Online Appendix for "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

Data Appendix

1. More details about Data Sources and Construction

For the firm data, the firm identifier in FirmStat is derived from the register "Old Firm Statistics" for the period 1995-1999 and from "General Firm Statistics" for the period 1999-2006. These two registers in combination allow us to track the same firm during the entire period 1995-2006 despite the structural break in 1999. For our firm-characteristics variables, the number of employees is from FirmStat and is calculated as the number of full-time equivalent workers in a year. Capital stock, measured as the value of land, buildings, machines, equipment and inventory is from the Accounting Statistics register. Gross output (net of taxes) is from the VAT register. Firm-level skill-intensities are computed using the educational attainment records of individual workers in IDA which are then aggregated to the firm-level using the matched worker-firm link (FIDA).

For the worker data, we use annual hours which is common in the literature (e.g. Christensen et al. 2005). A concern is that annual hours do not capture overtime work. For a portion of our sample in 2006 we have data for overtime work. A wage rate including overtime is correlated 0.86 with our main wage-rate variable, and overtime hours are uncorrelated with offshoring (0.015 for the full sample and -0.017 for the subsample of high-skilled workers). This suggests that our results are unlikely to be driven by the issue of overtime work. We measure labor market experience as actual time in employment since 1964. Other worker-level information regarding union membership and marriage are also derived from the IDA database. We experimented with breaking low-skilled workers into two subgroups, medium-skilled (those with a vocational education, defined as the final stage of secondary education that prepares students for entry into the labor market) and very-low-skilled (those with the equivalent of high school education or less). We obtained very similar results.

Our trade data, the Foreign Trade Statistics Register, consists of two sub-systems, Extrastat (trade with non-EU countries) and Intrastat (trade with EU countries). Extrastat has close-to-complete coverage as all extra-EU trade flows are recorded by customs authorities. Intrastat does not have complete coverage because firms are only obliged to report intra-EU trade if the annual trade value exceeds a threshold. In 2002 the thresholds were DKK 2.5 million for exports and DKK 1.5 million for imports.

After merging data on manufacturing workers, firms, and trade flows, we have 2.8 million worker-firm-year observations. We then trim our sample as follows. Since we have annual data we cannot investigate the changes in wage or employment status at weekly, monthly or quarterly frequencies. Thus we drop all the worker-firm-year observations of which the employment relationship, or job spell, lasts for a single year (about 200,000 observations). We also drop all the workers whose skill level changes in our sample period (about 35,000 observations), in order to get a clean identification of how the effects of offshoring vary across skill groups. We next drop the firms with fewer than 50 employees and less than 0.6 million DKK in imports, which corresponds to average annual wages for two manufacturing workers. This eliminates another 600,000 observations. This de minimis restriction eliminates from our sample very small firms who in some cases have imputed balance sheet variables and are more likely to fall below the reporting thresholds for intra-EU trade data.

For data on occupational characteristics, The occupation variable in IDA is based on a Danish version of the International Standard Classification of Occupations (ISCO-88) developed by the International Labour Office (ILO). We map the O*NET data into the ISCO-88 classification system using the crosswalk at the National Crosswalk center ftp://ftp.xwalkcenter.org/DOWNLOAD/xwalks/. For non-routine tasks we use the principal component of mathematical reasoning (O*NET task id 1.A.1.c.1), response orientation (1.A.2.b.3), gross body coordination (1.A.3.c.3), mathematics (2.A.1.e), thinking creatively (4.A.2.b.2), and organizing, planning, and prioritizing work (4.A.2.b.6). For routine tasks we use manual dexterity (1.A.2.a.2), finger dexterity (1.A.2.a.3), multilimb coordination (1.A.2.b.2), processing information (4.A.2.a.2), and evaluating information to determine compliance with standards (4.A.2.a.3). For social sciences we use the principal component of 2.C.1 (2.C.1.a, 2.C.1.b, etc.), 2.C.6, 2.C.7, 2.C.8, 2.C.9, 2.C.4.e, and 2.C.4.f. For natural sciences we use 2.C.2, 2.C.3, 2.C.5, 2.C.4.b, 2.C.4.c, 2.C.4.d, 2.C.4.g, 2.C.10, and 2.A.1.f. For communication and language we use 4.A.4.a, 2.B.1, 1.A.1.a, 4.C.1.a.4, 4.C.1.b.1, 2.A.1.a, 2.A.1.b, 2.A.1.c, and 2.A.1.d. For on-the-job hazards we use 4.C.2.c, 4.C.2.b.1, and 4.C.2.e.1.

2. Construction of the transport-cost instruments

The Danish trade data do not contain information on transportation costs paid by firms. To construct transportation costs we proceed in two steps. First, we use data on ad-valorem shipping costs taken from US sources to estimate costs as a function of transportation mode, product weight/value, fuel prices, and distances shipped. Second, we construct fitted cost measures using these same variables that are specific to Danish firms.

During our pre-sample years, 42% of Danish imports by value arrive by sea, 20% by air, 37% by truck, and 1% by rail. For sea and air transport we employ data on transportation costs taken from US Imports of Merchandise data for the 1995-2006 sample period. For sea transport we estimate

$$f_{ckt} / v_{ckt} = -4.16 + 0.393 \ln \frac{w_{ckt}}{v_{ckt}} + 0.351 \ln oil_t + 0.027 DIST_c + .0063 \ln oil_t * DIST_c$$

where c indexes exporters, k indexes HS6 products, t = year, f = transportation charge, v = value of shipment, and w = weight in kg, DIST = distance in 1000km measured to the nearest US coast. For air transport we estimate

$$f_{ckt} / v_{ckt} = -3.80 + 0.435 \ln \frac{w_{ckt}}{v_{ckt}} + 0.209 \ln JETFUEL_t + 0.033 DIST_c + .018 \ln JETFUEL_t * DIST_c$$

Note that this generates air shipping costs that are higher in levels, more sensitive to fuel prices, and more sensitive to the interaction between fuel prices and distance. Also, jet fuel prices, while correlated with crude oil prices, can vary from year to year as a function of differences in refining capacity and availability of high grade crude suitable for distilling light fuels.

For rail and truck transport we draw on transportation costs taken from the US Transborder Surface Freight data, which reports US state to Canadian province flows at the HS2 level monthly from 1994-2006. For truck transport we estimate

$$f_{spkt} / v_{spkt} = -8.18 + 0.234 \ln \frac{w_{spkt}}{v_{spkt}} + 0.862 \ln oil_t + 1.21 \ln DIST_{sp} - 0.373 \ln oil_t * \ln DIST_{sp}$$

Here, sp refers to a state-province pair. Note that we use log distance rather than levels as it provides a better fit to the land-based data. For rail transport we estimate

$$f_{spkt} / v_{spkt} = -4.37 + 0.54 \ln \frac{w_{spkt}}{v_{spkt}} + 0.079 \ln oil_t - 0.90 \ln DIST_{sp} - 0.224 \ln oil_t * \ln DIST_{sp}$$

With the exception of the rail cost function (which represents only 1 percent of our sample), these estimates are broadly consistent with estimates in the literature.

We then take the coefficients from this regression to construct the costs that would face a Danish firm with similar shipment characteristics. This is specific to each input purchased. Oil prices and distance are the same for all firms. We use data on transport mode used and weight/value ratio for all firms purchasing a particular c-k input; however to avoid introducing endogeneity we use pre-sample information in both variables. We construct transport costs for each input from the fitted equation as $\tau_{ckt} = \exp(f_{ckt} / v_{ckt})$ and aggregate over inputs using the share of each input in pre-sample trade for each firm.

To understand the source of variation generated by this approach, note that inputs travel different distances, have different bulk (product weight/value), and use different transport modes. Over time there are shocks to the level cost of each transport mode as a function of technological change and input prices. Further, oil prices fluctuate substantially in our sample, falling for 4 years and then rising sharply, as we show in Figure A1. Shocks to oil prices differentially affect costs depending on which mode is used and how far goods travel.

Online Appendix for “The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data”, by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

3. Measures of offshoring: details

3.1 Input-Output Tables

An alternative approach to identify imported inputs, commonly used in the literature, employs input-output (IO) tables. Under the “proportionality assumption” (see Feenstra and Jensen 2010 for limitations) that all firms in an industry use the same inputs and in the same ratios and that imports and domestic supplies have the same market share, one can use IO table input coefficients interacted with shocks to trade costs to generate industry-time variation in the desirability of offshoring. We do not employ this method for three reasons. One, we employ industry-time fixed effects in the estimation in order to control for demand shocks, and this eliminates all variation that can be exploited using an IO table. Two, unlike the literature we see the actual inputs purchased by firms and the data strongly reject the assumption of common input usage within an industry. Three, because the data indicate significant within-industry variation across firms in both inputs and source countries, we can use exporter-product-time variation in our instruments to better explain changes in offshoring.

Nevertheless, the IO tables provide a useful additional check on the firm level data. The most disaggregated Danish IO table is at the industry-level, covering 57 manufacturing industries, and does not distinguish inputs by source. This is in contrast to our country-product disaggregation, where we have over 13,500 distinct inputs. We make use of the import matrix of the Danish IO table for two cross-checks. First, are there instances where our trade data says that product k is an imported input for firm j, but the IO table disagrees (i.e. the industry of k is not an imported input for the industry of j)? This occurs for only 2 percent of cases.

As another cross-check on our firm-level data, we construct a hypothetical import matrix of the IO table using our trade data and compare it with the import matrix of the official Danish IO table. They are not identical, because the official IO table employ the proportionality assumption . Nevertheless, there is a broad correspondence between the inputs used by our firms and what we seem in the IO table. The input shares of our constructed import matrix have a 0.73 correlation with the official IO table.

3.2 Machinery Imports

Imports of machinery are potentially problematic in terms of interpretation. Access to foreign technology embodied in machinery imports may affect labor demand and wages (e.g. Hanson and Harrison 1999) but through a different channel than offshoring of material inputs that could have been produced by the firm. While we do not take a strong stand that we can completely separate the effects of offshoring material inputs versus technological change embodied in machinery imports, we do want to distinguish where such effects are likely to appear in our analysis.

The HS system classifies most types of machinery in HS84, "Nuclear reactors, boilers, machinery etc…", and HS85, "Electric machinery etc; sound equipment; TV equipment …". Our broad offshoring measures include imports of HS 84 and HS 85 for all firms, and this represents 16.9% of imports. Our narrow offshoring measure excludes machinery imports for all firms except for those who also produce machinery for sale. For firms that produce machinery for sale, narrow offshoring could potentially include machinery imports. The question for these firms is whether imports within HS 84, 85 represent machinery itself or parts for machinery. As an example, consider the five largest firms selling in HS 8413, "Pumps for liquids…". The top three import categories are HS 8413 itself, which could be machinery, and HS 8483, "Transmission shafts, bearings, gears…", and HS 8481, "Taps, cocks, valves…" which are clearly parts. We found similar results for the top five firms in HS 8481 and HS 8482, "Ball or roller bearings…".

At more disaggregated levels of data it is possible to distinguish machinery from parts of machinery. Looking over all firms and imports we ranked the value share for each six digit product within HS 84. Table A1 lists the top 20 products, comprising 59% of the imports of HS 84. All are parts, and not machinery itself. The largest HS6 import that is clearly a machine and not parts of a machine is HS 842240, "Packing or wrapping machinery…" It ranks 34th on the list and its share in imports is 0.007%. The results are similar for HS85. Therefore, even in those HS categories where machinery imports are concentrated, actual machinery accounts for a small share of total imports.

4. Robustness of Results to Sample Selection

To gauge the effects of our sample selection criteria, Table A2 compares the summary statistics of our estimation sample with the full sample. The full sample is the collection of workers and firms in the manufacturing sector we have before we implement our sample-selection criteria. The numbers under the heading "Estimation Sample" are identical to those in Table 1. As compared with the full sample our estimation sample has slightly larger firms that employ slightly more experienced workers with slightly higher wages, but the differences are very limited.

A related concern is that we include only the firms in the years in which they both import and export. If a large fraction of firms import/offshore but do not export, then our estimation sample may not be representative of the Danish manufacturing sector. Table A3 breaks down the trading firms (i.e. the firms that import/offshore, or export, or both) in the full sample (as defined for Table A2 above) by employment and trade categories, and shows the shares in employment (upper panel), output (middle panel) and import value (lower panel). Table A3 shows that offshoring-only firms account for 2% of employment, 1.6% of output and 0.7% of import value among all trading firms; in comparison, the firms that both offshore and export account for 86.6% of employment, 90% of output and 99% of import value. These results suggest that the offshoring-only firms are a small fraction of our Danish data.

In Table A4, we present the results of within-job-spell wage regressions with the offshoring-only firms added into the sample (columns 1-4). The number of observations (1.98 million) is very similar to our estimation sample (1.93 million, as in Table 5), consistent with the results of Table A2. Since the offshoring-only firms have export values of 0 we replace the two instrumented export variables in our main estimation with the un-instrumented variables of log(export value + 1) (columns 1 and 3), and export values as a share of output (columns 2 and 4). The coefficient estimates for offshoring and its interaction with high-skill dummy are similar to Table 5.

In Table A4 we also present the results of within-job-spell wage regressions for a balanced panel of firms that are in the sample in all years (columns 5-8). The balanced-panel sample has 40%, or roughly 800,000, fewer observations than in our estimation sample. Despite this reduction in sample size, the coefficient estimates for offshoring and its interaction with high-skill dummy are again similar to Table 5. However, the coefficient estimates for log exports are about twice their size than in Table 5, and there is a larger difference between the effects of exports on high- and low-skill wages, and a smaller difference for high and low-skill earnings. These are likely because we are unable to fit the log-export regression as well in the first-stage IV. To be specific, the WID instrument for export has a smaller coefficient estimate than in Table 4 and is not statistically significant when firm controls are included. When firm controls are not included, the WID instrument has a similar coefficient estimate to Table 4 but the WES instrument has a larger coefficient estimate. In addition, the F-statistics for the instruments are also smaller than in Table 4 (2.59 and 10.39, respectively, vs. 8.62 and 22.36 in Table 4). These are likely due to the reduction in sample size. The other first-stage IV regressions are similar to Table 4 (the first-stage IV results are available upon request).

5. Additional Robustness Exercises

To save space, we show the following two robustness exercises in Table A5. The other robustness exercises mentioned in note 34 of the text are available upon request.

In the text, we have emphasized narrow offshoring (imports purchased in same industry categories as the firm's sales) because these are more likely to be inputs the firm could have produced itself. In our next robustness check we use broad offshoring (all import purchases by the firm) instead. We find much larger effects of offshoring on wages, and more pronounced differences across skill types. A possible explanation is that broad offshoring includes inputs of all types and is therefore more likely to capture the effect of technological change operating through imports of machinery. Further, the estimation with firm controls yields a much larger wage drop than the estimation without firm controls. This is consistent with the view that the productivity effect, as distinct from the labor substitution effect, can be seen more clearly when imported inputs are different from those made by the firm.

Next, one may be concerned that firm level shocks originating within Denmark may have general equilibrium consequences for product prices in exporting and importing partners if Denmark is responsible for a large share of trade. We experiment with dropping trade flows where Denmark is responsible for more than 1% of trade with that partner and product. Results are very similar.

Online Appendix for "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

Theory Appendix

Generalizing the Production Function

To generalize our production function, equation (1), we have $Y_{jt} = A_{jt} K_{jt}^{\alpha} \prod_{f=1}^{F} C_{fjt}^{\alpha_f}$, where f = 1,2,…,F index types of labor, $C_{fjt} = \left(L_{fjt}^{\theta_f} + M_{jt}^{\theta_f} \right)^{1/\theta_f}$, $\theta_f = \frac{\sigma_f - 1}{\sigma_f}$, and $\sum_{f=1}^{F} \alpha_f = 1 - \alpha$

In words, the production function is Cobb-Douglas in capital (whose share is α) and composite inputs C_f (whose share is α_f). Each composite input C_f is produced with imported inputs M and type-f labor L_f using CES technology with the substitution elasticity $\sigma_f > 1$. σ_f may vary across labor types. Each labor type can be a skill group or an occupation, and different labor types enter into the production function symmetrically. We first show that

(A1) $\ln C_{jft} \approx c_{of} \ln M_{jt} + (1 - c_{of}) \ln L_{jt} + c_{1f}$

where c_{0f}, c_{1f} are constants and $0 < c_{0f} < 1$.

Proof Drop the subscripts j, f, and t, and let $y = \ln(L/M)$. Then $C = M(e^{y\frac{\sigma-1}{\sigma}} + 1)^{\frac{\sigma}{\sigma-1}}$ and $\ln C = \ln M + g(y)$, where $g(y) = \frac{\sigma}{\sigma-1} \ln(e^{y\frac{\sigma-1}{\sigma}} + 1)$. The first-order Taylor approximation for $g(y)$ is $g(y) = g(y_0) + g'(y_0)(y - y_0)$, where y_0 is a constant, and $g'(y_0) = \frac{e^{y_0\frac{\sigma-1}{\sigma}}}{e^{y_0\frac{\sigma-1}{\sigma}} + 1}$ lies between 0 and 1 for all values of y_0. Let $c_0 = g'(y_0)$ and $c_1 = g(y_0) - y_0 g'(y_0)$ and we have equation (A1). **QED**.

Similar to equation (2) in our paper, the marginal product of type-1 labor is $MPL_1 = (1-\alpha) A_{jt} K_{jt}^{\alpha} L_{1jt}^{-\frac{1}{\sigma_1}} C_{1jt}^{\frac{1}{\sigma_1} + \alpha_1 - 1} \prod_{f=2}^{F} C_{fjt}^{\alpha_f}$. Taking the log of MPL_1 and using equation (A1) we obtain

$$\ln MPL_1 = \ln[(1-\alpha) A_{jt} K_{jt}^{\alpha} L_{1jt}^{-\frac{1}{\sigma}}] + \sum_{f=2}^{F} \alpha_f c_{0f} \ln L_{fjt} + +(\frac{1}{\sigma_1} + \alpha_1 - 1) c_{01} \ln L_{1jt}$$
$$+[(\frac{1}{\sigma_1} + \alpha_1 - 1)(1 - c_{01}) + \sum_{f=2}^{F} \alpha_f (1 - c_{0f})] \ln M_{jt} .$$

If $c_{01} = c_{0f}$ for all f = 2,..,F, the coefficient for lnM_{jt} in the expression for lnMPL_1 simplifies to $[\frac{1}{\sigma_1} + \alpha_1 - 1 + \sum_{f=2}^{F} \alpha_f](1 - c_{01}) = (\frac{1}{\sigma_1} - \alpha)(1 - c_{01})$, where the equality uses $\sum_{f=1}^{F} \alpha_f = 1 - \alpha$. Therefore, an increase in M_{jt} increases the demand for type-1 (type-f) labor if $1/\sigma_1 - \alpha < 0$ ($1/\sigma_f - \alpha < 0$). This condition is analogous to what we have in section III.1. Since σ_f differs across labor types, this condition also suggests that an increase in M_{jt} may increase the wage for some labor types (those with small σ_f) but decrease the wage for the other types (those with large σ_f).

Derivation of Equation (3)

Assume that firm j faces the following supply curve for unskilled labor

(A2) $$w_{L,jt} = c(L_{jt})^{\gamma_{L,S}},$$

where $w_{L,jt}$ is the unskilled-labor wage for firm j in year t and γ_{Ls} is the elasticity of supply for unskilled labor. Equations (2) (in the text) and (A2) imply that the response of unskilled wages to offshoring (holding output constant) is

(A3) $$b_{L,M} = \frac{\partial \ln w_{L,jt}}{\partial \ln M_{jt}} |_{K \text{ constant}} = \frac{(1/\sigma - \alpha - \beta) c_0 \gamma_{L,S}}{\gamma_{L,S} - \gamma_{L,D}},$$

where $c_0 \in (0,1)$ is a constant and $\gamma_{L,D} < 0$ is the elasticity of labor demand implied by equation (2). Equation (A3) says that $b_{L,M} < 0$ if $1/\sigma < (\alpha + \beta)$, which is the same condition under which offshoring lowers labor demand. If labor supply is perfectly elastic, $\gamma_{Ls} \to \infty$, then shocks to labor demand will result in employment changes but not wage responses. A similar demonstration shows that offshoring raises skilled labor wages and exporting raises wages for both skilled and unskilled workers.

To derive equation (3), assume that each unskilled worker i has productivity h_{ijt} in year t and $h_{ijt} = \exp(\beta_1 x_{it} + \eta_{ij})$, where x_{it} represents observable worker characteristics (e.g. experience), β_1 is a vector of coefficients, and η_{ij} represents unobservable ability that is specific to the worker-firm match. Unskilled workers are the same up to the productivity term, so that worker i receives wage

(A4) $$w_{L,ijt} = w_{L,it} h_{ijt}.$$

A similar expression governs high skill labor wages. Then it is straightforward to derive equation (3) by solving for $\log w_{L,ijt}$ using equations (A2), (A4) and equation (2) in the text.

Online Appendix for "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

When the production function has multiple types of labor we can carry out the analyses in the same way. In particular, to derive the counterpart of (A3), let $\gamma_{f,S} > 0$ be the labor supply elasticity for type-f labor. Then the wage elasticity for type-f labor is $\frac{\partial \ln w_{f,jt}}{\partial \ln M_{jt}}|_{K \text{ and } L_f \text{ constant}} = \frac{(\sigma_1^{-1} - \alpha) c_{0f} \gamma_{f,S}}{\gamma_{f,S} - \gamma_{f,D}}$, where $\gamma_{f,D} = -[\frac{1}{\sigma_f} + (1 - c_{0f})(1 - \alpha_f - \frac{1}{\sigma_1})] < 0$ is the demand elasticity for type-f labor and c_{0f} is as defined in equation (A1). This expression is analogous to (A3).

The Productivity Effect

Finally, we use Figure A2 to illustrate the effects of offshoring on unskilled wage, with and without the productivity effect. LS is the supply curve for unskilled labor. Suppose that unskilled labor and imported inputs are highly substitutable; i.e. $\sigma > 1/(\alpha + \beta)$. An increase in offshoring shifts the unskilled labor demand curve from LD_0 to LD_1, holding constant physical capital, K_{jt}. This is the direct wage effect of offshoring. As the increase in foreign inputs makes the firm more profitable and the firm increases the use of all inputs in response, there is a secondary shift of the unskilled labor demand curve, rising from LD_1 to LD_2. This is the productivity effect of offshoring and it tends to increase unskilled wage. If the direct effect dominates the productivity effect, LD_2 lies between LD_1 and LD_0.

We now calculate the wage elasticity of unskilled labor inclusive of the productivity effect. We assume that firm j takes the rental rate for capital, r_t, as given, and that firm j increases capital input, K_{jt}, until its marginal revenue product equals the rental rate r_t, or that $r_t = \alpha \psi_{jt} A_{jt} K_{jt}^{\alpha-1} H_{jt}^{\beta} C_{jt}^{1-\alpha-\beta}$, which implies that $\frac{\partial \ln K_{jt}}{\partial \ln M_{jt}} = \frac{\partial \ln C_{jt}}{\partial \ln M_{jt}} \frac{1-\alpha-\beta}{1-\alpha} = c_0 \frac{1-\alpha-\beta}{1-\alpha} > 0$, where $0 < c_0 < 1$ is the same as defined in (A3). Using this expression and equation (2) we can show that $b_{L,M}^* = \frac{\partial \ln w_{jt}}{\partial \ln M_{jt}} = \frac{c_0 \gamma_{L,S}}{\sigma(\gamma_{L,S} - \gamma_{L,D}^*)}$, where $\gamma_{L,D}^* = -\frac{c_0}{\sigma} < 0$ is the elasticity of unskilled labor demand inclusive of the productivity effect. Comparing this expression with (A3) we show that $b_{L,M} < b_{L,M}^*$; i.e. the productivity effect tends to increase the wage for unskilled labor.

Instrumental Variables Strategy

Index Danish firms by j, years by t, exporting countries by c, products (measured at the HS-6 level) by k, and destination countries other than Denmark by d. For ease of exposition, assume that firm j only imports a single product k from a single destination country c (the case of multiple product x country is similar). Firm j's production function is given by equation (1) in the text, but we re-write firm j's imported inputs as.

(A5) $$M_{jt}^{\theta} = \lambda_{ck}(b_{ck}^{j})^{\frac{1}{\sigma-1}}\left(q_{ck}\right)^{\theta}, \theta = \left(\sigma-1\right)/\sigma,$$

In equation (A5), λ_{ck} is a preference shifter for a given exporter-product that is common across importers. This can be thought of as quality, or in cases where there are multiple firms providing product k in exporter c, this can be interpreted as variety. In addition, there are preference weights, b_{ck}^{j}, that are idiosyncratic to each firm j. As we note in Section III, a feature of the Danish data is that for a given firm j, b_{ck}^{j} is positive for a small number of inputs, and the modally occurring case is that for a given c,k, b_{ck}^{j} is positive for only one Danish firm.

By the production function (1) and equation (A5), the value of firm j's imported inputs equals

(A6) $$p_{ckt}q_{ckt} = m_{ckt}^{j} = \underbrace{(1-\alpha-\beta)}_{\text{Constant}}\underbrace{(p_{ckt}^{1-\sigma}\lambda_{ckt}^{\sigma-1})}_{\text{c-k Supply char.}}\underbrace{\tau_{ckt}^{-\sigma}}_{\text{Denmark Trans. Cost}}\underbrace{(\frac{b_{ck}^{j}E_{t}^{j}}{P_{t}^{j}})}_{\text{Firm-j Demand char.}}, P_{t}^{j} = b_{ck}^{j}p_{ckt}^{1-\sigma}\lambda_{ckt}^{\sigma-1} + w_{L,jt}^{1-\sigma}$$

In equation (A6) E_{t}^{j} is the output value of firm j, which reflects demand for j's outputs, and P_{t}^{j} represents the CES price index for j's imported inputs and low-skilled wage. Equation (A6) says that imports of a particular input ck can rise because there is a shock to supply characteristics (price, quality, variety), shocks to transport costs, or because overall demand for inputs rises. We want to isolate the component of the firm's changing import demand that arises from shocks to supply characteristics or to transport costs. We will measure transportation costs directly. We next discuss how we identify shocks to supply characteristics.

One possible approach is to use highly disaggregated gravity equations to separately identify shocks to supply and demand characteristics. This is similar in spirit to Redding and Venables (2004), except that they use aggregate bilateral trade in a single cross section and extract aggregate supply characteristics by exporter country. To see how this approach works, suppose that input demand in the rest of the world is of a similar form to demand in Denmark, though we write preference weights as idiosyncratic to a destination d. Imports into destination d are then

$$m_{ckt}^{d} = \underbrace{(1-\alpha-\beta)}_{\text{Constant}}\underbrace{(p_{ckt}^{1-\sigma}\lambda_{ckt}^{\sigma-1})}_{\text{c-k Supply char.}}\underbrace{(\tau_{ckt}^{d})^{-\sigma}}_{\text{country-d Trans. Cost}}\underbrace{\sum_{l}\frac{b_{ck}^{d,l}E_{t}^{d,l}}{P_{t}^{d,l}}}_{\text{country-d Demand char.}}, \tag{A7}$$

where the summation is over the firms l in destination country d. Assume that there is no idiosyncratic component to demand (i.e. if $b_{kt}^{d} = b_{ckt}^{d,l}, \forall c$), then equation (A7) can be rewritten as

(A8) $\ln m_{ckt}^{d} = \alpha_{ckt} + \alpha_{kt}^{d} - \sigma \ln\left(\tau_{ckt}^{d}\right)$, where $\alpha_{ckt} = (1-\sigma)\ln\left(p_{ckt}\lambda_{ckt}^{-1}\right)$ and $\alpha_{kt}^{d} = b_{kt}^{d}\sum_{l}\frac{E_{t}^{d,l}}{P_{t}^{d,l}}$.

In principal, one could estimate (A8) for every product and time period, using vectors of origin and destination fixed effects to capture the supply characteristics α_{ckt} and demand characteristics α_{kt}^{d}. After netting off the demand and trade cost components, one collects a vector of fixed effects α_{ckt} and expressing over time changes for a given country x product, we have shocks to exporter supply.

We do not pursue this approach for two reasons. One, this decomposition requires that demand and supply shocks be linearly separable, which is not the case if destination d has unusually strong preferences for the output of origin c, i.e. if b_{ckt}^{d} varies over exporters c. Our data for Denmark show that the idiosyncratic component to demand is a salient feature of the data. That is, two Danish firms within the same industry purchase very different input bundles. Adding the additional dimensions of variation across importing country and including differences in cross industry composition likely make this idiosyncratic component to demand even more important. Two, while Redding and Venables (2004) employs aggregate bilateral trade in a single cross section, we need the supply characteristics for each exporting country by HS6 product (over 5000 in total) in each year. The typical exporter ships an HS6 product to a small and changing number of destinations each year (i.e. the large majority of bilateral trade flows are zero for a given HS6 product). This means that the average conditional value of bilateral trade, α_{ckt}, is extremely sensitive to variation in the number of destinations, itself endogenous to the time varying supply and demand characteristics of interest. Tackling these two issues is beyond the scope of our paper and so we employ a different approach.

As we discussed in the text, our approach is based on the variable WES (world export supply). Using (A8) we can derive the expression for WES by summing import demands over all destinations worldwide other than Denmark,

Online Appendix for "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

$$WES_{ckt} = \sum\nolimits_{d \neq Denmark} m_{ckt}^{d} = \underbrace{(1-\alpha-\beta)}_{\text{Constant}} \underbrace{(p_{ckt}^{1-\sigma}\lambda_{ckt}^{\sigma-1})}_{\text{c-k Supply char.}} \underbrace{\sum\nolimits_{d \neq Denmark} (\tau_{ckt}^{d})^{-\sigma} \sum\nolimits_{l} \frac{b_{ck}^{d,l} E_{t}^{d,l}}{P_{t}^{d,l}}}_{\text{Rest of World Trans. Cost and Demand Char.}}, \tag{A9}$$

WES can change over time because of changes in supply, trade costs, or demand. (A9) and (A6) imply that:

$$\log m_{ckt}^{j} = \log WES_{ckt} - \sigma \ln \tau_{ckt} + \underbrace{\ln(\frac{b_{ck}^{j} E_{t}^{j}}{P_{t}^{j}})}_{\text{Firm-j specific variables.}} - \underbrace{\ln\{\sum\nolimits_{d \neq Denmark} (\tau_{ckt}^{d})^{-\sigma} \sum\nolimits_{l} \frac{b_{ck}^{d,l} E_{t}^{d,l}}{P_{t}^{d,l}}\}}_{\textit{unobserved}}. \tag{A10}$$

Equation (A10) motivates the specification of our first-stage IV regression. Using firm-j characteristics as controls we regress firm j's imports of inputs on WES and transportation costs, hoping to capture movements in supply characteristics, $p_{ckt}^{1-\sigma}\lambda_{ckt}^{\sigma-1}$, that are invariant to destinations and give rise to changes in exports for exporter c and product k over time, and movements in τ_{ckt}. (When focusing on instruments for exporting at the firm level, we follow a similar strategy. That is, we sum imports over all sources for a given destination to capture movements in demand that are invariant to sources and give rise to changes in imports for a given importer c and product k over time.)

We can now use the explicit form of the imports expression to describe threats to identification. The main concern for our instruments is that a worldwide shock to demand for product k will simultaneously affect world export supply and input demand for Danish firm j; i.e. in equation (A10), the unobserved variable is correlated with WES_{ckt} and the firm-j specific variables. Further, if there is overlap between the inputs that firm j uses and the products that it sells, this demand shock also affects the desirability of firm j's products and therefore labor demand, and the wage. For example, rising demand for consumer electronics affects both cell phone manufacturers and producers of memory chips, and potentially the wages of workers employed by cell phone makers. We have addressed these concerns in the text. To recap:

1. We experiment with omitting industries with obvious shocks to demand occurring in this period (housing, electronics) and obtain similar results.
2. We incorporate an industry-time fixed effect in all the wage specifications. Any time varying shock to demand at the industry level (e.g. electronics) is absorbed by the fixed effect, leaving only shocks that are idiosyncratic to firms.
3. We incorporate firm level exports, instrumented by world import demand in the wage regressions. This variable is used specifically to address the possibility that there might be idiosyncratic shocks to product demand affecting the wage equation. World import demand, constructed in a manner

symmetric to world export supply, captures movements in demand in the destination country that are invariant to sources. If there is a common shock to demand, it will affect WID, and exports. This control for shocks to demand is significant in the Danish context, where the average firm exports 45% of its output.

4. We experiment with excluding from our estimates any exporter-products where Denmark represents a large share of worldwide demand for the product. Because Denmark is small, we obtain similar results.

Online Appendix for "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

Figure A1 Fuel prices over time

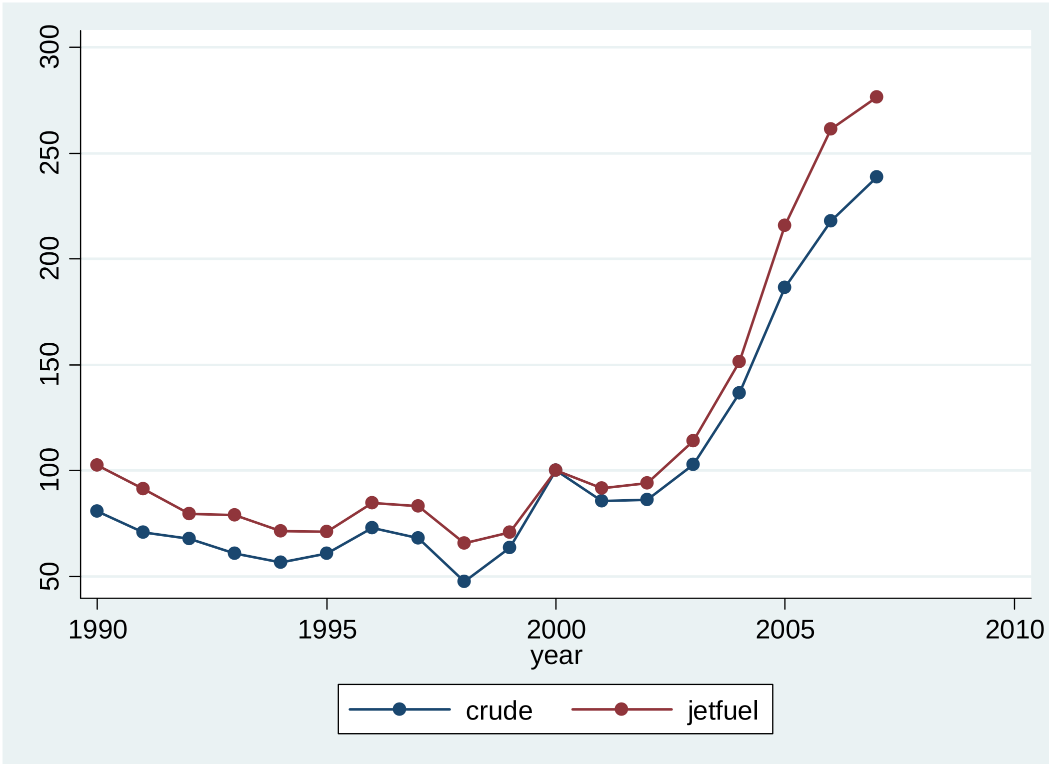
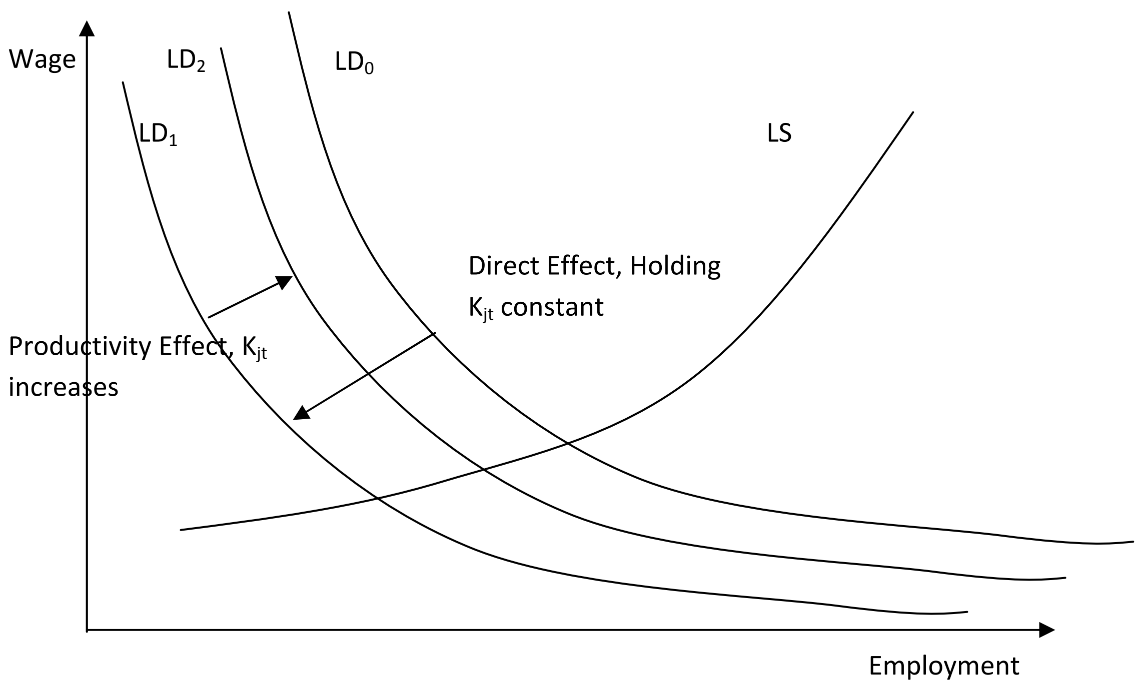


Figure A2. The Effects of Offshoring on Unskilled Wage



Online Appendix for “The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data”, by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

Table A1 The Top 20 HS6 Products in HS 84 (Machinery)

HS6	Description	Product Share	Cumul Share
848340	GEARS; BALL OR ROLLER SCREWS; GEAR BOXES, ETC	9.8%	9.8%
841391	PARTS OF PUMPS FOR LIQUIDS	8.8%	18.6%
848180	TAPS COCKS ETC F PIPE VAT INC THERMO CONTROL NESOI	6.3%	24.8%
840999	SPARK-IGNITION RECIPROCATING INT COM PISTN ENG PTS	5.3%	30.1%
848190	PTS F TAPS ETC F PIPE VAT INC PRESS & THERMO CNTRL	4.3%	34.4%
841290	ENGINE AND MOTOR PARTS, NESOI	3.2%	37.6%
840810	MARINE COMPRESS-IGNIN COMBUSTION PISTON ENGINE ETC	2.2%	39.8%
841370	CENTRIFUGAL PUMPS, NESOI	2.2%	41.9%
841899	REFRIGERATOR FREEZER AND HEAT PUMP PARTS NESOI	1.8%	43.7%
848210	BALL BEARINGS	1.8%	45.5%
848120	VALVES F OLEOHYDRAULIC OR PNEUMATIC TRANSMISSIONS	1.5%	47.0%
843390	PARTS FOR HARVESTER, GRASS MOWERS, SORTING EGG ETC	1.5%	48.5%
847990	PTS OF MACH/MECHNCL APPL W INDVDUL FUNCTION NESOI	1.4%	49.9%
843890	PARTS OF MACH OF CH 84, NESOI,IND PREP FOOD,DRINK	1.4%	51.3%
844900	MACH F MANUF OR FINISH NONWOVENS;HAT BLOCKS; PARTS	1.4%	52.7%
843149	PARTS AND ATTACHMENTS NESOI FOR DERRICKS ETC.	1.3%	54.0%
847330	PARTS & ACCESSORIES FOR ADP MACHINES & UNITS	1.2%	55.3%
847989	MACH & MECHANICAL APPL W INDIVIDUAL FUNCTION NESOI	1.2%	56.5%
841430	COMPRESSORS USED IN REFRIGERATING EQUIPMENT	1.2%	57.6%
848590	MACHINE PARTS WITH NO ELECTRIC FEATURES NESOI	1.1%	58.8%

Online Appendix for "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

Table A2 Summary Statistics for the Full Sample and the Estimation Sample

	Full sample			*Estimation sample*		
	Obs	Mean	Std. dev.	Obs	Mean	Std. dev.
Hourly wage	2,800,537	189.36	123.50	1,950,896	192.85	70.19
Log hourly wage	2,797,956	5.14	0.41	1,950,896	5.19	0.31
Log gross output	2,797,414	20.10	1.88	1,950,896	20.50	1.69
Log employment	2,800,537	6.06	1.69	1,950,896	6.44	1.49
Log capital per worker	2,779,424	12.53	0.94	1,950,896	12.59	0.89
High-skill	2,800,530	0.18	0.14	1,950,896	0.19	0.14
Experience	2,739,597	17.02	9.96	1,950,896	17.93	9.31
Union	2,791,622	0.83	0.37	1,950,896	0.88	0.33
Married	2,739,597	0.56	0.50	1,950,896	0.59	0.49

Notes: The full sample is the collection of workers and firms in manufacturing we have before we implement our sample-selection criteria. The estimation sample is the one we use in our paper, and the summary statistics for the estimations sample are identical to those reported in Table 1. All variables are calculated over the distribution of worker-year observations, which means firm characteristics such as output are repeated as many times as there are worker-years for that firm.

Online Appendix for “The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data”, by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

Table A3 Employment, Output and Import-Value Shares by Trade and Employment Categories

	Trade categories			
Number of Employees	Offshoring and exporting	Exporting only	Offshoring only	
	% in employment			
0-50	9.9%	6.7%	0.6%	
50+	76.7%	4.7%	1.4%	100%
	% of output			
0-50	9.5%	4.7%	0.4%	
50+	80.5%	3.6%	1.2%	100%
	% of imports			
0-50	12.6%	0	0.3%	
50+	86.7%	0	0.4%	100%

Notes: The sample consists of all the trading firms in the full sample. Trading firms either import/offshore, export, or both. The full sample is the same as in Table A1; i.e. the collection of firms we have before we implement our sample-selection criteria.

Table A4 Wage Regressions with Offshoring-Only Firms and for the Balanced-Panel Sample

	Offshoring-Only Firms				Balanced Panel			
Dependent Variable:	log hourly wage		log hourly wage		log hourly wage		log labor income	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log offshoring	-0.0138**	-0.0135**	-0.0355***	-0.0362***	-0.0179**	-0.0237***	-0.0354***	-0.0340***
	[-2.36]	[-2.32]	[-5.71]	[-5.69]	[-2.27]	[-3.32]	[-4.72]	[-4.98]
Log offshoring x high skilled	0.0526***	0.0530***	0.0462***	0.0462***	0.0732***	0.0734***	0.0494***	0.0552***
	[11.70]	[11.91]	[10.49]	[10.62]	[6.99]	[7.07]	[5.09]	[5.55]
Log exports					0.1019***	0.0812***	0.0875***	0.0991***
					[5.11]	[6.85]	[4.22]	[7.90]
Log exports x high skilled					-0.0337**	-0.0351**	0.0094	0.0049
					[-2.07]	[-2.11]	[0.62]	[0.30]
Log (exports + 1)	0.0010		0.0005					
	[1.57]		[0.80]					
Log (exports + 1) x high skilled	0.0008		0.0006					
	[1.55]		[1.07]					
Share, exports/output		-0.0001**		0.0005***				
		[-2.57]		[5.57]				
Firm Controls	No	No	Yes	Yes	Yes	No	Yes	No
Observations	1,976,883	1,976,883	1,976,883	1,976,883	1,124,449	1,124,449	1,124,443	1,124,443
Number of job spell fixed effects	388,575	388,575	388,575	388,575	191,653	191,653	191,653	191,653
R-squared	0.151	0.1509	0.153	0.153	0.1689	0.1669	0.1052	0.1036

Notes: Clustered (firm-year) t-statistics in brackets. *** p<0.01, ** p<0.05, * p<0.1. In columns (1)-(4) the sample is our main estimation sample plus the offshoring-only firms. The variables log (exports + 1), log(exports + 1) x high-skilled, and exports/output are not instrumented. Columns (5)-(8) are the same specification as Table 5, but the sample is a balanced panel of firms that are in the sample for the entire period.

Online Appendix for "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", by David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, September 2013.

Table A5 Additional Robustness Exercises

Dependent variable:	Log Hourly Wage		Log Hourly Wage	
Robustness check:	I. Broad offshoring		II. DK-dominant flows removed (1%)	
	FE-IV		FE-IV	
Log(offshoring)	-0.1064***	-0.0608**	-0.0179	-0.0187
	[-2.58]	[-2.31]	[-2.59]	[-3.50]
Log(offshoring) x high-skilled	0.1311***	0.1322***	0.0452***	0.0462***
	[8.93]	[9.32]	[7.32]	[7.35]
Log(exports)	0.0720***	0.0742***	0.0405***	0.0437***
	[6.32]	[5.11]	[2.88]	[4.47]
Log(exports) x high-skilled	-0.0597***	-0.0567***	0.0103	0.0123
	[-4.05]	[-3.81]	[1.03]	[1.17]
First stage IV *F*-statistics:				
log offshoring	2.55	9.03	5.20	9.70
... *x* high skill	33.79	31.93	28.10	21.80
log exports	6.08	10.97	3.95	7.75
... *x* high skill	16.29	15.26	17.26	14.38
Other firm-level controls	Yes	No	Yes	No
Obs	1,950,896	1,950,896	1,928,599	1,928,599
No. job spells	384,257	384,257	383,035	383,035
R2	0.1543	0.1525	0.155	0.1533

Notes: Table A5 presents the results from worker-level Mincer regressions, using log hourly wage as the dependent variable. All specifications include job spell, industry-year and regional fixed effects. Log offshoring, log exports and their skill interactions are instrumented using world export supply (WES), world import demand (WID) and transport costs. Robust T-statistics in brackets. Standard errors clustered at firm-year levels. *** p<0.01, ** p<0.05, * p<0.1.