Literature review

* Industry 4.0 = AI, big data, iot, mass production, reduction of human involvement , connectivity of machines
* Industry 5.0 = reintegration of human creativity, mass tailored consumerism, robotic management

Acemoglu & Autor (2010)

* Canonical model and also cross country wise in advanced countries good, but has drawbacks discussed in paper

1. Skill != task
2. Skill bias is not always the same

* Job polarization (low high)
* Skill bias produce higher demand for high skilled labour
* Skill for wages
* Technologies substituting or complementary to human (that varies over time)
* Reallocation of skills across occupations
* Increasing divergence of return to education between 3rd and 2nd educ
* Skilled/unskilled -> 3rd/2nd educ definition used!! (p.7)
* fixed-weighted averages of the relevant sub-group means to capture heterogenous group differences (p.7)
* historic overview of returns to education
* polarization = simultaneous increase of low and high skill wages but not in middle range
* Employment rose for lowest three wages deciles, decreased in 4-8 and stagnated in the top decile
* managerial, professional and technical occupations = highly-educated & highly-paid occupations
* More affordable ICT cause for job polarization
* s partly endogenous responses to changes in the supplies of
* Nordhaus (2007) describes steep fall of costs for computation (60 6o 75 % annually)
* Computers are good at routine tasks
* Substation of middle skill occupations vs complementation in high skill occupations since analytical jobs depend heavily on information which can be more easily gathered with computers
* Non routine cognitive tasks (abstract) are mostly used in managerial, professional and technical occupations
* magnitudes of differences, task measures are standardized to have a mean of zero and a cross-occupation standard deviation of one
* Routine task intensity measure: routine cognitive, routine manual; Task composition is splitted into coarser categories (substi: blue collar = routine manual, sales = routine coginitive, Non substi: care = manual)
* declining middle-skill employment on the female occupational distribution is distinct -> women more affected since tendency to routine cognitive task whereas males tend to routine manual tasks
* loss in middle occupation was for women offset by high skill job by 85 % and for men 55%
* women were more likely to attain different job esp. highly educated
* heterogeneity of polarization of occupation in EU
* industry sectors has shifted and thus different demand for tasks and not because of different composition of jobs
* share of employment in sector changes either by shifting sectors (industry composition changes) or due to within sector happenings
* Figure with employment change of sector splitted into 10 categories (Fig.12, p.30)
* Before 1979 cross industry effects prevailed, since then within industry effects
* Industries changes within toward more service oriented tasks, in history more occupation changes caused by structural (cross) industry changes
* Increasing importance of occupation to determine wage rather than education
* METHOD: years of completed schooling; dummy variables for highest completed educational category, dummy variables 10 occupational categories used above (Table 2, dummies 11 industry categories used in Table 6. For each set of regressors, calculate partial Ρ2 value
* e explanatory power of the dummies rises substantially more than does the linear term, reflecting **the convexification of the return to education**
* ran several OLS regression wage = **X** (X = either educ level, occupation dummy, industry dummy..)
* occupation plays an increasingly important role in the evolution of employment and (here) earnings; it is not simply a proxy for either education or industry
* perform this benchmark by comparing the partial Ρ2 value: task dummies have more explanatory power than occupation
* demand for skills increases over time because changes in technology are assumed to be “skill biased,” and are more complementary to high skill workers
* **canonical model**: two types of worker, substitution between them constant (inherently wrong since they are imperfect substitutable) & competitive labour markets
* high and low skill workers with distributed efficiency units (not everyone uses their skills in the same manner)
* technology in canonical model as factor augmenting either, or both multiplicated to input of high or low skill worker
* 3 cases:
  1. 1 good, workers imperfect substitutes
  2. 2 goods, each good produced by one skill (either high skill product or low skill product)
  3. 2 sectors producing many goods (imperf. substi.) and workers are emplyed in both sectors
* technological improvements of any sort will lead to higher wages for both skill groups in the canonical model
* high skill technology increases skill premium (p.39), low skill technology decreases skill premium
* college wage premium for workers with under 20 years of experience is quite responsive
* n significant changes in inequality among workers with the same education
* wage variation between workers within a group (high or low skil workers) -> for my research maybe not all highly educated person (women, migration) benefit in the same extent of technology improvement
* no longer the college premium! (UNCERTAINTY), since not all college workers have high skill and not all non-college workers have low skill.
* Return to educ more dispersed among high educ individuals p44
* Canonical model best if technolgy change is stead (log) linear change
* technical change imply that technology should become more skill biased following increases in the supply of high skill workers -> technical change might be partly a response to the steady increase in the supply of skills (improved technology just possible due to higher skilled people)
* **computer technology** is particularly **well suited for automating routine** tasks
* canonical model drawbacks:
  1. does not explain WHY earnings decrease in low skill jobs
  2. does not provide framework for the employment polarization
  3. does not distinguish between skill, task and employment
  4. silent about importance of occupations as predictors of earnings have increased over time. (p.45)
  5. does not consider that new technologies are labor replacing instead focus on factor augmenting characteristics
  6. treats technology as exogenous (its likely to be endogenous)
  7. does not consider offshoring and outsourcing
* a task is a unit of work activity that produces output. A skill is a worker’s endowment of capabilities for performing various tasks
* skills are used in tasks and tasks are output producing in exchange for wages
* a given skill level can potentially perform a variety of tasks and, moreover, can change the set of tasks that they perform in response to changes in supplies or technology.
* At least three skill groups
* e substitution of skills across tasks.
* endogenous technology relative demand curves can be upward sloping rather than downward sloping. contrasts with the necessarily downward sloping relative demand for skills in the canonical model and Ricardian model studied. If the induced response of technology is sufficiently strong to make the endogenous relative demand curves upward sloping, then the increase in the skill premium a response to the large increase in the supply of skills
* Riccardian model the response of the economy to any **increase in the supply of a factor** will be to undergo an **endogenous change in technology that weakly increases demand for that** factor => increase in a supply factor creates endogenously an increase in demand for that factor == wage will increase
* Proposition 8 Under regularity conditions an increase in the supply of factor φ (for φ ∈ {Λ; Μ; Η}) will induce technical change strongly biased towards that factor–thus increasing the wage of that factor–if and only if the aggregate production possibilities set of the economy is locally nonconvex in factor φ and technology ∝.
* local nonconvexity condition implies, loosely, that if we double both the supply of factor φ and the quality or quantity of technology ∝, output will more than double
* high skill workers bias technology in their favour to make themself even more productive and hence replace demand for middle skilled workers (p.75)
* abstract, routine and manual intensive tasks (gammy 1-3) be the employment shares. S denotes gender, e denotes education group, j denotes age group, and k denotes region of the country
* demographic groups as skill groups, and the gamma parameters as reflecting their patterns of comparative advantage in 1959.
* EMPIRICAL APPROACH: p 77
* systematic, non-monotone changes in the distribution of employment across occupations of various skill levels
* similar technological forces have altered occupational structures comparable over countries, labor market (wage schedules) far from identical no satisfactory understanding of root. adoption of new technologies either replacing or complementing workers in certain tasks requires investments and incentives for adopting technologies are affected by existing regulations. possible that firms select different technologies in different countries

Bittarello Kramarz (2018)

* n increase in the supply of skilled labor affects task assignment within and between occupations.
* We find **higher average education**al associated **with more routine,** fewer cognitive / social tasks within occupations and **with fewer routine**, more cognitive and more social tasks across occupations. -> two effects of higher education
* Tasks are the building blocks of production.
* Three main points: , job content is heterogeneous within occupations, , university graduates hold a comparative advantage cognitive duties e skill premium fell with the expansion of higher education
* schooling, migration and labor supply are endogenous, we project the graduate share on the basis of previous surveys to construct instruments. -> Which jobs creates certain demands, higher educ supply change occupation(?)
* market conditions influence task assignment within occupations
* vars used: : female; married; foreign born; age and age squared; tenure and tenure squared; multiple jobs; part-time job; fixed-term contract; and civil servant
* fixed effects for education,2 occupation and region of residence.
* Dictionary of Occupational Titles and the O\*NET -> task to job description
* routine, cognitive and social -> three classifications of tasks
  1. Routine tasks denote a lack of autonomy or a submission to machinery.
  2. Cognitive tasks involve decision making.
  3. Social tasks require interaction with clients or the public
* graduate share among employed workers
* e skill premium declined
* Routine activities became common for both groups, but the proportional increase was larger for graduates. (this seems relative but from which base level off??)
* Job content is heterogenous within occupation
* routine tasks are more frequent in the bottom of the wage distribution by our measure, whereas social and especially cognitive tasks are more common in the top. Third, this pattern has become looser over time.
* Negative relationship found between share of university graduates and skill premium (in France between 1990-2010) -> higher digitalisation wasn’t a thing back then!! TIME UNTIL TECHNOLOGY GETS BROADLY USED
* I DON’T FIND THE PAPER THAT CONVINCING…..
* P. 13 how to determine composition effect / substitution effect with a regression
* explore variation in job content within occupations in greater depth
* occupations evolve -> take care of that if I choose 10 years or longer!!

**Huber Lechner Wunsch (2012)**

* semiparametric methods have advantage in contrast to parametric ones that covariates can be added in a more flexible way
* PS = probability of being observed in one of two subsamples conditional on the covariates
* propensity score is based on a parametric model, but the relationship between the outcome variables and the propensity score is non-parametric
* 4 classes of estimators: parametric estimators (OLS, probit, DR), inverse probability weighting estimators, direct matching estimators and kernel matching
* kernel-matching based on local regressions with finite sample adjustments (local ridge regression) performs best vs
* inverse probability weighting (IPW) has the best properties (when using normalized weights for estimation)
* two branches: asymptotic properties on small sample properties AND Monte Carlo Evidence
* Lechner et al do a Empirical Monte Carlo Study
* ‘placebo treatments’ among the non-treated
* we exploit the actual dependence of the outcome of interest on the covariates on which selection is based in the data rather than making assumptions on this relation when specifying the data generating process.
* combining matching with weighted regression
* relevance of trimming to improve the finite sample properties of all estimators, trimming those obs which reveive a “too” large weight
* bias-adjusted radius (or calliper) matching estimators perform best on average
* parametric approaches perform almost as well in smaller sample bc of precision gain compensates for their larger bias, and parametric approaches become even better if the sample becomes large
* Coarse instruction for PSM procedure:
  1. estimate the propensity score in data and use as true propensity score for the simulations.
  2. draw a sample of control observations from actual (large) data, simulate a (placebo-) treatment for this draw, estimate the effects with the different estimators for this sample. By definition, the true effect of this treatment is zero.
  3. repeat the second step many times to evaluate the performance of the estimators.
* In step 2 they chose sample sizes with 300, 1200, and 4800N vs 114 Control N’s! If they are with this scale close to its asymptotic distribution, we expect it to perform similarly or even better for larger sample sizes
* ATET: expected potential outcome for treated under treatment vs treated under no treatment = high skill but not in cognitive non-routine tasks
* assumption required to interpret θ as a causal parameter is called either unconfoundedness, the conditional independence assumption (CIA) or selection on observed variables
* to obtain a causal effect of ATET many covariates required to make CIA plausible
* Other PS based covariate adjustments**: instrumental variable estimator** proposed by Frölich (2007a), the **decomposition-type of approach** suggested by DiNardo et al. (1996) and **semiparametric versions of the difference-in-difference estim**ator (e.g., Abadie, 2005; Blundell et al., 2004; Lechner, 2010).
* Reweighting is required to make the non-treated comparable to the treated (Done by distance measuremets -> giving weight to those which are close)
* IPW
  1. This estimator directly reweights the non-treated outcomes to control for differences in the propensity scores
  2. parametric propensity score is used, inference for IPW is straightforward, because one could either rely on the GMM methodology (Hansen, 1982), or on the bootstrap
  3. IPW is attractive because it is computationally easy, fast, and probably close to being asymptotically efficient and does not rely on any tuning parameters
  4. IPW may be sensitive to large values of pˆ(x) that might lead to fat tails in its distribution
  5. PS does not guarantee ATET but inverse probability tilting restores this causal interpretation, Graham et al (2011) show that this estimator is locally efficient
  6. s exact balancing property of the propensity scores = key difference for IPW estimator
* Direct matching
  1. Pair matching (1:1)Unfortunate that all other controls receive a weight of 0 -> loss of efficiency, reduce bias since only closest is used, More robust to ps misspecifications
* 1:m
  1. 1:m -> no ideal way of how to choose m
  2. 1:m by radius / calliper, more efficient and less biased as 1:m or 1:1 since in the latter distance is not defined thus also unappropiate control could be the closest one still
  3. Abadie and Imbens (2006) show that for a 1 : M matching estimator (directly on X) non-parametric regression can be used to remove the bias from the asymptotic distribution that may occur when X is more than one-dimensional
  4. a distance metric that not only includes the propensity score, but in addition those covariates that are particularly good predictors of the outcome (in addition to the treatment) -> MAHALONOBIS
  5. Lechner 2011: combines the features of radius matching with additional predictors and linear or non-linear regression adjustment
  6. distance-weighted radius matching (which could be interpreted as kernel matching
* improving naïve propensity score matching is to use a distance metric that not only includes the propensity score, but in addition those covariates that are particularly good predictors of the outcome (in addition to the treatment
* combines the features of radius matching with additional predictors and linear or non-linear regression adjustment.
* Propensity score kernel matching:
  1. kernel matching estimators and found the estimator that is based on ridge regressions to have the best finite sample properties
  2. Ridge regression may be considered an extension to local linear kernel regression -> ridge superior in terms of boundary bias, but is prone to variance problems entailing rugged regression curves when data are sparse or clustered
  3. Binary outcome variable has no ridge but local logit exist, bootstrap usually applied for that
* Parametric models:
  1. Logit or linear model
  2. Using PS covariates directly in regression
  3. Parametric estimators have the advantage that they are very easy to compute, their asymptotic properties are well known, inference procedures are known and reliable, they are efficient if correctly specified, do not depend on tuning parameters
  4. disadvantage is their sensitivity to the correct specification of the models involved
* TRIMMING:
  + From Eq. (1) we see that all estimators can be written as the mean outcome of the treated minus the weighted outcome of the non-treated observations
  + If particular values of p(x) are rare among the controls and common among the treated, such control observations receive a very large weight in all estimators of the ATET -> problem if treated values have no counterpart in control group
  + Semiparametric estimators where all treated have a p(x)=0,99 and only one non-treated has the same ps then the one nontreated will receive a weight of 1 and all the other non treated units have a weight of almost zero. This results in estimators with an infinite variance bc they are based on the mean of effectively only one observation. -> problem can be solved if adding more obs with non treated display the same ps.
  + removing such observations comes at the cost of incurring potential asymptotic bias s, if it does not disappear fast enough with increasing sample size
  + When **treated observations are removed** based on a fixed cut-off value of p(x), the population underlying the definition of the ATET changes.
  + When **control observations are removed**, we may not be able to reweight the controls successfully towards the distribution of the covariates observed for the treated
  + Three step correction:
    - * 1. Control weights and set them to zero if share of the sum of all weights is larger than t%
        2. Remaining weights are normalized
        3. Remove treated observations with a value of p(x) larger then smallest value of p(x) among the controls removed in the first step to avoid a severely unbalanced sample induced by trimming
* estimators, whether they are parametric or semiparametric, are treated exactly the same way: control observations are removed if their IPW weights are above the threshold and the treated sample is adjusted accordingly to enforce common support of the propensity scores in finite samples.
* Common support issues
* Trimming rule vanishes in large samples
* Thin-support regions. Occurs when one of the covariates has infinite support
* trimming changes the finite samples properties only, because there is no asymptotic support problem
* for t% they use the weight from the IPW estimator

Inverse probability weighting and tilting

* no tuning parameterst o choose
* IPW weights are normalized after trimming -> otherwise wont add up to one in the trimmed controll sample
* IPT propensity scores (based on the method of moments) are estimated after trimming such that the momentas are balanced wrt trimmed treated sample

RESULTS:

* conclusions stem from root mean squared error analysis
* specification problem is relevant as the variables are jointly highly significant in the propensity score as well as in the outcome equations based on Wald-statistics

Direct matching

* Following psm estimators: pair-matching, radius matching and radius matching with linear and non-linear post-matching regressions
* Common ground between all: measure the distance between observations
* Also consider linear index, which matter since this monotone transformation may matter at the boundary of the propensity score where the cdf is highliy non linear (bordercases behave often differently)
* Ps supplemented by covariates, covariates looked at that they are jointly significant in outcome variable and both outcome equations based on Wald tests
* Radius matching: defining radius in terms oft he largest distance calculated from pair-matching (using multiples and shares oft hat distance). If radius remains empty use nearest neighbor
* Ps + two covariates were always superior at radius matching estimators

Kernel matching:

* Ridge regression matching: like linear regression but tries to prevent overfitting. Used to produce ps and afterward again for computing ATE
* Bandwidth can be varied
* They use Silverman (1986) type rule of thumb fort he Epanechnikov kernel

Parametric models:

* Two versions:
  + 1. Just applied to non treated (for treated simply their sample average outcome is computed)
    2. Separate parametric model for the treated gets conducted
* Linear regression, heckit, tobit model, probit, flexible datadriven OLS and probit estimation that selectively adds higher order and interactoion terms to chose the optimal model wrt to correcte AIC
* Table contain the coefficient oft he regression results for the root mean squared error
* RMSE increases in the strength of selection and the sources appear to be both the bias and the precision of the estimators.
* Balanced sample leads tot he lowest RMSE. In particular fort he sample with very few control observations. Significant small sample bias for all types of estimators

Functional missepcification of the propensity score:

* Functional misspecification which leads to an inconsistent estimation of the propensity score leads to an increase of the bias and to a reduction oft he variance. In smaller samples gain in precision. Larger sample bias dominates
* moving from no trimming to the 6% trimming rule leads to a considerable reduction in the RMSE

trimming seems to be very effective in cases where it is most needed, while it does not hurt much in the other scenarios.

Estimator specific issues:

* Medium sized sample IPW > IPT since IPW has better bias and variance properties, for small and large sample differences vanish
* Direct matching: NN matching not successfull in terms of RMSE, BUT as n>>> better bias properties become more important -> but still dominated by other matching methods
* Radius matching: smaller radius less bias but more variance, if n>>> regression adjustment becomes more attractive
* Logit adjustment > linear regression adjustment at least for smaller samples.
* Mahalanobis matching: inclusion of additional covariates in Maha matching variance get reduced but bias increases
* Kernel matching: largest bandwidth preferred way currently, BINARY: local logit or local linear regression -> local logit performs better in larger samples but in general local linear regressions dominate over all
* Parametric models:
  + standard probit & OLS preferred in terms of RMSE
  + other parametric models: tobit, heckit, worse performance in smaller/medium samples due to their comparably large variances -> especially when share of treated is high.

DR: less attractive due to larger variability compared to standard regression

Flexible estimation based on corrected AIC also worse RMSE

* large values of the kurtosis are also a good indicator of estimator instability leading to important outliers

Winners

* estimators have to be in the best group in at least half of the cases and never be in the worst group. Among that group, we choose the best estimators in terms of average RMSE -> according to this procedure the *regression-adjusted radius matching with additional predictors based on the linear index and using the large radius.* (either logit or linear dependent on y)
* Simple pair matching based on Ps
* Kernel matching not clear thus they used smallest and largest bandwidth
* Local linear > local logit
* Parametric: non weighted OLS and probit, followed by their DR versions
* Trimming important among correctly specified but also among misspecified ones (except IPT did not profit from it in terms of bias, precision, skewness and kurtosis)
* Gains from trimming come mainly from the DGP’s with strong selection and few controls
* gains are probably larger for the correctly specified model because the propensity score of this model contains additional interaction terms that should lead to a ‘better’ individual prediction. Since such a prediction is likely to increase the (unconditional) variance of the propensity score, it becomes more likely that the weights are above the threshold
* Bias worry of estimator: could not be confirmed, especially not for small sample and not worsened for the others
* Comparing the estimators to each other shows that most appear to lie within a reasonable distance to the respective best estimator, with the exception of pair matching, which is never competitive in terms of the RMSE due to its large variance
* Moreover, when the propensity score is correctly specified, IPT does often worse than pair matching
* For the case of a correctly specified model, probit and OLS appear to be the best estimators in terms of RMSE, while for the misspecified propensity score, logit and OLS adjusted radius matching are best.
* Distributional properties oft he estimators are dependent on the outcome considered
* binary employment outcome, the best performing logit adjusted radius matching and the probit estimators also have ‘good’ higher order moments

Acemoglu Restrepo (2019)

* production requires task (only production or also other jobs??)
* new technologies increase productivity and impact the allocation of tasks to these factors of production (labor and capital)

Goos, Manning and Salomons (2009a)

* Share of employment in Middle wage jobs decreased in all 16 EU countries (1993-2007)
* Skill-biased technological change (more by Autor, Katz, 1999) -> augmented that middle skill is decreasing and low and high skill is increasing
* Define tasks as abstract, routine and service tasks and collapse it then into one single routine intensity index

Autor & Handel (2009)

* Tasks differs between and within occupations
* Task are determining for wage differences
* Autor levy & Murnane (2003) posed hypothesis of polarization of jobs can be attributed to computerization

Mihaylov & Tijdens (2019)

* Routine intensity indicator: RTI = R – NR
* Table of five measures of tasks
* Address potential problems of misclassified tasks
  + 1. Routine task to wrong routine task
  + 2. Non routine task to wrong non routine task
  + 3. Routine task to non routine task / vica versa // less likely
* Review different task measures (Acemoglu & autor, frey & osborne
* Describing procedure how to arrive at 5 classifications

J.G. MacKinnon, M.Ø. Nielsen and M.D. Webb HOW TO CLUSTER

* Thus, before specifying any clustering structure, we need to think about how intra-cluster correlations may arise and why independence across clusters may, or may not, be plausible for that structure
* fine clusters within each coarse cluster are just the individual observations. Under the assumption of fine clustering, the terms on the right-hand side of (16) are all equal to zero. Under the assumption of coarse clustering, however, at least some of them are non-zero, and (16) must therefore be estimated. If we cluster at the fine level when coarse clustering is appropriate, the CRVE is inconsistent. On the other hand, if we cluster at the coarse level when fine clustering is appropriate, the CRVE has to estimate (16) even though it is actually zero. This makes the CRVE less efficient than it should be, leading to loss of power, or, equivalently, to confidence intervals that are unnecessarily long, especially when the number of coarse clusters is small.

Cameron & Miller CLUSTER VCE VERY NICE

* e bias-variance trade-off that is common in many estimation problems – larger and fewer clusters have less bias but more variability
* consensus is to be conservative and avoid bias and use bigger and more aggregate clusters when possible

ALM (2003)

NOT PAPER

* Mcgee & Gauch (2020)
* FIR demands high skilled workers
* Relevance for competitivness of industry
* Replacement of jobs which are predictable physical and routine tasks
* Shiftment of resources and power
* Just in time learning
* tech literate workforce that can function in a multidisciplinary work environment
* FIR affects directly (emerging tech sector), indirectly (spillover effects in other sectors such as healthcare) and broad (income effect) (p.281)
* larger samples also contain the more problematic DGP’s with 10% and 90% treated. Furthermore, note that because specifications with incorrectly specified propensity scores are also included, they are not expected to be unbiased.
* Standard deviation is smaller if correctly specified -> especially for probit bias disappear -> probit looks good
* Performance of OLS suggests that the linear model is not flexible enough and thus misspecified
* Similar for kernel estimators but bias is smaller
* Standard deviations are approximately reduced by half when quadrupling the sample size -> more obvious for binary outcome
* Greater sample size increases relative differences in the RMSE of estimators
* Larger sample size probit dominates, 2nd place regression adjusted radius matching

# O\*net

Arntz (2018):

* Expectation-Maximization (EM) algorithm

Acemoglu & Autor (2011) Skills, Tasks and Technologies: Implications for Emplo

* Linking occupations with tasks
* Clustered jobs in four major groups: managerial:
  + professional and technical occupations (abstract non routine)
  + sales, clerical and administrative support occupations (routine cognitive)
  + production, craft, repair,and operative occupations (routine manual tasks)
  + service occupations (non routine manual tasks))
  + simultaneous growth of high and low-skill occupations, male educational attainment -> employed males who had lower education fell from 57 to 42 percent, while fraction with degree rose from 20 to 28 percent
* Routine tasks are characteristic of many middle skilled cognitive and manual jobs
* Process of automation and offshoring of routine tasks, in turn, raises relative demand for workers who can perform complementary non-routine tasks.
* Non routine = abstract or manual (interpersonal or just manual)
  + Cognitive : Abstract tasks are activities that require problem-solving, intuition, persuasion, and creativity, found in occupations such as professional, managerial, technical and creative occupations
  + Manual: situational adaptability, visual and language recognition
    - Interpersonal
    - Manual
* Abstract and non-routine are less susceptible to automation!!!!
* More managerial (NRI) and technical (NRC) while less RM, NRM
* More ICT more demand for highskill -> bigger difference in outcome if more ICT?
* shifts in industrial composition do not explain the observed polarization of employment across occupations.
* Tasks
  + RC = Importance of Repeating same tasks, importance of being exact or accurate, structured vs unstructured
  + RM = Pace Determined by Speed of Equipment, Controlling Machines and Processes, and Spend Time Making Repetitive Motions
  + NRA = Analyzing Data or Information, Thinking Creatively, and Interpreting the Meaning of Information for Others
  + NRI = Establishing and Maintaining Interpersonal Relationships, Guiding, Directing, and Motivating Subordinates, and Coaching and Developing Others
  + NRM = Feel Objects, Tools, or Controls, Manual Dexterity, and Spatial Orientation.
  + Skill biased technologies are expected after increase of skilled labor supply

**Educational upgrading, structural change and the task composition of jobs in Europe**

**(Hardy, Keister, Lewandowski, 2018)**

* transforming O\*net soc into isco
* estimate the task content of jobs, by mapping o\*net items to the corresponding occupations in SOC and then using official ILO crosswalk
* shift of isco88 ot isco 08 in 2011!! -> especially in farming workers and retail workers
* approximation of the general task intensity distribution across occupations

tasks and skills in europe

* We used the crosswalk available at the ILO website: <http://www.ilo.org/public/english/bureau/stat/isco/isco08/> WHERE I GOT ALL THE DOCS for crosswalks FROM