Microeconometrics replication project: Valuing Alternative Work Arrangements Mas & Pallais, (2017)

Guillem Tobías Larrauri

November 2023

1 Summary

Compensating wage differentials are hard to measure in observational data. Lavetti (2023) provides a very good summary of the main empirical challenges. Mas and Pallais (2017) use an experiment to see how workers value different work arrangements/amenities. The answer is that it depends on the arrangement, they value working from home, dislike intensely being scheduled at employer discretion and are indifferent to scheduling flexibility. The valuations are very heterogeneous across workers. There are important gender differences when it comes to working from home, women value it more.

2 Experimental design

The authors were staffing a call center for unrelated purposes and decided to run this experiment. Ex ante there was no information on the job location, schedule or duration. This was done to minimize possible sample selection issues. Each worker that applied was given a choice between a standard 9-5 job and a job with an "amenity": flexible hours, work from home, schedule set by the employer (different workers had different choices). The wage difference between these two jobs was selected randomly from a uniform distribution. Applicants were assured that the choices they made would not have any effect on the probability of them getting the job. With the observed choices for given wage differences and some econometric theory, the authors then estimated the distribution of willingness to pay for the different amenities amongst these applicants.

3 Econometric theory

We have a sample of 1,...,N individuals. People have a choice between 2 jobs, these jobs have different "schedules" and pay different wages. Δw is defined as $w_1 - w_0$ and it is distributed uniformly over a discrete set: from -5 to 5 with more concentration close to zero.

Each individual in this population has a true underlying preference over the two jobs, Y_i (define it as job preference):

$$Y_i = \begin{cases} 0 & \text{if } WTP_i < -\Delta w := w_0 - w_1, \\ 1 & \text{if } WTP_i \geq -\Delta w := w_0 - w_1. \end{cases} \tag{1}$$

That means, if individual i's Willingness To Pay (WTP) for the amenity job 1 provides is larger than the wage difference between jobs, the individual will want the job with the amenity. Then we can define

$$\forall i Pr(Y_i = 1) := Pr(WTP_i > \Delta w) = P_{\Delta w}$$

But some of our individuals (2α) are "inattentive", that means they make their choice $Bernoulli(0.5)$, while still having WTP_i or underlying preferences that are from the same distribution as the others. This is an assumption. If what causes the inattentiveness is related to their willingness to pay distribution, the reasoning falls apart. For the inattentive there is always a 50% chance that they pick something that goes against their underlying preferences. We care about job preferences, not job acceptance, which is defined as A_i, job preferences and job acceptance would be the same if all people were attentive. However, as there are inattentive people, we do not observe job preference for any individual. Correcting for the existence of these inattentive people we can make inference on the underlying preferences. If we aggregate these 2 kinds of people we can get:

$$Pr[A_i = 1|\Delta w] = P_{\Delta w}(1 - 2\alpha)[1] + (1 - P_{\Delta w})\alpha[2] + P_{\Delta w}(\alpha)[3]$$

Term 1 is the fraction of attentive people, they will only choose 1 if $WTP_i > \Delta w$. Term 2 is half the attentive people, they make the right choice on accident. Term 3 is the other half of the inattentive people who at this wage differential would pick the standard job, so they make a mistake. Rearranging:

$$Pr[A_i = 1|\Delta w] = P_{\Delta w}(1 - 2\alpha) + \alpha \tag{2}$$

Now assume that the WTP_i follows a logistic distribution.

$$F(WTP) = \frac{e^{\frac{(WTP-\mu)}{s}}}{1 + e^{\frac{(WTP-\mu)}{s}}}$$

It is important to note that this distribution is our ultimate goal. This can be rewritten as:

$$F(\Delta w) = \frac{1}{1 + e^{-(c+b\Delta w)}}$$

where c $=\mu/s$ and b=1/s. We are interested in the distribution of WTP not on the c and b coefficients per se. The final trick is to show that $Pr(Y_i = 1) := F(c+b\Delta w)$, this follows easily from the Logistic distribution. I define $X_i := (1, \Delta w)'$ and $\beta := (c, b)'$, then:

$$Pr(Y_i = 1) = Pr(X_i'\beta + WTP_i > 0) = Pr(WTP_i > -X_i'\beta) = 1 - F_{WTP}(-X_i'\beta)$$

$$= 1 - \frac{e^{-X_i'\beta}}{1 + e^{-X_i'\beta}} = \frac{1}{1 + e^{-X_i'\beta}} \equiv F(WTP)$$

[1]

For the estimation we will rely on the fact that we have iid observations and the key assumption is the fact Δw is independent from WTP. A couple comments are warranted on this independence. First, it holds. This is why we do an experiment in the first place. Then, it is important to note that we are not trying to predict/describe how this preferences are formed. You could say being a woman with children matters, how far you live from the job, your education level, the difference in pay between the jobs, etc. In this paper, everything that is not wage differences is put into the error term and called willingness to pay, so WTP can be due to being a woman with children, how far away you live, etc. And this WTP-which is traditionally the error term ϵ- will be independent of Δw by random assignment[2], then we can estimate this. This might be obvious, but it really confused me because this is a bit different from the reasoning traditionally used in discrete choice models, however the purpose of this model is also quite different from the usual motivation of discrete choice models in Train (2009). Summarizing the assumptions I found[3]:

- Inattentive and attentive have the same underlying preferences over the two jobs.

[1] This follows Wooldridge (2010).

[2] If we have a good enough randomization, in any case, correcting for covariates doesn't seem to be as straightforward as in the linear case.

[3] For now assume the α parameter is known, I will explain how they estimate it in one of my extensions

≡ Observations are iid.

- The WTP distribution is logistic.
- Δw_i is independent of WTP_i, which comes from the experiment.

With these assumptions we can easily estimate equation 2 via MLE, all standard errors are bootstraped. The procedure the authors and I follow is to first estimate the alpha and then estimate using MLE with alpha as a given. In the boostrap they reestimate the alphas, so the variability in the estimation of alpha is taken into account for the standard errors.

4 Replication results

This shows the main table from the paper and the graph for flexible scheduling. All specifications have observations ranging from 600-700. As I previewed, mean WTP for scheduling flexibility is very low. For working from home; specification 3 and 4, the mean WTP is high and statistically significant.[4] Even more striking is the mean WTP to avoid having your schedule set by your employer. Finally, there is significant heterogeneity amongst workers, this can be seen in the interquantile range, for all amenities.

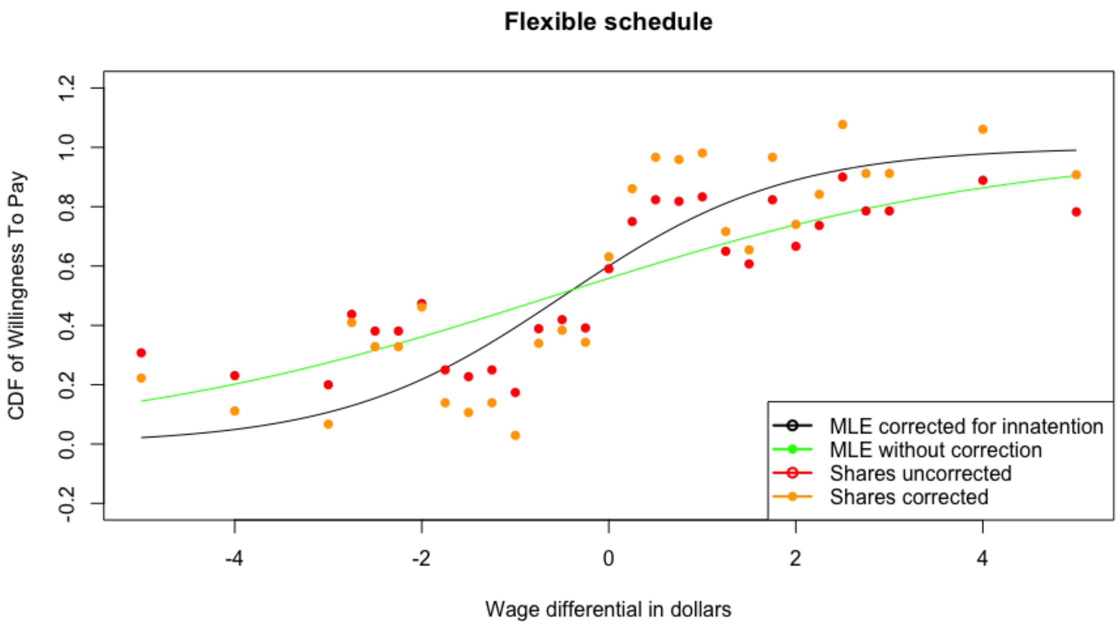
The graph shows a similar picture, the black line is the one reported in the paper, MLE corrected for inattention, it can be thought of as the logit regression fitted to the orange points; which are innatention corrected shares of acceptance of the job with the amenity for every wage difference. I also provide the green line which is MLE without correction and it can be thought to be the logit plotted to the red uncorrected dots. I plot the uncorrected to show that correcting for the inattentiveness has the effect of

[4] It could have been clearer for the authors to report confidence intervals, in my code there is an option to produce this table with intervals, however to homogeneize the analysis I report all tables with standard errors.

reducing the variance of the distribution. A larger slope in the middle will mean less mass at the tails. The tails can be a policy relevant quantity.

	Mean	SD	25th quantile	50th quantile	75th quantile
Flexible schedule	0.48	2.15	-0.82	0.48	1.79
Standard deviation	0.24	1.10	0.56	0.24	0.83
Flexible number of hours	-0.22	2.24	-1.58	-0.22	1.14
Standard deviation	0.22	0.94	0.54	0.22	0.67
Work from home	1.33	1.86	0.20	1.33	2.45
Standard deviation	0.27	0.82	0.50	0.27	0.62
Combined flexible	1.17	2.33	-0.25	1.17	2.58
Standard deviation	0.31	0.81	0.47	0.31	0.68
Employer discretion	3.41	2.95	1.63	3.41	5.20
Standard deviation	0.45	0.86	0.49	0.45	0.84

Table 1: Replication of Table 5 from the paper: Moments of WTP for different amenities.



5 Proposal 1: Better observational study

The paper showcased the difficulty of computing compensating wage differentials from observational data -CPS data from 2001 and 2004- by running some linear regressions of wages on different work arrangements and controls. On average people who worked from home earned more after the controls, while the theory by Rosen, without frictions, predicts the opposite,Lavetti (2023).

While their observational study is just illustrative, I wanted to see if I could do something better with the methods that we learned from class. **Both** analysis are based on the same identifying assumption, **selection on observables**, or conditional independence based on covariates. The double machine learning methods will not solve any potential issues with unobservables, if people who work at home have higher unobserved ability that would still be a problem. However the machine learning methods relaxes the assumptions made on functional forms. My main issue with their analysis was on the lack of flexibility from OLS. In their experiment the different amenities are disjoint, there is no option for working from home with a flexible schedule, reality is not that simple, people can have more than one amenity in the same job, therefore flexibility is paramount. When they estimated the differentials for a given amenity, the variables that contained information on other amenities were not included in the regression.

What I did was to keep the other amenities as covariates, and then estimate the "causal" effects. I used both the PLR, which we saw in class: $Y = D\theta_0 + g_0(X) + \zeta, E[\zeta|D, X] = 0, D = m_0(X) + V, E[V|X] = 0$ and the IRM, which is a similar, even more flexible model: $Y = g_0(D, X) + U, E[U|X, D] = 0, D = m_0(X) + V, E[V|X] = 0$. To implement this methodology I used the DoubleML package in R,Bach et al. (2023). I used their clean data directly[5] and I made some choices, I didn't use the weights from the CPS, I didn't augment the dataset further, there were 60 variables, and all data was

[5] I was under time constraints and thought that cleaning data was just time consuming

declared numerical, even some categorical variables.[6] This more flexible estimation[7] approach does not solve the main issues with the estimation of these differentials. People who work from home still earn more, even after controlling more flexibly. A nice result, taking into account all the issues I previously described, is that the people that work with irregular inconsistent schedules do seem to earn more on average; this is an improvement from their baseline estimation.

	Estimate	Std. Error	t value	Pr(>|t|)
Flexible hours	36.36	8.00	4.55	0.00
Any work from home	98.39	33.04	2.98	0.00
Formal work from home	11.78	30.68	0.38	0.70
Any irregular	37.74	14.60	2.59	0.01
Irregular but consistent	2.17	14.83	0.15	0.88
Irregular and inconsistent	63.24	19.83	3.19	0.00

Table 2: Results from Partially Linear Regression, weekly wages on observed characteristics with data from the CPS in 2004 and 2001. Irregular means bad hours, night shifts. Inconsistent means unpredicatability.

6 Proposal 2: Inattentiveness measures

The authors have 3 measures of inattentiveness, I don't fully agree with any of them, however the one they use in the main specification is decidedly the worst of them:

- Some individuals are given a fake choice, one of the options says that the option is unavailable and to pick the other one. Some people still pick the one "unavailable". It could be that people think there is something wrong with the website, because why give a choice if there is no choice?

- After their choice some people are told whether they picked the classic 9-5 or an alternative, the people that answer wrongly are inattentive. It could be that people were paying attention only to which option paid more money.

[6] I am not fully familiar with how this is properly done in R, however I did discover a package called "recipes" that implements this, in the future I would use this.

[7] It is worth noting that if I took logs of the weekly earnings the estimation completely collapsed.

- The third one is based on preferences, and it is the one used for the results in the paper.[8]

To compute this third measure they assume that people who pick the job without the amenity when the job with amenity pays the maximum(5$) are not paying attention, they are making a "dominated choice". Formally, $\alpha = 1 - \hat{E}[Y|\Delta w = 5]$. They estimate the CEF by OLS $Y_{\Delta w} = \gamma + \beta\Delta w + u_i$ for values of Δw that go from 2 to 5. In the end $\hat{\alpha} = 1 - (\hat{\gamma} + 5\hat{\beta})$. This estimation is done using averages of individuals for a given wage differential and it is not optimal, both in the linearity and the way they implement it. But most importantly by what they assume, there are no strong preferences for a regular 9-5 job at the office by any agent. This can be understood for employer discretion, no one would pick a job where the employer sets the schedule if there are better paid options. However, for the other options there may be heterogeneous preferences, people may not like working from home for various reasons, the same can apply to the other options.

In the paper the authors report $\hat{\alpha}$ is 0.13,0.133,0.145, for the 3 methods respectively, but in the main specification these numbers are misleading, their true numbers are bigger. These are the estimated inattentiveness rates for different amenities:

Specification	Flex schedule	Flex hours	Work from home	Combined flex	Employer discretion
Estimated α	0.15	0.18	0.24	0.24	0.13

Table 3: Estimated inattentiveness rates in the main specification.

Alpha is defined as the share of inattentive individuals who make the "wrong choice", the total share of inattentive individuals is 2*α, so these shares are as high as 48%. These shares seem wrong, intuitively, for two reasons. First, the shares seem too high, surely there are people applying quickly without paying too much attention but that these are half the people applying seems too high. Second, more importantly, for the econometric analysis we have assumed that we are drawing individuals from

[8] The authors show robustness to the other two forms of estimation in the appendix.

the same population, it makes no sense that our chosen measure of inattentiveness gives different rates for different treatments when ex ante the individuals are assigned randomly to the different treatments. The sample is small but this seems to be beyond what sampling error could explain. The way the alpha is estimated could have bad effects for the point estimates and also for their standard errors.

I extended and improved their analysis estimating alphas in more robust ways, in general the mean willingness to pay for the different amenities doesn't change, however, the variance of the WTP and consequently the quantiles do change. The standard deviations are also smaller as a consequence of reducing the variance in the estimated alphas. I estimated 5 specifications. First, I estimated the results assuming all individuals are attentive[9] i.e. $\alpha = 0$. Secondly, I estimated the alphas by 1-mean of people who accepted the job with the amenity for $\Delta w \geq 2$ for each different specification; my reasoning here was that at high levels of Δw the acceptance rate seems to stabilize, so estimating this using OLS didn't make much sense. Third, I used the same methodology as in 2 but now the alphas were estimated from a pooled sample of the 5 main specifications. This was done because beyond some sampling error, alphas should be the same for all sub populations if the sample was iid. Fourth, I used the same methodology as in 2 but using the employer discretion sample. My reasoning here was that if people chose the employer discretion job while the baseline paid more, then they were inattentive, this seems like a weaker assumption, no one should prefer that their schedule is set by their boss.

As I previously mentioned, the mean WTP is largely the same as the baseline for all specifications. In specifications 2 and 3 all estimates are very similar to the baseline, however the precision of these estimates does improve. This is due to the fact that the variance of my alpha estimates goes down. In specifications 1 and 4 results are different, the variance and the 75th quantile also increase significantly. However, the

[9]They have this table in their appendix but it seemed a natural place to start

standard errors are a bit larger.

Finally, I modeled attentiveness as a choice rather than a state, I posited that individuals were more likely to be inattentive if the monetary difference between the two jobs was small. A simple way to test this was to assume:

$$\alpha_i = a + \beta|\Delta w|$$

To compute this linear function I assumed that $\alpha_i(0) = 0.24$ and $\alpha_i(5) = 0.13$, these numbers are the alpha from work from home and the alpha from employer discretion respectively. Modifying the likelihood function slightly allowed for estimation of this specification. In the end the results derived from this assumption were very similar to the baseline.

The tables for different specifications are in the appendix.

7 Policy implications

It is interesting to think about what the policy relevant quantity is here, while the mean WTP is very robust, under the frictionless theory and perfect sorting, the compensating wage differential in the market will be the valuation of the marginal worker who works in a job with the amenity. In the UAS (Understanding America Study) survey data[10] 20% of hourly workers report scheduling flexibility and 10% have formal work from home arrangements. So, under these restrictive assumptions, these high percentiles (80th-90th) may be the relevant policy quantity as they will determine the compensating wage differential in equilibrium. In extensions 1 and 4, these percentiles are significantly higher than in their baseline, and so they imply higher compensating wage differentials in the real world.

[10]They essentially run the same experiment in this survey but with choice being in a hypothetical, non marketplace setting.

That being said, these assumptions are restrictive, once one moves away from frictionless perfect competition into matching models the policy relevant quantity may become a different one.

8 Conclusion and looking ahead

Mas and Pallais (2017) estimate compensating wage differentials in an experimental but nonetheless marketplace setting. While their experiment is as realistic as possible, it is still unlikely that a company offers a choice like the one in the experiment to workers. In other words, the external validity of the experiment is suspect, it seems realistic that search and matching play a determinant role in setting these differentials. Pure cross-sectional methods are not likely to yield good estimates, even when using the most modern methods, however with matched employer employee data these methods could come in handy to flexibly estimate the confounding factors. Under perfect sorting and perfect competition, the policy relevant quantity is likely to be a high quantile, the estimation of these quantiles is sensitive to the way we measure inattentiveness, and less stringent assumptions, such as measuring the alpha from the employer discretion data, suggest that their results may be an underestimate. Alternative measures of alpha also yield more precise estimators.

On a personal note, I really enjoyed doing this project as it allowed me to learn about compensating wage differentials and discrete choice models at the same time, even if the methodology the authors used was distinct from most of the topics we have discussed in class. In the future I would like to fully understand this DoubleML theory and R package as it seems it has a lot of potential uses. It could also be worth it to invest time into understanding the data preparation packages in R and maybe I will when I have some free time.

Happy holidays!

References

Bach, P., Chernozhukov, V., Kurz, M. S., and Spindler, M. (2023). Doubleml - an object-oriented implementation of double machine learning in r. *Journal Name*. Preprint.

Lavetti, K. (2023). Compensating wage differentials in labor markets: Empirical challenges and applications. *Journal of Economic Perspectives*, 37(3):189–212.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–59.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, UK.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2nd edition.

9 Appendix

	Estimate.	Std. Error	t value	Pr(>|t|)
Flexible hours	24.28	3.40	7.14	0.00
Any work from home	126.01	2.55	49.46	0.00
Formal work from home	118.52	2.75	43.11	0.00
Any irregular	0.54	2.65	0.20	0.84
Irregular but consistent	-24.71	2.58	-9.58	0.00
Irregular and inconsistent	29.55	2.62	11.30	0.00

Table 4: Results from Interactive Regression Model, weekly wages on observed characteristics with data from the CPS in 2004 and 2001. Irregular means bad hours, night shifts. Inconsistent means unpredictability.

	Mean	SD	25th quantile	50th quantile	75th quantile
Flexible schedule	0.59	4.51	-2.14	0.59	3.32
Standard deviation	0.23	0.52	0.33	0.23	0.44
Flexible number of hours	-0.13	4.87	-3.07	-0.13	2.82
Standard deviation	0.24	0.61	0.43	0.24	0.45
Work from home	1.44	6.38	-2.42	1.44	5.31
Standard deviation	0.38	1.08	0.53	0.38	0.93
Combined flexible	1.26	6.60	-2.73	1.26	5.26
Standard deviation	0.36	1.05	0.52	0.36	0.89
Employer discretion	3.74	5.43	0.45	3.74	7.03
Standard deviation	0.56	0.87	0.30	0.56	1.05

Table 5: Moments of WTP distribution for different amenities assuming α=0.

	Mean	SD	25th quantile	50th quantile	75th quantile
Flexible schedule	0.36	1.33	-0.45	0.36	1.17
Standard deviation	0.22	0.75	0.34	0.22	0.63
Flexible number of hours	-0.27	1.41	-1.13	-0.27	0.58
Standard deviation	0.20	0.46	0.30	0.20	0.38
Work from home	1.32	1.69	0.29	1.32	2.34
Standard deviation	0.29	0.48	0.33	0.29	0.48
Combined flexible	1.17	2.29	-0.22	1.17	2.56
Standard deviation	0.30	0.54	0.33	0.30	0.53
Employer discretion	3.42	3.04	1.57	3.42	5.26
Standard deviation	0.43	0.55	0.35	0.43	0.68

Table 6: Moments of WTP distribution for different amenities computing α using 1-mean acceptance for those individuals with $\Delta w \geq 2$. Individuals are those with a given amenity

	Mean	SD	25th quantile	50th quantile	75th quantile
Flexible Schedule	0.25	0.52	-0.07	0.25	0.56
Standard deviation	0.20	0.63	0.30	0.20	0.53
Flexible number of hours	-0.26	1.72	-1.30	-0.26	0.78
Standard deviation	0.20	0.48	0.32	0.20	0.38
Work from home	1.34	2.21	-0.01	1.34	2.68
Standard deviation	0.28	0.61	0.39	0.28	0.53
Combined flexible	1.18	2.77	-0.50	1.18	2.86
Standard deviation	0.32	0.65	0.39	0.32	0.60
Employer discretion	3.40	2.01	2.18	3.40	4.62
Standard deviation	0.50	0.45	0.37	0.50	0.72

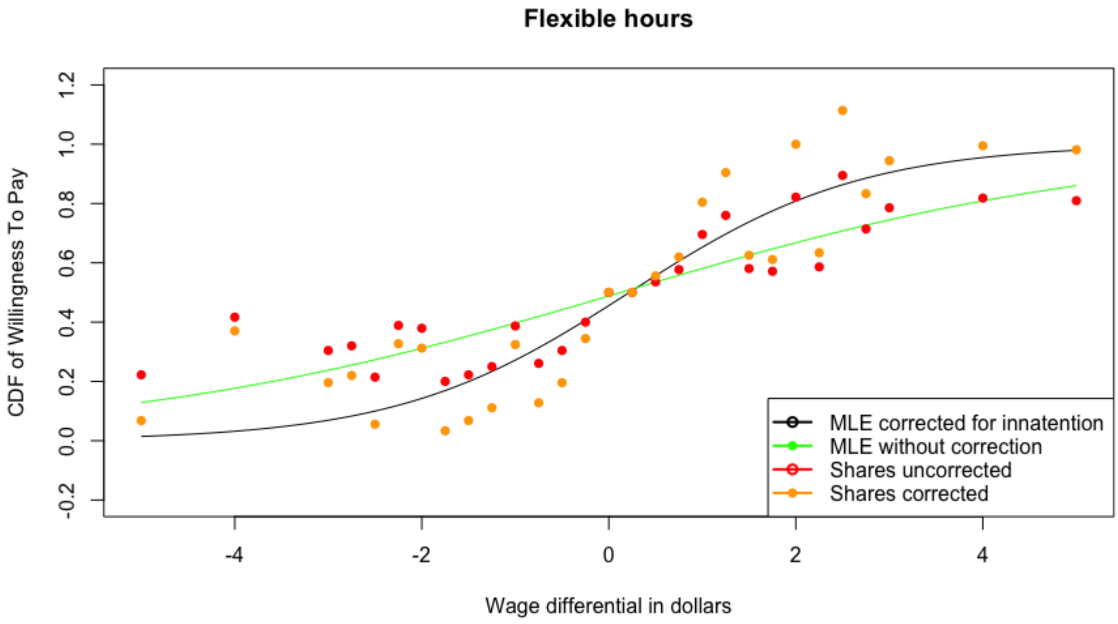
Table 7: Moments of WTP distribution for different amenities computing α using 1-mean acceptance for those individuals with $\Delta w \geq 2$. Individuals are pooled from the 5 main specifications.

	Mean	SD	25th quantile	50th quantile	75th quantile
Flexible schedule	0.53	2.68	-1.10	0.53	2.15
Standard deviation	0.21	0.66	0.38	0.21	0.52
Flexible number of hours	-0.17	3.14	-2.07	-0.17	1.73
Standard deviation	0.22	0.66	0.43	0.22	0.48
Work from home	1.39	4.01	-1.04	1.39	3.82
Standard deviation	0.32	0.98	0.57	0.32	0.77
Combined flexible	1.22	4.40	-1.44	1.22	3.89
Standard deviation	0.33	1.04	0.59	0.33	0.82
Employer discretion	3.42	3.04	1.57	3.42	5.26
Standard deviation	0.46	0.69	0.38	0.46	0.79

Table 8: Moments of WTP distribution for different amenities computing α using 1-mean acceptance for those individuals with $\Delta w \geq 2$. Individuals are from the employer discretion specification.

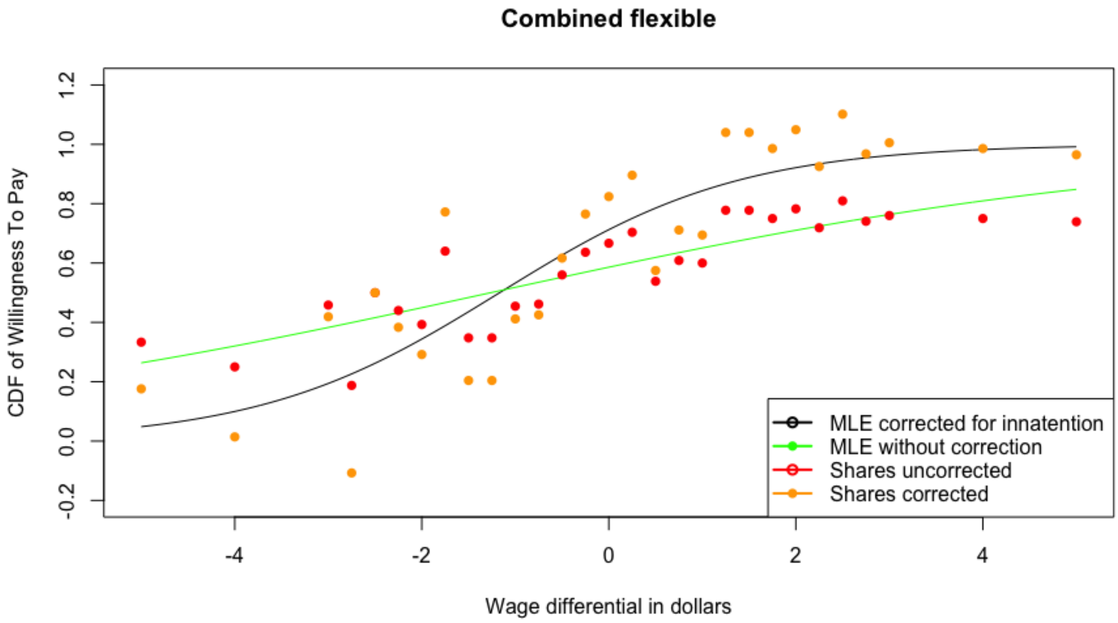
	Mean	SD	25th quantile	50th quantile	75th quantile
Flexible schedule	0.45	1.57	-0.51	0.45	1.40
Standard deviation	0.26	0.84	0.37	0.26	0.72
Flexible number of hours	-0.23	2.17	-1.54	-0.23	1.09
Standard deviation	0.22	0.57	0.37	0.22	0.45
Work from home	1.46	2.84	-0.26	1.46	3.18
Standard deviation	0.32	0.86	0.51	0.32	0.70
Combined flexible	1.32	3.39	-0.73	1.32	3.38
Standard deviation	0.36	0.78	0.45	0.36	0.71
Employer discretion	3.60	2.25	2.24	3.60	4.97
Standard deviation	0.45	0.39	0.34	0.45	0.63

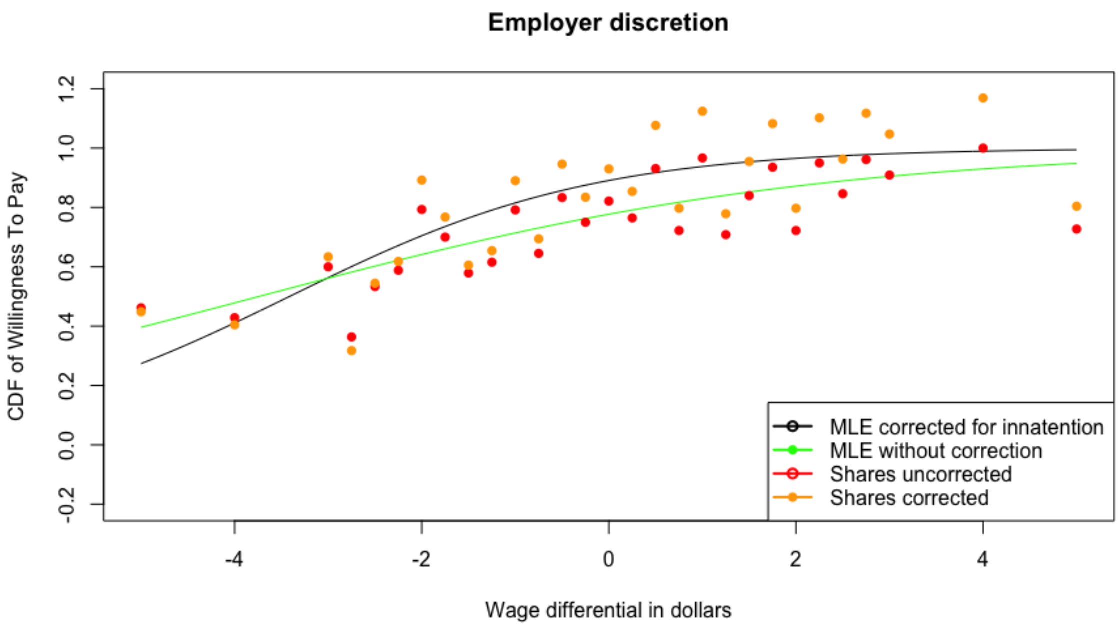
Table 9: Moments of WTP distribution for different amenities with alpha decreasing for larger $|\Delta w|$.



Flexible hours
CDF of Willingness To Pay
Wage differential in dollars
MLE corrected for innatention
MLE without correction
Shares uncorrected
Shares corrected

Work from home
CDF of Willingness To Pay
Wage differential in dollars
MLE corrected for innatention
MLE without correction
Shares uncorrected
Shares corrected


Combined flexible
CDF of Willingness To Pay
Wage differential in dollars
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
-4
-2
0
2
4
MLE corrected for innatention
MLE without correction
Shares uncorrected
Shares corrected


Employer discretion
CDF of Willingness To Pay
Wage differential in dollars
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
-4
-2
0
2
4
MLE corrected for innatention
MLE without correction
Shares uncorrected
Shares corrected