Answers to Referee 2

We thank the reviewer for his or her questions and suggestions. To simplify the discussion, we repeat the reviewer's main points in bold typeface, followed by our replies in standard typeface.

His first point regards the specification of the model and data generation:

> **The authors could easily have utilised their data much more efficient. Now the authors are using occupation-year cells. That means that they are throwing away all firm-level variation and instead base their analysis on aggregate variables. One solution would have been to make shares of low-skilled workers at firm level. Then the authors could utilize firm characteristics. Due to the way the authors set up the model, they cannot use (the workhorse) OLS and instead have to rely their empirical analysis a grouped probit model. This is mainly caused by their use of aggregate unemployment rate for Germany. But the authors know where the firms are located. Thus, they could easily build up more disaggregated unemployment rates and therefore complement their current analysis with OLS estimates (which would be more easily understood by the general reader). Using more disaggregated unemployment rates may also be more relevant if the regional mobility among workers is low.**

In summary, the reviewer suggests to conduct the analysis at a more disaggregate level to gain precision and to make estimation by OLS possible. A closer inspection of the issue reveals that in principle *aggregation is not necessary at all*. The grouped probit model is aimed to handle data which were aggregated from individual binary decisions.[1] As we have individual level data, we could use a standard probit model to explain the binary dependent variable (indicating whether the hired worker is skilled or not) on the regressors. The reviewer emphasizes that the grouped probit model is required *only* because of the aggregation and that it could be replaced by OLS after changing to a more disaggregate analysis. Here we disagree: *Both* the grouped probit and the standard individual level probit can be replaced by OLS, but OLS is an approximation in both cases because the individual level dependent variable is binary.[2] Let us remark that we have run individual level regressions (linear probability models) in the experimentation

[1] The employer decides whether to hire a worker with higher or lower qualification.

[2] Note that we present OLS regression results based on the cell data, and that our results recommend OLS as a good approximation.

stage of the project but abandoned it because it would have increased the size of the paper without adding substantial information.

The individual level analysis has two possible advantages: First, it allows to control for individual level characteristics such as sex, age and establishment size. Our test regressions show, however, that controlling for individual level characteristics has a negligible impact on the coefficient of the unemployment rate. For example, the unemployment estimates for the high-skilled share reported in our paper are 0.098 for the linear grouped data OLS and 0.102 for the grouped probit model. The corresponding estimates obtained from the *individual data* linear probability model is 0.096. Adding a female dummy and second order polynomials in age and log establishment size to this model increases the coefficient slightly to 0.108. That the inclusion of individual level controls practically does not matter is intuitively clear because omitted variable bias requires correlation between regressors. The correlation between *individual characteristics* like sex, age and the establishment size and the aggregate unemployment rate is tiny.

A second advantage of the individual level analysis may be due to the possibility to include the regional (district level) unemployment rate instead of the aggregate unemployment rate. At first glance this seems to raise the precision by increasing the degrees of freedom from 1 to 326 per year (the number of observations of the unemployment rate per year is then equal to the number of districts, 326).

A closer look reveals, however, that the change of the regressor implies a *change of the meaning* of the unemployment coefficient. Reder's theory is aimed to explain the response of employers to *aggregate cyclical fluctuations* which may be different from responses to idiosyncratic, regional or sectoral shocks. If a single firm (or region) grows above average, bidding up wages to attract workers from competitors (or neighbor regions) may be more sensible than if the same firm faces an aggregate cyclical shock in product demand. This is intuitively clear since the firm- or region-specific supply of skills is more elastic than the *aggregate one* which should be practically *fixed* in the short run as it takes months or years to 'produce' these skills. Because of that skill proportions regressions based on regional unemployment rates will deliver *a mixture of responses to aggregate and regional shocks*. Note that e.g. the large empirical U.S. literature on wage cyclicality is based on *aggregate* unemployment rates for that reason (although regional unemployment rates were available). Although it were interesting per se to test whether employers' wage setting depends on the 'aggregation' level of shocks, the results are not in the focus of our paper and cannot be used for a comparison with Devereux's U.S. results.[3]

[3]It is possible to isolate responses to purely regional shocks by the application of a two-step procedure that is commonly used in the wage curve literature: In the first stage skill proportions are regressed on year dummies for each district separately. The coefficients of the year dummies are then regressed on the regional unemployment rates and complete

In summary, we would prefer not to shift the analysis to a more disaggregate level, because (1) individual level controls are practically irrelevant, (2) the coefficient on regional unemployment rates has a different meaning which is not relevant in our investigation, and (3) the comparison with the U.S. results has to mimic Devereux's specification. Nevertheless it appears sensible to present the results from our individual level robustness checks in a future version of the paper.

> **In their writing the authors are positioning themselves very closely to Devereux op.cit. Thus, the reader gets the impression that this paper is just a replication, not something new.**

This is correct and was intended. We have deliberately not added something original, as this would hinder easy comparison. The value of our contribution rests not so much in theoretical originality, but rather relates to the quality of the data set that we were able to use (and which seems to be superior to the one used by Devereux), and to the attention we devoted to various technical details. In our view, such replication for different countries are warranted for the effects in question, just as is the case with Phillips curves or wage curves. An important issue in the context of Reder competition is to check whether it captures general aspects of hiring decisions and wage setting which are not closely related to specific institutional environments. In this regard, Germany suggests itself as a quite interesting comparison case for the U.S.

> **What I find interesting is that changes in skill composition over the business cycle also takes place within the German labour market, commonly known as heavily standardised and regulated with strict certification requirements for several occupations. But this might also be the curse of the blessing; i.e. the authors need to argue much better why the occupational change over the business cycle is possible in such a labour market. I therefore recommend a substantial rewriting and rephrasing.**

From the point of view of Reder's theory the high unemployment coefficient is not surprising: According to the central assumption employers adjust hiring standards instead of bidding up wages. If wage setting is more

sets of district and year dummies. This generates purely regional responses by sweeping out the effects of aggregate unemployment shocks.

restricted by collective agreements in Germany, adjustment of hiring standards appears to be an *obvious* alternative. If German wages are more rigid (compared with the U.S.) in that sense, we would expect even more pronounced effects in Germany. That we observe lower effects instead is likely to be caused by a more structured occupational qualification system in Germany which implies less permeable barriers between qualification groups. This is outlined in our paper already but the point should be stated more clearly. Nevertheless we deliberately avoided to add further speculation regarding the impact of the occupational qualification system which cannot be substantiated empirically with our data.

> **Furthermore, the paper is quite long and should be shortened. The current version is 23 pages.**

We agree that we should try to shorten the paper and shift details of data description to the appendix. As a further simplification we will base the comparison with Devereux's results on the linear grouped data model and describe the grouped probit in a more compact manner or put the details into the appendix.