(p_m^1\right)^2\right) \tag{2}$$

where $\eta_j\left(p_f, p_m^j\right)$ for $j \in \{0,1\}$ is the probability that a woman with productivity p_f matches with a man of productivity p_m^j.

Consider a man with quality p_m^1 sampling two women with productivities $p_f^1 > p_f^0$. Naturally, he will prefer to match with the woman with the highest quality p_f^1. If the other woman with productivity p_f^0 has drawn a less productive man $p_m^0 < p_m^1$, she will wait until it is known whether the man with productivity p_m^1 actually matches with p_f^1 or not. If p_m^1 and p_f^1 do not match, then, it must be the case that p_f^1 has a better offer $p_m^2 > p_m^1$, and p_m^2 and p_f^1 will match. Since the man with productivity p_m^1 cannot match with his preferred draw, he will, instead match p_f^0.

We can assign probabilities to such matching process. With probability $H\left(p_m^1\right)$ the preferred woman with productivity p_f^1 accepts p_m^1 and the two match. However, with probability $1 - H\left(p_m^1\right)$ the preferred woman samples a better man and with probability $Q\left(p_f^1\right)$ the woman and the other man match. It can also be the case that the woman of productivity p_f^1 does not get the preferred man with $p_m^2 > p_m^1$. For a moment let us stop the matching process here and calculate the probability that the man of productivity p_m^1 matches with his preferred woman of productivity p_f^1. This probability is $H\left(p_m^1\right) + \left(1 - H\left(p_m^1\right)\right)\left(1 - Q\left(p_f^1\right)\right)$. However, this calculation is only true conditional on the man with quality p_m^2 matches with his preferred woman. The probability of that event is $H\left(p_m^2\right) + \left(1 - H\left(p_m^2\right)\right)\left(1 - Q\left(p_f^2\right)\right)$ given that if the woman with productivity p_f^2 draws a higher $p_m^3 > p_m^2$ he gets his preferred woman of productivity $p_f^3 > p_f^2$. Hence, we have that the probability that a man of productivity p_m^1 matches with his best alternative is

$$\begin{aligned}
\lambda_1\left(p_m^1, p_f^1\right) &= E\left[\left. \begin{array}{c} H\left(p_m^1\right) + \left(1 - H\left(p_m^1\right)\right)\left(1 - Q\left(p_f^1\right)\right) \\ \times \left\{ \begin{array}{c} H\left(p_m^2\right) + \\ \left(1 - H\left(p_m^2\right)\right)\left(1 - Q\left(p_f^2\right)\right) \left\{ \begin{array}{c} H\left(p_m^3\right) + \\ \left(1 - H\left(p_m^3\right)\right)\left(1 - Q\left(p_f^3\right)\right) \left\{ \begin{array}{c} and \\ so\ on \end{array} \right\} \end{array} \right\} \end{array} \right\} \end{array} \right| p_m^1, p_f^1 \right] \\
&= E\left[\left. \prod_{i=1}^{K-1} \left(1 - Q\left(p_f^i\right)\right)\left(1 - H\left(p_m^i\right)\right) + \sum_{i=1}^{K} H\left(p_m^i\right) \prod_{j=1}^{i-1} \left(1 - Q\left(p_f^j\right)\right)\left(1 - H\left(p_m^j\right)\right) \right| p_m^1, p_f^1 \right] \\
&= E\left[\left. \sum_{i=1}^{K} H\left(p_m^i\right) \prod_{j=1}^{i-1} \left(1 - Q\left(p_f^j\right)\right)\left(1 - H\left(p_m^j\right)\right) \right| p_m^1, p_f^1 \right] \qquad (3)
\end{aligned}$$

where K is the maximum number of rounds it takes before all $2N$ agents are assigned, and where we can abstract from the first term $\prod_{i=1}^{K-1}\left(1-Q\left(p_f^i\right)\right)\left(1-H\left(p_m^i\right)\right)$, since it goes to zero as K becomes large. We will approximate $\lambda_1\left(p_f^1, p_m^1\right)$ in equation (3) by abstracting from terms for $K \geq 5$. The expressions for $i = 1, 2, 3$ can be solved exact, while we will approximate the term for $i = 4$. All these derivations are relegated to Appendix A.[1] Dropping the superscripts such that $p_m = p_m^1$ and $p_f = p_f^1$ we have that

$$\begin{aligned}\lambda_1\left(p_m, p_f\right) &= H\left(p_m\right) + \frac{1}{2}\left(1 - Q\left(p_f\right)\right)\left(1 - H\left(p_m\right)^2\right) \\ &\quad + \frac{1}{12}\left(1 - Q\left(p_f\right)\right)^2\left(2 - 3H\left(p_m\right) + H\left(p_m\right)^3\right) \\ &\quad + \frac{1}{288}\left(1 - Q\left(p_f\right)\right)^3\left(1 - H\left(p_m\right)\right)^3\left(\frac{37}{6} + H\left(p_m\right)\right) \qquad (4)\end{aligned}$$

The following proposition characterize the basic properties of $\lambda_1\left(p_m, p_f\right)$.

Proposition 1 *(i) The higher quality of the man, the larger is the probability of matching with the best of the two drawn women, that is $\partial\lambda_1\left(p_m, p_f\right)/\partial p_m > 0$. (ii) The better the quality of the (preferred) drawn woman the smaller the chance is of matching, that is $\partial\lambda_1\left(p_m, p_f\right)/\partial p_f < 0$. (iii) The cross-derivative of $\lambda_1\left(p_m, p_f\right)$ is positive.*

Proof. See Appendix B-D. ■

Conditional on not matching with the best alternative the probability of matching with the other alternative is identical to the probability of matching with the best alternative, that is $\lambda_1\left(p_m, p_f\right) = \lambda_0\left(p_m, p_f\right)$ and the results in Proposition 1 also hold for $\lambda_0\left(p_m, p_f\right)$.

The overall share of unmatched agents u is

$$u = 1 - \int_{\underline{p}_m}^{\bar{p}_m}\int_{\underline{p}_f}^{\bar{p}_f}\left(\lambda_1\left(\tilde{p}_m, \tilde{p}_f\right) + \left(1 - \lambda_1\left(\tilde{p}_m, \tilde{p}_f\right)\right)\frac{1}{Q\left(\tilde{p}_f\right)}\int_{\underline{p}_f}^{\tilde{p}_f}\lambda_0\left(\tilde{p}_m, \tilde{p}_f^0\right) dQ\left(\tilde{p}_f^0\right)\right) d\left(Q\left(\tilde{p}_f\right)^2\right) dH\left(\tilde{p}_m\right) \quad (5)$$

Proposition 2 *The share of matched men conditional on type is increasing in p_m and, similarly, the share of matched women conditional on type is increasing in p_f.*

Proof. See Appendix E. ■

We want to consider the allocation of agents for this one-shot game. Let $G\left(p_m, p_f\right)$ denote the mass of men with quality p_m or less who match with women of quality p_f or less. When $S = 2$ there are two types of inflow to $G\left(p_m, p_f\right)$; a man with productivity below p_m matches with the preferred partner, who has a productivity below p_f or a man with productivity below p_m is not accepted with the preferred partner, but is accepted by the least preferred partner

[1] It is easy to assess the quality of the approximation since it is independent of the shape of $H\left(p_m\right)$ and $Q\left(p_f\right)$. The maximum approximation errror for $\lambda_1\left(p_m^1, p_f^1\right)$ is when $p_m^1 = p_f^1 = 0$. In this case the approximation error is approximately 0.001.

with a productivity below p_f. Hence, we have

$$(1-u)\,G\,(p_m,p_f) =$$
$$\int_{\underline{p}_m}^{p_m}\left(\int_{\underline{p}_f}^{p_f}\lambda_1\,(\tilde{p}_m,\tilde{p}_f)\,d\left(Q\,(\tilde{p}_f)^2\right)+\int_{\underline{p}_f}^{p_f}\lambda_0\,(\tilde{p}_m,\tilde{p}_f^0)\int_{\tilde{p}_f^0}^{\bar{p}_f}(1-\lambda_1\,(\tilde{p}_m,\tilde{p}_f))\,\frac{1}{Q\,(\tilde{p}_f)}d\left(Q\,(\tilde{p}_f)^2\right)dQ\,(\tilde{p}_f^0)\right)dH\,(\tilde{p}_m) \tag{6}$$

Differentiating equation (6) with respect to p_m and subsequently p_f gives us

$$\begin{aligned}\hat{g}\,(p_m,p_f) &= \frac{2}{1-u}\left(\lambda_1\,(p_m,p_f)\,Q\,(p_f)+\lambda_0\,(p_m,p_f)\int_{p_f}^{\bar{p}_f}(1-\lambda_1\,(p_m,\tilde{p}_f))\,dQ\,(\tilde{p}_f)\right)\\ &= \frac{2}{1-u}\lambda_1\,(p_m,p_f)\left(1-\int_{p_f}^{\bar{p}_f}\lambda_1\,(p_m,\tilde{p}_f)\,dQ\,(\tilde{p}_f)\right)\end{aligned} \tag{7}$$

where we use the normalized density $\hat{g}\,(p_m,p_f)\equiv g\,(p_m,p_f)/\left(h\,(p_m)\,q\,(p_f)\right)$ in order to be able to examine the allocation of the game. Due to property (i) the first part of equation (7) is increasing in p_m whereas the second part is decreasing in p_m. This implies that $\hat{g}\,(p_m,p_f)$ conditional on p_f has a global maximum in p_m.

For examination of the allocation of agents we need the following definition of assortative matching also used in Lentz (2008) and le Maire and Scheuer (2008).

Definition 1 *Consider two women A and B, where $p_f^A > p_f^B$. Assortative matching implies that $G^B\,(p_m)\equiv\frac{1-u}{1-u\left(p_f^B\right)}\int_{\underline{p}_m}^{p_m}\hat{g}\left(\tilde{p}_m,p_f^B\right)dH\,(\tilde{p}_m)\geq\frac{1-u}{1-u\left(p_f^A\right)}\int_{\underline{p}_m}^{p_m}\hat{g}\left(\tilde{p}_m,p_f^A\right)dH\,(\tilde{p}_m)\equiv G^B\,(p_m)$ for $p_m\in\left]\underline{p}_m;\bar{p}_m\right[$.*

Proposition 3 *The model features assortative matching.*

Proof. Differentiating $\hat{g}\,(p_m,p_f)$ gives

$$\frac{\partial\hat{g}\,(p_m,p_f)}{\partial p_m}=\frac{2}{1-u}\left\{\begin{array}{c}\frac{\partial\lambda_1(p_m,p_f)}{\partial p_m}\left(1-\int_{p_f}^{\bar{p}_f}\lambda_1\,(p_m,\tilde{p}_f)\,dQ\,(\tilde{p}_f)\right)\\ -\lambda_1\,(p_m,p_f)\int_{p_f}^{\bar{p}_f}\frac{\partial\lambda_1(p_m,\tilde{p}_f)}{\partial p_m}dQ\,(\tilde{p}_f)\end{array}\right\} \tag{8}$$

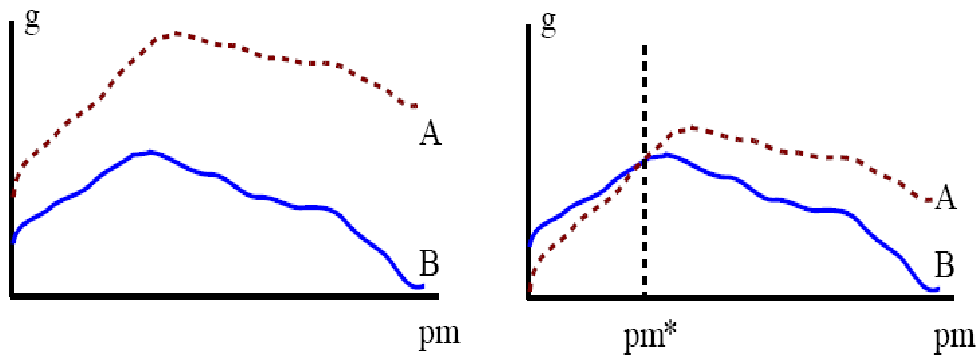
Due to property (i) the first part of equation (8) is always positive, while the second part is always negative. It is clear that for p_f sufficiently close to $\bar{p}_f$ we must have that $\frac{\partial\hat{g}\left(p_m,p_f\right)}{\partial p_m}>0$ for all $p_m<\bar{p}_m$, since the negative part approaches 0. Next, consider decreasing p_f from $\bar{p}_f$. The first term of the positive part decreases when p_f decreases due to property (iii). Furthermore, the second term of the positive part also decreases as p_f decreases. At the same time, decreasing p_f implies that the negative part in absolute terms increases. Hence, decreasing p_f decreases $\frac{\partial\hat{g}\left(p_m,p_f\right)}{\partial p_m}$. These results correspond to the cross-derivative being positive, that is

$$\frac{\partial^2\hat{g}\,(p_m,p_f)}{\partial p_m\partial p_f}=\frac{2}{1-u}\left\{\begin{array}{c}\frac{\partial^2\lambda_1\left(p_m,p_f\right)}{\partial p_m\partial p_f}\left(1-\int_{p_f}^{\bar{p}_f}\lambda_1\,(p_m,\tilde{p}_f)\,dQ\,(\tilde{p}_f)\right)+\frac{\partial\lambda_1\left(p_m,p_f\right)}{\partial p_m}\lambda_1\,(p_m,p_f)\,q\,(p_f)\\ -\frac{\partial\lambda_1\left(p_m,p_f\right)}{\partial p_f}\int_{p_f}^{\bar{p}_f}\frac{\partial\lambda_1\left(p_m,\tilde{p}_f\right)}{\partial p_m}dQ\,(\tilde{p}_f)+\lambda_1\,(p_m,p_f)\,\frac{\partial\lambda_1\left(p_m,p_f\right)}{\partial p_m}q\,(p_f)\end{array}\right\}$$

which is positive since all terms are positive due to property (i),(ii) and (iii). Hence, if decreasing p_f makes $\frac{\partial \hat{g}(p_m,p_f)}{\partial p_m}$ negative, it will not become positive again. We cannot without further examination know whether $\frac{\partial \hat{g}(p_m,p_f)}{\partial p_m}$ becomes negative, but we can see that when p_m comes sufficiently close to $\bar{p}_m$ and p_f is sufficiently close to $\underline{p}_f$ we must have that $\frac{\partial \hat{g}(p_m,p_f)}{\partial p_m} < 0$, since the positive part approaches zero. From equation (7) we can see that $\hat{g}\left(\bar{p}_m, p_f^A\right) > \hat{g}\left(\bar{p}_m, p_f^B\right)$, since $\lambda_1\left(\bar{p}_m, p_f\right) = 1$, so that increasing p_f just increases the second term of equation (7).

By proposition 2 we have that $\frac{\partial(1-u(p_f))}{\partial p_f} > 0$, which implies that $\int_{\underline{p}_m}^{\bar{p}_m} \hat{g}\left(\tilde{p}_m, p_f^A\right) d\tilde{p}_m > \int_{\underline{p}_m}^{\bar{p}_m} \hat{g}\left(\tilde{p}_m, p_f^B\right) d\tilde{p}_m$. We cannot know whether $\hat{g}\left(\underline{p}_m, p_f^A\right) \geq \hat{g}\left(\underline{p}_m, p_f^B\right)$ or $\hat{g}\left(\underline{p}_m, p_f^A\right) < \hat{g}\left(\underline{p}_m, p_f^B\right)$ and we need to consider each case separately. Below, the two cases are illustrated.

Case 1: $\hat{g}\left(\underline{p}_m, p_f^A\right) \geq \hat{g}\left(\underline{p}_m, p_f^B\right)$ Case 2: $\hat{g}\left(\underline{p}_m, p_f^A\right) < \hat{g}\left(\underline{p}_m, p_f^B\right)$



1. Consider the case where $\hat{g}\left(\underline{p}_m, p_f^A\right) \geq \hat{g}\left(\underline{p}_m, p_f^B\right)$. Since $\frac{\partial \hat{g}(p_m,p_f^A)}{\partial p_m} > \frac{\partial \hat{g}(p_m,p_f^B)}{\partial p_m}$ we must have that $\hat{g}\left(p_m, p_f^A\right) > \hat{g}\left(p_m, p_f^B\right)$ for all $p_m \in \left]\underline{p}_m, \bar{p}_m\right]$. Furthermore, $\frac{\partial \hat{g}(p_m,p_f^A)}{\partial p_m} > \frac{\partial \hat{g}(p_m,p_f^B)}{\partial p_m}$ implies that the vertical distance between $\hat{g}\left(p_m, p_f^A\right)$ and $\hat{g}\left(p_m, p_f^B\right)$ will be increasing in p_m. Hence, a larger share of the type A women will be matched to high type men, such that $\frac{1-u}{1-u(p_f^A)} \int_{\underline{p}_m}^{p_m} \hat{g}\left(\tilde{p}_m, p_f^A\right) dH\left(\tilde{p}_m\right) < \frac{1-u}{1-u(p_f^B)} \int_{\underline{p}_m}^{p_m} \hat{g}\left(\tilde{p}_m, p_f^B\right) dH\left(\tilde{p}_m\right)$, that is the allocation features assortative matching.

2. Consider the case where $\hat{g}\left(\underline{p}_m, p_f^A\right) < \hat{g}\left(\underline{p}_m, p_f^B\right)$ such that in a range $\left[\underline{p}_m, p_m^*\right]$ $\hat{g}\left(p_m, p_f^A\right) < \hat{g}\left(p_m, p_f^B\right)$ and in a range $[p_m^*, \bar{p}_m]$ $\hat{g}\left(p_m, p_f^A\right) > \hat{g}\left(p_m, p_f^B\right)$. Since more type A women are matched than type B women, we must have that

$$\int_{p_m^*}^{\bar{p}_m} \left[\hat{g}\left(\tilde{p}_m, p_f^A\right) - \hat{g}\left(\tilde{p}_m, p_f^B\right)\right] dH\left(\tilde{p}_m\right) > \int_{\underline{p}_m}^{p_m^*} \left[\hat{g}\left(\tilde{p}_m, p_f^B\right) - \hat{g}\left(\tilde{p}_m, p_f^A\right)\right] dH\left(\tilde{p}_m\right)$$

Such allocation features assortative matching, since it implies that $\frac{1-u}{1-u(p_f^A)} \int_{\underline{p}_m}^{p_m} \hat{g}\left(\tilde{p}_m, p_f^A\right) dH\left(\tilde{p}_m\right) < \frac{1-u}{1-u(p_f^B)} \int_{\underline{p}_m}^{p_m} \hat{g}\left(\tilde{p}_m, p_f^B\right) dH\left(\tilde{p}_m\right)$.

■

The source of assortative matching in Lentz (2008) and le Maire and Scheuer (2008) is the endogenous search effort, while here the only source of assortative matching is the possibility for both sides of the market to consider more than one candidate at a time.

In Lentz' model assortative matching arises, since more productive workers have a higher probability of matching since they search more. In the model of le Maire and Scheuer, assortative matching arises since more productive agents have more alternatives to choose from, also since they search more. In Lentz' model more productive agents are less likely to be unmatched, but the agents are not more likely to match with more productive partners, since assortative matching only arises since good agents, from search when already being matched, change partner more frequently. In contrast to this, in le Maire and Scheuer's model the share of unmatched agents is independent of the productivity, but high productive workers are more likely to match with high productive employers. Hence, in the latter model on-the-job search is not needed for assortative matching to arise.

The present model combines the attractive features from Lentz and le Maire and Scheuer in that a) the share of unmatched agents is higher for low productive agents and b) more productive agents are more likely to match with more productive agents on the other side of the market, so that on-the-job search is not needed. Furthermore, in the present framework assortative matching also arises, when the production function is only supermodular compared to strictly supermodular. This implies that unlike in the models by Lentz and le Maire and Scheuer assortative matching is the result with a production function of the type $f(p_m, p_f) = p_m + p_f$.

In both the models of Lentz and le Maire and Scheuer, the type of assortative matching is asymmetric, since when assortative matching arises due to differences in search effort on the worker side, it can only be proven that more productive workers on average will match with more productive firms. The opposite that more productive firms on average match with more productive workers cannot be proven. Interestingly, the two-sided sampling in this model implies that the assortative matching is symmetric, which we summarize in the following proposition.

Proposition 4 *Assortative matching is symmetric.*

Proof. By symmetry since $\eta_i(p_f, p_m)$ for $i = \{0, 1\}$ satisfies property (i)-(iii). ■

The symmetry of the equilibrium allocation can also be illustrated by considering a man with productivity p_m^A and a woman with productivity p_f^A, where $H(p_m^A) = Q(p_f^A)$ and two other points of the distribution $p_m^B > p_m^A$ and $p_f^B > p_m^B$, where $H(p_m^B) = Q(p_f^B)$. Then we must have that the man with productivity p_m^A has the same probability of matching as the woman of quality p_f^A with the preferred partner, that is $\lambda_1(p_m^A, p_f^B) = \eta_1(p_f^A, p_m^B)$. Furthermore, for the equilibrium allocation we obtain the symmtric result that $g(p_m^A, p_f^B) / (h(p_m^A) q(p_m^B)) = g(p_m^B, p_f^A) / (h(p_m^B) q(p_m^A))$.

Since the sample size S is exogenous to the agents, the allocation is not influenced by the shape of the distribution functions for the unmatched agents $H(p_m)$ and $Q(p_f)$. Hence, if we double the productivity of women in the highest decile, the matching will not change. Clearly, this would not have been the case if we allowed men to choose their sample size. Doubling the productivity of women in the highest decile would imply higher returns to search and, thereby, a larger optimal search sample.

Earlier studies of assortative matching relied on examining the sets of matches that are acceptable by both workers and firms, whereas for $S = 2$ assortative matching only happens on average. The frictionless case with perfect segregation along the 45 degree line is obtained when $S = N$, but what happens when S increases towards N? To answer this question consider $S = N - 1$ for $N \geq 3$. Next, consider the most productive man and the second most productive man. If the most productive man samples all but the most productive woman, and the second most productive man samples the $N - 1$ most productive women, the most productive woman and the second most productive man will match and the second most productive woman and the most productive man will match. Hence, for $S = N - 1$ we obtain a segregated market.

In conclusion, there is not assortative matching when $S = 1$, but for $S > 1$ there will be assortative matching on average, although as S gets sufficiently close to N the allocation will be segregated.

3 Conclusion

This paper has presented a new way of achieving assortative matching in a search context even though the production function is only supermodular. Whereas Lentz (2008) and le Maire and Scheuer (2008) relied on endogenous search effort, we show that this is not needed. We just need that both sides of the market sample more than one potential partner at a time.

The continuous time approximation has shown itself very useful in the search literature, but this approximation is less innocent when two-sided heterogeneity is introduced in equilibrium search models. In fact, it is crucial for the properties of the equilibrium allocation of agents whether the agents can consider more than one candidate at a time.

In typical real life matching situations agents are able to consider more than one potential partner. For example, a firm aiming at filling a vacancy often invites a small number of candidates for job interviews. Empirically, it seems very likely that the degree of assortative matching in the labor market is enhanced by on-the-search, but it is less compelling that assortative matching in the labor market arises solely as a result of search by already matched agents. Also therefore, it seems to be very relevant to develop discrete time search models with two-sided sampling.

References

[1] Atakan, A.E. (2006): Assortative Matching with Explicit Search Costs, *Econometrica* **74-3**, p. 667-680.

[2] Becker, G.S. (1973): A Theory of Marriage: Part I, *Journal of Political Economy* **81**, p. 813-846.

[3] Bohrnstedt, G.W. and A.S. Goldberger (1969): On the Exact Covariance of Products of Random Variables, *Journal of the American Statistical Journal* **64-328**, pp. 1439-1442.

[4] Chade, H. and L. Smith (2006): Simultaneous Search, *Econometrica* **74-5**, pp. 1293-1307.

[5] Gale, D. and L.S. Shapley (1962): College Admissions and the Stability of Marriage, *American Mathematical Monthly* **69-1**, pp. 9-15.

[6] Lentz, R. (2008): An Equilibrium Model of Wage Dispersion with Sorting, working paper, University of Wisconsin-Madison.

[7] le Maire, D. and C. Scheuer (2008): Job Sampling and Sorting, working paper, University of Copenhagen.

[8] Shimer, R. and L. Smith (2000): Assortative Matching and Search, *Econometrica* **68-2**, p. 343-369.

[9] Stigler, G.J. (1961): The Economics of Information, *Journal of Political Economy* **69-3**, p. 213-225.

[10] Stigler, G.J. (1962): Information in the Labor Market, *Journal of Political Economy* **70-5**, p. 94-105.

Appendix

Appendix A: Detailed derivation of $\lambda_1\left(p_m^1, p_f^1\right)$

The first term of equation (3) for $i = 1$ is obviously simply $E\left[H\left(p_m^1\right)\middle| p_m^1, p_f^1\right] = H\left(p_m^1\right)$. The term for $i = 2$ is

$$\begin{aligned} E\left[H\left(p_m^2\right)\left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)\right)\middle| p_m^1, p_f^1\right] &= \left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)\right)E\left[H\left(p_m^2\right)\middle| p_m^1, p_f^1\right] \\ &= \frac{1}{2}\left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)^2\right) \end{aligned}$$

where we have used that $E\left[H\left(p_m^2\right)\middle|\, p_m^1, p_f^1\right] = \left(1+H\left(p_m^1\right)\right)/\,2.$

To derive the the third term for $i = 3$ we use that $H\left(p_m^i\right)$and $Q\left(p_f^j\right)$ are uncorrelated and that on average it is only half of the variation in the previous $H\left(p_m^{i-1}\right)$ which affects the expected value of $H\left(p_m^i\right)$. We have

$$
\begin{aligned}
E\left[H\left(p_m^3\right)\prod_{j=1}^{2}\left(1-Q\left(p_f^j\right)\right)\left(1-H\left(p_m^j\right)\right)\middle|\, p_m^1, p_f^1\right] &= \begin{array}{c}\left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)\right)\times \\ E\left[1-Q\left(p_f^2\right)\middle|\, p_m^1, p_f^1\right]\times \\ \left\{\begin{array}{c} E\left[1-H\left(p_m^2\right)\middle|\, p_m^1, p_f^1\right]E\left[H\left(p_m^3\right)\middle|\, p_m^1, p_f^1\right] \\ -cov\left(H\left(p_m^2\right), H\left(p_m^3\right)\middle|\, p_m^1, p_f^1\right)\end{array}\right\}\end{array} \\
&= \begin{array}{c}\left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)\right)\left(1-\frac{1+Q\left(p_f^1\right)}{2}\right)\times \\ \left\{\frac{1-H\left(p_m^1\right)}{2}\frac{3+H\left(p_m^1\right)}{4}-\frac{1}{2}var\left(H\left(p_m^2\right)\middle|\, p_m^1, p_f^1\right)\right\}\end{array} \\
&= \frac{1}{12}\left(1-Q\left(p_f^1\right)\right)^2\left(2-3H\left(p_m^1\right)+H\left(p_m^1\right)^3\right)
\end{aligned}
$$

Before, we compute the fourth term we will separately consider the variance of $H\left(p_m^3\right)$ For this we need to define v which is standard uniform. By using the formula for the variance of a product (see Bohrstedt and Goldberger (1969)) we find

$$
\begin{aligned}
var\left(H\left(p_m^3\right)\right) &= var\left(\left(1-H\left(p_m^2\right)\right)v\right) \\
&= \left[E\left(1-H\left(p_m^2\right)\right)^2 var\left(v\right)+E\left(v\right)^2 var\left(1-H\left(p_m^2\right)\right)+var\left(1-H\left(p_m^2\right)\right)var\left(v\right)\right] \\
&= \left[\left(\frac{1-H\left(p_m^1\right)}{2}\right)^2\frac{1}{12}+\frac{1}{4}\frac{\left(1-H\left(p_m^2\right)\right)^2}{12}+\frac{\left(1-H\left(p_m^2\right)\right)^2}{12}\frac{1}{12}\right] \\
&= \frac{7}{144}\left(1-H\left(p_m^1\right)\right)^2
\end{aligned}
$$

The fourth term is

$$
E\left[H\left(p_m^4\right)\prod_{j=1}^{3}\left(1-Q\left(p_f^j\right)\right)\left(1-H\left(p_m^j\right)\right)\middle|\, p_m^1, p_f^1\right]
$$

$$
\begin{aligned}
&= && \begin{array}{c} \left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)\right)\times \\ E\left[\left(1-Q\left(p_f^2\right)\right)\left(1-Q\left(p_f^3\right)\right)\Big|\, p_m^1,p_f^1\right]\times E\left[\left(1-H\left(p_m^2\right)\right)\left(1-H\left(p_m^3\right)\right)H\left(p_m^4\right)\big|\, p_m^1,p_f^1\right] \end{array} \\
&= && \begin{array}{c} \left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)\right)\times \\ \left\{E\left[1-Q\left(p_f^2\right)\Big|\, p_m^1,p_f^1\right]E\left[1-Q\left(p_f^3\right)\Big|\, p_m^1,p_f^1\right]+cov\left(Q\left(p_f^2\right),Q\left(p_f^3\right)\Big|\, p_m^1,p_f^1\right)\right\}\times \\ \left\{\begin{array}{c} \left\{\begin{array}{c} E\left[\left(1-H\left(p_m^2\right)\right)\big|\, p_m^1,p_f^1\right]\times E\left[\left(1-H\left(p_m^3\right)\right)\big|\, p_m^1,p_f^1\right] \\ +cov\left(H\left(p_m^2\right),H\left(p_m^3\right)\big|\, p_m^1,p_f^1\right) \end{array}\right\}\times E\left[H\left(p_m^4\right)\big|\, p_m^1,p_f^1\right]+ \\ cov\left[\left(1-H\left(p_m^2\right)\right)\left(1-H\left(p_m^3\right)\right),H\left(p_m^4\right)\big|\, p_m^1,p_f^1\right] \end{array}\right\} \end{array} \\
&\simeq && \begin{array}{c} \left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)\right)\times\left\{\frac{1}{8}\left(1-Q\left(p_f^1\right)\right)^2+\frac{1}{24}\left(1-Q\left(p_f^1\right)\right)^2\right\}\times \\ \left\{\begin{array}{c} \left\{\frac{1}{8}\left(1-H\left(p_m^1\right)\right)^2+\frac{1}{24}\left(1-H\left(p_m^1\right)\right)^2\right\}\times\frac{1}{8}\left(7+H\left(p_m^1\right)\right) \\ -E\left[\left(1-H\left(p_m^2\right)\right)\big|\, p_m^1,p_f^1\right]cov\left(H\left(p_m^3\right),H\left(p_m^4\right)\big|\, p_m^1,p_f^1\right) \\ -E\left[\left(1-H\left(p_m^3\right)\right)\big|\, p_m^1,p_f^1\right]cov\left(H\left(p_m^2\right),H\left(p_m^4\right)\big|\, p_m^1,p_f^1\right) \end{array}\right\} \end{array} \\
&\simeq && \begin{array}{c} \frac{1}{6}\left(1-Q\left(p_f^1\right)\right)^3\left(1-H\left(p_m^1\right)\right)\times \\ \left\{\frac{1}{48}\left(1-H\left(p_m^1\right)\right)^2\left(7+H\left(p_m^1\right)\right)-\frac{7}{576}\left(1-H\left(p_m^1\right)\right)^3-\frac{1}{192}\left(1-H\left(p_m^1\right)\right)^3\right\} \end{array} \\
&\simeq && \frac{1}{288}\left(1-Q\left(p_f^1\right)\right)^3\left(1-H\left(p_m^1\right)\right)^3\left(\frac{37}{6}+H\left(p_m^1\right)\right)
\end{aligned}
$$

where we have used what Bohrstedt and Goldberger (1969, p. 1441) term the "conventional asymptotic approximation procedure" in order to compute $cov\left[\left(1-H\left(p_m^2\right)\right)\left(1-H\left(p_m^3\right)\right),H\left(p_m^4\right)\big|\, p_m^1,p_f^1\right]$.

Inserting all this gives us

$$
\begin{aligned}
\lambda_1\left(p_m^1,p_f^1\right) &= E\left[\sum_{i=1}^{4}H\left(p_m^i\right)\prod_{j=1}^{i-1}\left(1-Q\left(p_f^j\right)\right)\left(1-H\left(p_m^j\right)\right)\Bigg|\, p_m^1,p_f^1\right] \\
&= H\left(p_m^1\right)+\frac{1}{2}\left(1-Q\left(p_f^1\right)\right)\left(1-H\left(p_m^1\right)^2\right) \\
&\quad +\frac{1}{12}\left(1-Q\left(p_f^1\right)\right)^2\left(2-3H\left(p_m^1\right)+H\left(p_m^1\right)^3\right) \\
&\quad +\frac{1}{288}\left(1-Q\left(p_f^1\right)\right)^3\left(1-H\left(p_m^1\right)\right)^3\left(\frac{37}{6}+H\left(p_m^1\right)\right)
\end{aligned}
$$

Appendix B: Proof of property (i) in proposition 1

Proof. Differentiate $\lambda_1(p_m, p_f)$ from equation (3) with respect to p_m

$$\begin{aligned}
\frac{\partial \lambda_1(p_m, p_f)}{\partial p_m} &= h(p_m) - (1-Q(p_f))H(p_m)h(p_m) \\
&\quad -\frac{1}{4}(1-Q(p_f))^2\left(1-H(p_m)^2\right)h(p_m) \\
&\quad -\frac{3}{288}(1-Q(p_f))^3(1-H(p_m))^2\left(\frac{37}{6}+H(p_m)\right)h(p_m) \\
&\quad +\frac{1}{288}(1-Q(p_f))^3(1-H(p_m))^3h(p_m) \\
&= h(p_m) - (1-Q(p_f))H(p_m)h(p_m) \\
&\quad -\frac{1}{4}(1-Q(p_f))^2\left(1-H(p_m)^2\right)h(p_m) \\
&\quad -\frac{1}{288}(1-Q(p_f))^3\left[\frac{35}{2}-31H(p_m)+\frac{19}{2}H(p_m)^2+4H(p_m)^3\right]h(p_m) \\
&= h(p_m)\left\{\begin{array}{c}
1+\frac{1}{576}(1-Q(p_f))^2[-179+35Q(p_f)] \\
+\frac{1}{288}(1-Q(p_f))\left[-257-62Q(p_f)+31Q(p_f)^2\right]H(p_m) \\
+\frac{1}{576}(1-Q(p_f))^2[125+19Q(p_f)]H(p_m)^2 \\
-\frac{1}{72}(1-Q(p_f))^3H(p_m)^3
\end{array}\right\}
\end{aligned}$$

Consider the part inside the curly brackets and abstract for a moment from the final term. Then, we have a quadratic equation. It is immediately apparent that the coefficient in front of the quadratic term is positive and, hence, the minimum of the function is where

$$\begin{aligned}
H^*(p_m) &= \frac{-\frac{1}{288}(1-Q(p_f))\left[-257-62Q(p_f)+31Q(p_f)^2\right]}{\frac{2}{576}(1-Q(p_f))^2[125+19Q(p_f)]} \\
&= \frac{257+62Q(p_f)-31Q(p_f)^2}{125-106Q(p_f)-19Q(p_f)^2}
\end{aligned}$$

and it is obvious that $H^*(p_m|Q(p_f)=0)=\frac{257}{125}$ and that the minimum $H^*(p_m)>\frac{257}{125}$ for all $Q(p_f)\in[0,1]$. Hence, the minimum of the quadratic equation is in $H(p_m)=1$. Therefore, it suffices to evaluate $\partial\lambda_1(p_m,p_f)/\partial p_m$ in this point and verify that it is positive

$$\begin{aligned}
\frac{\partial \lambda_1(p_m, p_f|p_m=\bar{p}_m)}{\partial p_m} &= h(p_m)\left\{\begin{array}{c}
1+\frac{1}{576}(1-Q(p_f))^2[-179+35Q(p_f)] \\
+\frac{1}{288}(1-Q(p_f))\left[-257-62Q(p_f)+31Q(p_f)^2\right] \\
+\frac{1}{576}(1-Q(p_f))^2[125+19Q(p_f)] \\
-\frac{1}{72}(1-Q(p_f))^3
\end{array}\right\} \\
&= h(p_m)\left\{1+\frac{1}{576}\left(1-Q(p_f^1)\right)\left\{\begin{array}{c}
-54+108Q(p_f)-54Q(p_f)^2 \\
-514-124Q(p_f)+62Q(p_f)^2 \\
-8+16Q(p_f)-8Q(p_f)^2
\end{array}\right\}\right\} \\
&= h(p_m)Q(p_f)
\end{aligned}$$

which obviously is strictly positive for $0<Q(p_f)\leq 1$. ■

Appendix C: Proof of property (ii) in proposition 1

Proof. Straightforward differentiation of $\lambda\left(p_m, p_f\right)$ with respect to p_f gives

$$\begin{aligned}
\partial\lambda_1\left(p_m,p_f\right)/\partial p_f &= q\left(p_f\right)\left\{\begin{array}{c} -\frac{1}{2}\left(1-H\left(p_m\right)^2\right)-\frac{1}{6}\left(1-Q\left(p_f\right)\right)\left[2-3H\left(p_m\right)+H\left(p_m\right)^3\right] \\ -\frac{1}{96}\left(1-Q\left(p_f\right)\right)^2\left(1-H\left(p_m\right)\right)^3\left(\frac{37}{6}+H\left(p_m\right)\right)\end{array}\right\} \\
&= -q\left(p_f\right)\left\{\begin{array}{c} \frac{1}{2}Q\left(p_f\right)\left(1-H\left(p_m\right)^2\right)+\frac{1}{6}\left(1-Q\left(p_f\right)\right)\left[5-3H\left(p_m\right)-3H\left(p_m\right)^2+H\left(p_m\right)^3\right] \\ +\frac{1}{96}\left(1-Q\left(p_f\right)\right)^2\left(1-H\left(p_m\right)\right)^3\left(\frac{37}{6}+H\left(p_m\right)\right)\end{array}\right\} \\
&= -q\left(p_f\right)\left\{\begin{array}{c} \frac{1}{2}Q\left(p_f\right)\left(1-H\left(p_m\right)^2\right)+\frac{1}{6}\left(1-Q\left(p_f\right)\right)\left(1-H\left(p_m\right)\right)\left(5-H\left(p_m\right)^2+2H\left(p_m\right)\right) \\ +\frac{1}{96}\left(1-Q\left(p_f\right)\right)^2\left(1-H\left(p_m\right)\right)^3\left(\frac{37}{6}+H\left(p_m\right)\right)\end{array}\right\}
\end{aligned}$$

where it can be immediately seen that all three terms in the curly brackets are positive, such that $\partial\lambda_1\left(p_m,p_f\right)/\partial p_f<0$. ■

Appendix D: Proof of property (iii) in proposition 1

Proof.

$$\begin{aligned}
\frac{\partial^2\lambda_1\left(p_m,p_f\right)}{\partial p_m\partial p_f} &= h\left(p_m\right)q\left(p_f\right)\left\{\begin{array}{c} -\frac{2}{576}\left(1-Q\left(p_f\right)\right)\left[-179+35Q\left(p_f\right)\right] \\ +\frac{35}{576}\left(1-Q\left(p_f\right)\right)^2 \\ -\frac{1}{288}\left[-257-62Q\left(p_f\right)+31Q\left(p_f\right)^2\right]H\left(p_m\right) \\ +\frac{1}{288}\left(1-Q\left(p_f\right)\right)\left[-62+62Q\left(p_f\right)\right]H\left(p_m\right) \\ -\frac{2}{576}\left(1-Q\left(p_f\right)\right)\left[125+19Q\left(p_f\right)\right]H\left(p_m\right)^2 \\ +\frac{19}{576}\left(1-Q\left(p_f\right)\right)^2H\left(p_m\right)^2 \\ \frac{3}{72}\left(1-Q\left(p_f\right)\right)^2H\left(p_m\right)^3\end{array}\right\} \\
&= \frac{1}{576}h\left(p_m\right)q\left(p_f\right)\left\{\begin{array}{c} \left(1-Q\left(p_f\right)\right)\left[393-105Q\left(p_f\right)\right] \\ +\left[390+372Q\left(p_f\right)-186Q\left(p_f\right)^2\right]H\left(p_m\right) \\ -\left(1-Q\left(p_f\right)\right)\left[231+57Q\left(p_f\right)\right]H\left(p_m\right)^2 \\ 24\left(1-Q\left(p_f\right)\right)^2H\left(p_m\right)^3\end{array}\right\}
\end{aligned}$$

where it is clear that for $0\leq H\left(p_m\right),Q\left(p_f\right)\leq 1$ the first two terms and the fourth term are positive while third term is negative. Next consider only the first and the third terms. Together, they will have a minimum for $H\left(p_m\right)=1$, since this assigns the highest weight on the negative third term. Setting $H\left(p_m\right)=1$ and adding the two terms give $162\left(1-Q\left(p_f\right)\right)^2$, which will never be negative. Hence, we can conclude that $\partial^2\lambda_1\left(p_m,p_f\right)/\partial p_m\partial p_f>0$ for $0<H\left(p_m\right),Q\left(p_f\right)\leq 1$. ■

.1 Appendix E: Proof of increasing share of matched indviduals 2

Proof. The share of matched men conditional on type is given by

$$\left(1-u\left(p_m\right)\right)=2\int_{\underline{p}_f}^{\bar{p}_f}\left(\lambda_1\left(p_m,\tilde{p}_f\right)Q\left(\tilde{p}_f\right)+\int_{\underline{p}_f}^{\bar{p}_f}\left(1-\lambda_1\left(p_m,\tilde{p}_f\right)\right)\int_{\underline{p}_f}^{\tilde{p}_f}\lambda_0\left(p_m,\tilde{p}_f^0\right)dQ\left(\tilde{p}_f^0\right)\right)dQ\left(\tilde{p}_f\right)$$

and for women by

$$(1-u(p_f)) = 2\int_{\underline{p}_m}^{\bar{p}_m} \left(\eta_1(\tilde{p}_m, p_f) H(\tilde{p}_m) + \int_{\underline{p}_m}^{\bar{p}_m} (1-\eta_1(\tilde{p}_m, p_f)) \int_{\underline{p}_m}^{\tilde{p}_m} \lambda_0(\tilde{p}_m^0, p_f)\, dH(\tilde{p}_m^0) \right) dH(\tilde{p}_m)$$

Differentiating each in turn gives us

$$\begin{aligned}
\frac{\partial(1-u(p_m))}{\partial p_m} &= 2\int_{\underline{p}_f}^{\bar{p}_f} \begin{pmatrix} \frac{\partial \lambda_1(p_m, \tilde{p}_f)}{\partial p_m} Q(\tilde{p}_f) - \frac{\partial \lambda_1(p_m, \tilde{p}_f)}{\partial p_m} \int_{\underline{p}_f}^{\tilde{p}_f} \lambda_0\left(p_m, \tilde{p}_f^0\right) dQ\left(\tilde{p}_f^0\right) \\ + \int_{\underline{p}_f}^{\bar{p}_f} (1-\lambda_1(p_m, \tilde{p}_f)) \int_{\underline{p}_f}^{\tilde{p}_f} \frac{\partial \lambda_0\left(p_m, \tilde{p}_f^0\right)}{\partial p_m} dQ\left(\tilde{p}_f^0\right) \end{pmatrix} dQ(\tilde{p}_f) \\
&= 2\int_{\underline{p}_f}^{\bar{p}_f} \begin{pmatrix} \frac{\partial \lambda_1(p_m, \tilde{p}_f)}{\partial p_m} \left[Q(\tilde{p}_f) - \int_{\underline{p}_f}^{\tilde{p}_f} \lambda_0\left(p_m, \tilde{p}_f^0\right) dQ\left(\tilde{p}_f^0\right) \right] \\ + \int_{\underline{p}_f}^{\bar{p}_f} (1-\lambda_1(p_m, \tilde{p}_f)) \int_{\underline{p}_f}^{\tilde{p}_f} \frac{\partial \lambda_0\left(p_m, \tilde{p}_f^0\right)}{\partial p_m} dQ\left(\tilde{p}_f^0\right) \end{pmatrix} dQ(\tilde{p}_f) > 0
\end{aligned}$$

and

$$\frac{\partial(1-u(p_f))}{\partial p_f} = 2\int_{\underline{p}_m}^{\bar{p}_m} \begin{pmatrix} \frac{\partial \lambda_1(\tilde{p}_m, p_f)}{\partial p_f} \left[H(\tilde{p}_m) - \int_{\underline{p}_m}^{\tilde{p}_m} \eta_0\left(\tilde{p}_m^0, p_f\right) dH\left(\tilde{p}_m^0\right) \right] \\ + \int_{\underline{p}_m}^{\bar{p}_m} (1-\eta_1(\tilde{p}_m, p_f)) \int_{\underline{p}_m}^{\tilde{p}_m} \frac{\partial \eta_0\left(\tilde{p}_m^0, p_f\right)}{\partial p_f} dH\left(\tilde{p}_m^0\right) \end{pmatrix} dH(\tilde{p}_m) > 0$$

■