THE SUBSTITUTABILITY OF YOUNG AND OLD WORKERS

By

Abuzar Ali

Submitted to

Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of

Master of Arts in Economics

Supervisor: Professor Gábor Kézdi

Budapest, Hungary

2017

Abstract

In this study I estimate the elasticity of substitution between young and old workers in the U.S. using Constant Elasticity of Substitution (CES) models. The current CES models in the literature impose the same elasticity of substitution between any age group pairs. My contribution to the literature is that I develop CES models that in principle allow the elasticity of substitution between young and old workers to be different from the elasticity of substitution between other age groups. Estimating these models using 1996 to 2016 March CPS data yields estimates for the elasticity of substitution between young and old workers that are lower than the elasticity of substitution between arbitrary age group pairs, implying that young and old workers are less substitutable than other age group pairs. This result contradicts the conventional wisdom that greater labor force participation of old workers leads to less favorable labor market opportunities for youth. Another result of my study is that the substitutability of young and old workers increases within more homogenous skill groups.

Acknowledgements

This study would not have been possible without the proper guidance and constant encouragement of my advisor Professor Gábor Kézdi. You have the quality to bring out the best in others, a quality which I really appreciate and would very much like to have myself. I would also like to thank Professor Thomas Rooney who assisted me with the revisions of various drafts of my thesis. Your comments and suggestions have made the final version much more readable. Lastly, I want to thank my teachers, classmates, and the staff at the Economics Department of CEU. I learned a lot, improved as a person, and enjoyed my two years stay in Budapest in your supportive company.

Table of Contents

Abstract .... **i**

Acknowledgements .... **ii**

Table of Contents .... **iii**

List of Figures .... **v**

List of Tables .... **vi**

1 Introduction **1**

2 Literature Review **4**

3 Models and their Identification **9**

3.1 Model A 9

3.2 Model B 14

3.3 Model C 18

3.4 Comparison of the Models 21

4 Data and Variables **24**

5 Results **28**

5.1 Results for Model A 28

5.2 Results for Model B 30

5.3 Results for Model C 31

5.4 Summary of Results 32

6 Interpretation of Results **34**

6.1 Comparing Estimates of the Models 34

6.2 Interpretation of the Elasticity of Substitution between Young and Old 35

6.3 Relevance of the Elasticity of Substitution to Public Policy 39

7 Robustness Checks **43**

8 Conclusions **47**

Appendix **50**
Section A1. Derivation of Elasticity of Substitution for CES production function 50
Section A2. Data Appendix 53
Section A3. Tables 57
Section A4. Derivation of Standard Error for the Estimate of Elasticity of Substitution 72
Section A5. Three Step Procedure for Model B 74
Section A6. Robustness Checks for Model C 77
References **81**

List of Figures

Figure 1. Possible Shapes for Isoquant with Old (O) and Young (O) labor inputs 36
Figure 2. Young-Old Relative Demand and Substitution Elasticity..37
Figure 3. Graphical Interpretation of Elasticity of Substitution between young and old 38

Figure A 1. College-High School Relative Demand Curve...52

List of Tables

Table 1. Main Estimates of Model A 29
Table 2. Main Estimates of Model B 30
Table 3. Main Estimates of Model C 31
Table 4. Estimates and Their Confidence Intervals 33
Table 5. Estimates From All Models of this Study 47

Table A 1. Model A - College-HS log relative wages 57
Table A 2. Model A - College-HS log relative hours supplied 57
Table A 3. Model B - College-HS log relative wages 58
Table A 4. Model B - College-HS log relative hours supplied 59
Table A 5. Model C - Old-Young log relative wage 60
Table A 6. Model C - Old-Young log relative hours supplied 60
Table A 7. Model A – Estimates from First Stage 61
Table A 8. Model A - Estimates from Second Stage 62
Table A 9. Estimates for Model B 63
Table A 10. Estimates for Model C 64
Table A 11. Model A Robustness Check 1 - College-HS log relative wages 65
Table A 12. Model A Robustness Check 1 - College-HS log relative hours supplied 65
Table A 13. Estimates for Model A – Robustness Check 1 66
Table A 14. Model B Robustness Check 2 - College-HS log relative wages 67
Table A 15. Model B Robustness Check 2 - College-HS log relative hours supplied 67
Table A 16. Second Stage Estimates for Model B – Robustness Check 2 68
Table A 17. Third Stage Estimates for Model B – Robustness Check 2 69
Table A 18. Estimates for Model B - Robustness Check 3 70
Table A 19. Estimates from Robustness Checks 4 and 5 for Model C 71

1 Introduction

I estimate the elasticity of substitution between young and old workers using different Constant Elasticity of Substitution (CES) models. The elasticity of substitution between young and old workers shows the ease of substitutability of the two age groups in the production process of the whole economy and helps in answering two policy relevant questions. Does encouraging early (late) retirement for old workers create more (less) employment opportunities for young workers? Does providing tax and other monetary incentives to employers to retain old workers will lead to lower relative demand for young workers? These questions can be answered using atheoretical methods and program evaluation tools. But these are fundamentally questions related to the production process itself and the approach used in my thesis can provide answers that are grounded in economic theory.

CES type modeling has been widely used in the Immigration and Inequality literature where the elasticity of substitution between different age groups has secondary importance (Welch, 1979; Card and Lemieux, 2001; Borjas, 2003; Autor et al., 2008; Ottaviano and Peri, 2012). In these studies, an 'average' elasticity of substitution between age groups is estimated; the elasticity is the same between age groups close and far away from each other. Having same elasticity between all age group pairs is a restriction not easily justifiable. My contribution to the literature is that I enrich the CES type models used in the labor economics literature by allowing the elasticity of substitution between young and old workers to be different from the rest of the age group pairs.

I develop and compare three different CES models. I estimate these models using 1996 to 2016 March CPS data and obtain elasticity of substitution between young and old workers in the range of 3.4 to 5.7[1]. The usual range for the 'average' elasticity of substitution between age groups is between 5 to 10 (Card and Raphael, 2013, p. 32)[2]. This leads to my first conclusion: young and old workers are less substitutable than other age group pairs. In other words, as compared to other age group pairs, it is not easy for firms to replace young with old workers and vice versa. The second conclusion is that the substitution elasticity between young and old increases within more narrowly defined (or more homogenous) skill groups, a result which is not surprising and makes intuitive sense.

There are atheoretical and program evaluation studies which argue against the conventional wisdom that greater employment for old workers have large negative impacts on labor market opportunities for youth (Gruber and Wise, 2010; Munnell and Wu, 2012). The strength of these studies is that they exploit exogenous variations in variables (like employment) and as a result have higher internal validity. These studies are usually confined to the context of a particular time and/or place and look at results at more aggregated levels. In these studies extrapolating or comparing results from one country to another country for instance might be problematic because the distribution of education levels in age groups might differ considerably between countries. Therefore they have lower external validity.

The CES models that I use in my thesis allow different substitutability between young and old workers as compared to other age groups. As a result the elasticity estimates from these models can help critically analyze the conventional wisdom mentioned above. The advantage of the approach used in my thesis is that I calculate the substitutability of young and old *within* education levels, which eases the concern of different educational distribution between age groups in different countries and hence my method has higher external validity.

[1] These estimates are from models that explicitly allow substitution between young and old to differ from other age groups.
[2] I also estimate a model which calculates this average elasticity (i.e. a model which assumes substitution between all age groups to be the same) and got an estimate in the range of 6.1 to 6.5.

The rest of the paper is organized as follows. The next section presents the past literature, its shortcomings, and my motivation for enriching a core set of studies. Section 3 builds, lays out the identification mechanism for, and lists the limitations of the models used in my thesis. This section also compares the models. Section 4 introduces the data and the steps used in the construction of variables. The results from the estimated models are presented in section 5. Section 6 interprets the estimates and discusses it implications for labor market policies. Robustness checks are performed in section 7 to test the stability of my results. The last section provides conclusions and directions for future research.

2 Literature Review

In this section[3] I first discuss a set of core papers that utilize CES type modeling. These papers contain models that I use and extend for my thesis. Then I summarize studies from the literature that have estimated the elasticity of substitution between age groups. After that I mention a couple of atheoretical studies that reject the hypothesis that increased employment of old workers leads to fewer labor market opportunities for youth. I identify the shortcomings of the existing literature and provide motivations for the importance of my contribution.

CES methodology has been widely used in the immigration and inequality literature. In their study explaining the trends of relative wages, Katz and Murphy (1992) estimate a simple model based on CES production function and obtain an estimate of 3 for the elasticity of substitution between old and young ‘equivalents’. Card and Lemieux (2001) in their influential paper enrich the model of Katz and Murphy by assuming that different age groups with the same level of education are imperfect substitutes. They estimate the elasticity of substitution between different age groups to be between 4 and 5. Ottaviano and Peri (2012) use a nested CES model to estimate the effect of immigration on the wages of natives and get an average estimate of 5 for the elasticity of substitution between different age groups[4]. But the (2001) and (2012) studies restrict the elasticity of substitution between all age group pairs to be the same. I enrich them to explicitly allow the elasticity of substitution between young and old workers to be different from the other age group pairs.

There are other studies that look at the substitutability of age groups using CES type modeling. Welch (1979, Table 8) uses a career phase model in which different phases are

[3] Unless explicitly mentioned, all the studies mentioned in this section use U.S. data.
[4] Ottaviano and Peri (2012) use different models with different samples to get estimates in the range of 3.3 to 12.5.

imperfect substitutes and estimates the elasticity of substitution between age (experience) groups with same level of education. He obtains an estimate in the range of 5 to 10. In his paper about the effect of immigration on native wages, Borjas (2003) uses a nested CES method to obtain an estimate of 3.5 for the substitution elasticity between age groups. Fitzenberger and Kohn (2006) explain trend in skill wage premium using German data and get estimates between 5 and 20. Autor, Katz, and Kearney (2008) also use a CES model to explain wage inequality between high skilled and low skilled workers and obtain an elasticity estimate of 3.6 for age groups. Lanot and Sousounis (2016) use UK data and exploit the exogenous variation in relative wages due to minimum wage changes to show that young and old workers are complements. To my knowledge, Lanot and Sousounis (2016) is the only study that employs a CES type modeling to estimate different elasticity of substitution between different pairs of age groups. But in their study education groups are implicitly assumed to be perfect substitutes, an assumption which can create problems[5].

All the studies using a CES modeling strategy mentioned above restrict the substitution elasticity between age groups to be the same; they do not allow the elasticity between young and old workers to be different from other age groups. The only exception is Lanot and Sousounis (2016) which, however, ignores education groups and is problematic as a result. My contribution is that I build and estimate CES models which allow the elasticity between young and old workers (within education groups) to be different from the other age groups.

Flexible functional forms have also been used to estimate substitutability between age groups. Grant and Hamermesh (1981) use a Translog function and Borjas (1986) uses a Leontief function to show that young workers and adult females are strong substitutes,

[5] Studies like Welch (1979), Katz and Murphy (1992), Card and Lemieux (2001), and Ottaviano and Peri (2012) clearly show that education groups are not perfect substitutes.

implying that an increased supply of adult females is associated with adverse labor market outcomes for youth. Guest and Jansen (2016) use a CRESH (Constant Ratios for Elasticity of Substitution, Homothetic) model, which is a generalization of a CES model, to generate age-wage profiles consistent with data. The use of CRESH model in the literature has been limited because of the difficulty associated with the estimation of its parameters. The CRESH model allows the substitution elasticity to be different for each pair of age groups. Guest and Jansen argue that middle age group should be less substitutable with other age groups and young and old should be more substitutable with each other as well as with other age groups. So they "choose" rather than estimate parameters of the model and run simulations to test if such chosen parameters produce good fit for age-wage profiles. Using Australian data they conclude that estimates in the range of 10 to 20 for the substitution elasticity between young and old produce good fit for age-wage profiles. These are large values for substitution elasticity between young and old but it is important to keep in mind that these elasticities are not estimated rather they are "chosen" by researchers themselves.

If there are flexible functional forms that allow varying substitutability across age groups then why are they not used instead of CES models in the literature? One reason may be that CES models are relatively simple and easy to work with theoretically. Working with fewer parameters like in CES model can be more manageable which allows the researcher to focus on the larger picture without getting sidetracked. This is usually the case with immigration and inequality literature where the focus is on immigrants versus natives or high skilled versus low skilled workers. In both cases we have two main groups of interest and CES is not too restrictive with only two groups. Further nesting levels (like age groups) can be included which enriches the model while still keeping the model manageable. Another reason why CES models are preferred is because they have fewer parameters and can be estimated more easily and more precisely. Flexible functional forms like Leontief, Translog, and CRESH have many parameters and are more difficult to estimate precisely. And because these functional forms have many parameters, they are more likely to yield parameter estimates which are inconsistent with theory. Due to these reasons I also use CES models in my thesis.

There have been atheoretical studies which argue that increased employment opportunities for old workers do not lead to worse labor market opportunities for youth. These studies are inspired by the dual problem of youth unemployment and population aging. Population aging is putting enormous pressure on social security systems in the developed world. One obvious proposed solution is that old workers delay retirement and continue working. This solution has been attacked by arguments claiming that if old workers do not retire then young workers will not be able to find jobs or will have lower quality jobs (in terms of wages and hours worked). Using studies from twelve developed countries, Gruber and Wise (2010) disprove this argument against delaying retirement of old workers. Munnell and Wu (2012) build on one of these studies and argue that we can put causal interpretation on the effect of employment of old workers on labor market opportunities of youth. Using data from the U.S. they show that increased employment of the old has no negative effects on labor market opportunities of youth. The strength of these studies is that they exploit exogenous variations in variables and hence have high internal validity. One drawback of these studies is the difference in the distribution of education levels between age groups (especially between young and old) in different countries. These differences can lower the comparability and external validity of atheoretical studies compiled by Gruber and Wise (2010)[6].

[6] By saying that these studies are atheoretical I simply mean that they do not employ explicit theoretical models in order derive regression equations from them. These studies may very well be credible and have intuitive measurement strategies derived from a conceptual framework. The main problem I want to identify in these studies is the “local” nature of their results in a sense that we do not know much about their external validity.

The elasticity of substitution between young and old workers provides a summary measure about the ease of substitutability of young and old workers for firms in the whole economy. This can be an alternate measure to counter arguments against higher employment for old workers. Low estimates show that it is difficult to substitute young with old workers and large estimates show easy substitutability. The advantage of the CES type modeling I am employing here is that I estimate the elasticity of substitution between young and old *within* education groups. Therefore, if there are studies from many countries using the CES type models used in my thesis, then the comparability and external validity of these studies would be much higher because we would be less worried about the differences in the distribution of education levels between age groups in these countries. Another advantage of the CES type modeling is that results are backed up by economic theory.

Studies complied by Gruber and Wise (2010) intuitively show low substitutability between young and old. However, CES type models can yield pretty high estimates for elasticity of substation between age groups (Welch, 1979; Card and Lemieux, 2001; Ottaviano and Peri, 2012). To address the contradiction in these two types of studies I will enrich existing CES models so that I could test if old and young are less substitutable than other age groups. This endeavor can ease the diverging results between atheoretical and CES type studies.

3 Models and their Identification

In this section I develop, lay out the identification procedure for, and list the limitations of the three CES models used in my thesis. I also compare these models with each other. They are built on a core set of studies that use CES type modeling to explain either (a) the income inequality between high skilled (more educated) and low skilled (less educated) workers or (b) the effect of immigration on the wages of native workers. In these studies, substitutability between age groups is of secondary importance and the models they use impose same substitution elasticity between all age group pairs. I enrich the models used in these studies by using a trick similar to Krusell et al. (2000)[7]. The trick is to add extra nesting level in the age nest which allows young and old workers to be studied in isolation and to have different substitutability than other age group pairs.

3.1 Model A

This is the same model as used by Card and Lemieux (2001) and will serve as a benchmark for the remaining two models. I replicate this model using more recent data[8]. The replication will ensure that my understanding of a standard CES model and its identification method are correct. However, this model forces the elasticity of substitution between all age group pairs to be the same; Models B and C will address this issue by using a more flexible approach.

[7] Krusell et al. (2000) use this trick to test capital-skill complementarity (i.e. capital and high skilled labor are less substitutable than capital and low skilled labor).

[8] My motivation to work with recent data and ignore the past data used by previous studies like Card and Lemieux (2001) and Katz and Murphy (1992) is twofold. First of all the parameters of interest may have changed since the period analyzed by previous studies. Secondly, past and more recent CPS data are hard to harmonize because of constant changes to the measurement methodology and definition of key variables. The same recent data is used to estimate Models B and C (see section 4 for more details of the data used in my thesis).

Card and Lemieux (2001) assume a CES aggregate production function for the economy where aggregate output in period t, y_t, is a function of college labor C_t, high school labor H_t, and the technology efficiency parameters θ_{ht} and θ_{ct}:

$$y_t = \left(\theta_{ht} H_t^{\rho} + \theta_{ct} C_t^{\rho}\right)^{1/\rho} \qquad (1)$$

where $-\infty < \rho \leq 1$ is a function of the elasticity of substitution (σ_E) between college and high school graduates and $\rho = 1 - 1/\sigma_E$ (See appendix Section A1 for derivation of this formula). By construction (σ_E) is non-negative. Values of (σ_E) greater than 0 but not too large imply that the two education groups are imperfect substitutes. A value of (σ_E) greater than 1 would mean that the two groups are 'gross substitutes' and value lower than 1 would imply that these two education groups are 'gross complements' (Autor, 2012). Special cases of (σ_E) are:

- $\sigma_E = 1$ (Cobb-Douglas).
- $\sigma_E = 0$ (Leontief function) i.e. the two groups are perfect complements.
- $\sigma_E = \infty$ (Linear) i.e. the two groups are perfect substitutes.

Card and Lemieux (2001) further assumes that different age groups (j) with the same level of education are imperfect substitutes and hence aggregate college and high school labor cannot be obtained by simply adding the labor provided by different age groups. In this case we have two CES sub-aggregates for high school and college labor:

$$H_t = \left[\sum_j \left(\alpha_j H_{jt}^{\eta}\right)\right]^{1/\eta} \qquad (2)$$

$$C_t = \left[\sum_j \left(\beta_j C_{jt}^{\eta}\right)\right]^{1/\eta} \qquad (3)$$

where α_j and β_j are relative efficiency parameters (assumed to be constant over time[9]), and $-\infty < \eta \leq 1$ is a function of the partial elasticity of substitution (σ_A) between different age groups j with same level of education $\left(\eta = 1 - \frac{1}{\sigma_A}\right)$. In this case the marginal product of different age-education groups are:

$$\frac{\partial y_t}{\partial H_{jt}} = \frac{\partial y_t}{\partial H_t} \times \frac{\partial H_t}{\partial H_{jt}} = \theta_{ht} H_t^{\rho-1} \Psi_t \times \alpha_j H_{jt}^{\eta-1} H_t^{1-\eta} = \theta_{ht} H_t^{\rho-\eta} \Psi_t \times \alpha_j H_{jt}^{\eta-1} \qquad (4)$$

$$\frac{\partial y_t}{\partial C_{jt}} = \frac{\partial y_t}{\partial C_t} \times \frac{\partial C_t}{\partial C_{jt}} = \theta_{ct} C_t^{\rho-1} \Psi_t \times \beta_j C_{jt}^{\eta-1} C_t^{1-\eta} = \theta_{ct} C_t^{\rho-\eta} \Psi_t \times \beta_j C_{jt}^{\eta-1} \qquad (5)$$

where $\Psi_t = (\theta_{ht} H_t^{\rho} + \theta_{ct} C_t^{\rho})^{\frac{1}{\rho}-1}$.

Assuming that different age-education groups are utilized efficiently will imply that the wage of high school workers in age group j at time period t, w_{jt}^h , will be equal to equation (4) and wage of college workers in age group j at time period t, w_{jt}^c, will be equal to equation (5). As a result, the relative wage of college workers as compared to high school workers in age group j and time period t is given by:

$$\frac{w_{jt}^c}{w_{jt}^h} = \left(\frac{\theta_{ct}}{\theta_{ht}}\right)\left(\frac{C_t}{H_t}\right)^{\rho-\eta}\left(\frac{\beta_j}{\alpha_j}\right)\left(\frac{C_{jt}}{H_{jt}}\right)^{\eta-1} \qquad (6)$$

Taking logarithm on both sides of (6) and noting that $\rho - \eta = \left[\left(\frac{1}{\sigma_A}\right) - \left(\frac{1}{\sigma_E}\right)\right]$ and $\eta - 1 = -\left(\frac{1}{\sigma_A}\right)$ will give us:

$$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right) = log\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + log\left(\frac{\beta_j}{\alpha_j}\right) + \left[\left(\frac{1}{\sigma_A}\right) - \left(\frac{1}{\sigma_E}\right)\right] log\left(\frac{C_t}{H_t}\right) - \left(\frac{1}{\sigma_A}\right) log\left(\frac{C_{jt}}{H_{jt}}\right) \qquad (7a)$$

An alternative form of equation (7a) is:

[9] Without this assumption we cannot estimate the model because then number of parameters to be estimated will be greater than the number of observations.

$$log\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right) =$$

$$log\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + log\left(\frac{\beta_j}{\alpha_j}\right) - \left(\frac{1}{\sigma_E}\right) log\left(\frac{C_t}{H_t}\right) - \left(\frac{1}{\sigma_A}\right)\left[log\left(\frac{C_{jt}}{H_{jt}}\right) - log\left(\frac{C_t}{H_t}\right)\right] \qquad (7b)$$

Estimating equation (7a) can give us the elasticity of substitution between different age groups (σ_A). Estimating equation (7b) can give us the elasticity of substitution between college and high school workers(σ_E). But equation (7b) cannot be estimated directly because we need measures of H_t and C_t which depend on (σ_A), (α_j), and (β_j).

<u>Identification of Model A</u>

Card and Lemieux (2001) use a two-step procedure for identification of the model. In the first step (σ_A), (α_j), and (β_j) are estimated in order to construct H_t and C_t. And in the second step, (7b) is estimated to get (σ_E). The second step estimates (σ_A) as well and in principle, the estimated (σ_A) should be similar in both steps. I am mainly interested in (σ_A) but calculate (σ_E) just to make sure my estimates are consistent with theirs.

In the first step, equation (7a) cannot be estimated directly but when $log\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right)$ is regressed on $log\left(\frac{C_{jt}}{H_{jt}}\right)$, age effects (which absorb $log\left(\frac{\beta_j}{\alpha_j}\right)$), and time effects (which absorb $log\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + \left[\left(\frac{1}{\sigma_A}\right) - \left(\frac{1}{\sigma_E}\right)\right] log\left(\frac{C_t}{H_t}\right)$) then (σ_A) can be recovered (the remaining wage gap is captured by error term (u_{jt})):

$$r_{jt} = b_j + d_t - \left(\frac{1}{\sigma_A}\right) log\left(\frac{C_{jt}}{H_{jt}}\right) + u_{jt} \qquad (8)$$

where $b_j$and d_t are age and year effects (these are just age and year dummies) respectively. Once (σ_A) is estimated then (α_j) and (β_j) can be recovered using equations (4) an (5) respectively. By efficient utilization of resources, equations (4) and (5) imply that:

$$log(w_{jt}^{h}) + \frac{1}{\sigma_A} H_{jt} = log(\theta_{ht} H_t^{\rho-\eta} \Psi_t) + log(\alpha_j) \qquad (9a)$$

$$log(w_{jt}^{c}) + \frac{1}{\sigma_A} C_{jt} = log(\theta_{ct} C_t^{\rho-\eta} \Psi_t) + log(\beta_j) \qquad (9b)$$

Since (σ_A) is estimated using (8), so the left hand sides of $(9a)$ and $(9b)$ are known. The first term on the right hand side can be absorbed by a set of year dummies. The age effects (age group dummies) of regressions based on $(9a)$ and $(9b)$ will give us (α_j) and (β_j). Once we know (σ_A),(α_j), and (β_j), we can get measures of H_t and C_t using (2) and (3). Card and Lemieux (2001) assume that (a) relative productivity term $log\left(\frac{\theta_{ct}}{\theta_{ht}}\right)$ follows a linear trend and (b) relative supplies are exogenous. With these estimates and assumptions, we can now estimate equation (7b) by OLS to get the elasticity of substitution between college and high school workers (σ_E). This step will also give us an estimate for (σ_A) which, in principle, should be close to the first step estimate of (σ_A).

Limitations of Model A

The limitations of Model A are:

i. The assumption that the aggregate production function for the economy has Constant Elasticity of Substitution (CES) form, with constant returns to scale.
ii. The relative supplies are assumed to be exogenous[10].
iii. The partial elasticity of substitution (σ_A) between different age groups is assumed to be the same for college and high school graduates. Like (Card and Lemieux (2001)), I later relaxed this assumption[11].

[10] In the case of immigration we can think of supply influx of immigrants as exogenous (Borjas, 2003). Or exogenous supply can be due to demographic changes like birth and death rate changes 20 to 25 years earlier.
[11] Card and Lemieux (2001) relax this assumption about (σ_A) during robustness checks and conclude that the assumption is valid. In my robustness check for (σ_A) for Model A, the estimates were far away from each other; the estimate for high school group had the wrong sign and was insignificant and the estimate for college degree was insignificant and equal to 26. These estimates are available on request. Welch (1979) observes that young

iv. There is no capital in the production function, which may bias the parameters estimates (Hamermesh, 1986).

v. The model forces that the elasticity of substitution between age groups having equal education (σ_A) to be the same (Hamermesh, 1986). This means that the model assumes that the substitutability between age groups is independent of the distance between them i.e. age groups close to each other are assumed to have same degree of substitutability as age groups far away from each other. In Models B and C, I address this limitation.

vi. The nested CES model used here also assumes that the substitution between age groups with high school degree is unaffected by the amount of labor supplied by workers with college degree and vice versa (Hamermesh, 1986).

vii. The standard error estimates are biased downwards in the second stage because H_t and C_t are estimated themselves (via first stage estimates of (σ_A), (α_j), (β_j)), whereas the second stage treats them as data. Correct standard errors can be obtained if we bootstrap the whole procedure.

3.2 Model B

This model is based on Card and Lemieux (2001) and adds one more nesting level to Model A. I employ a similar trick used by Krusell et al. (2000)[12] in order to find a different elasticity of substitution between young and old workers. Like in model A, I assume that the aggregate output produced in the economy, q_t, has a CES production function form:

$$q_t = \left(\theta_{ht} H_t^{\rho} + \theta_{ct} C_t^{\rho}\right)^{1/\rho} \qquad (10)$$

and old workers with a high school degree are more substitutable than young and old workers with college degree. My estimates provide inconclusive results as they are insignificant.

[12] Krusell et al. (2000) use this trick to test capital-skill complementarity (i.e. capital and high skilled labor are less substitutable than capital and low skilled labor).

where the inputs and parameters are defined as in Model A.

The difference between Model A and B is that I add one more nesting level for age groups in Model B which allows young and old workers (age groups 1 and 2) to be studied in isolation and to have different substitution elasticity. More precisely, I further assume that[13]:

$$H_t = \left[\alpha_a A_t^{\eta} + \sum_{j=3}^{N} \left(\alpha_j H_{jt}^{\eta}\right)\right]^{1/\eta} \qquad (11)$$

$$A_t = \left[\gamma_1 H_{1t}^{\varphi} + \gamma_2 H_{2t}^{\varphi}\right]^{1/\varphi} \qquad (12)$$

$$C_t = \left[\beta_b B_t^{\eta} + \sum_{j=3}^{N} \left(\beta_j C_{jt}^{\eta}\right)\right]^{1/\eta} \qquad (13)$$

$$B_t = \left[\xi_1 C_{1t}^{\varphi} + \xi_2 C_{2t}^{\varphi}\right]^{1/\varphi} \qquad (14)$$

where

$\eta = 1 - \frac{1}{\sigma_A}$ and $-\infty < \eta \leq 1;$

$\varphi = 1 - \frac{1}{\sigma_{1,2}}$ and $-\infty < \varphi \leq 1;$

$\sigma_A = Elasticity\ of\ substitution\ between\ age\ groups\ j;$

$\sigma_{1,2} = Elasticity\ of\ substitution\ between\ age\ groups\ 1\ and\ 2\ (Young\ and\ Old);$

$N = number\ of\ age\ groups\ j;$

$H_{jt} = Labor\ provided\ by\ workers\ with\ high\ school\ degree\ in\ age\ group\ j, time\ period\ t;$

$C_{jt} = Labor\ provided\ by\ workers\ with\ college\ degree\ in\ age\ group\ j, time\ period\ t;$ and

$\alpha_a, \alpha_j, \gamma_1, \gamma_2, \beta_b, \beta_j, \xi_1, \xi_2 =$
$Technology\ Efficiency\ Parameters\ (unobserved\ and\ assumed\ to\ be\ time\ invariant).$

[13] Note that I add one more CES nesting level for age groups. Comparing (2) with (11) and (12) for instance shows the extra age nesting level in Model B.

I also assume only age groups 1 and 2 (young and old workers) can be aggregated like in (12) and (14). This assumption ensures that we cannot aggregate other age group pairs like young and old workers, otherwise different age group aggregations will automatically contradict my assumptions for functional forms (11) to (14). It also makes sense to only aggregate young and old workers (which are farthest away from each other) and estimate their substitutability differently. The middle age groups are closer to each other and also to young and old groups, so assuming the same substitutability between pairs of middle age groups and between middle age groups and young-old aggregator ($A_t \; and \; B_t$) is more plausible.

These assumptions allow me to focus only on old and young (age groups 1 and 2 respectively) in order to estimate a different elasticity of substitution ($\sigma_{1,2}$) between them. The marginal product of high school labor in age group 1 and time period t is:

$$\frac{\partial q_t}{\partial H_{1t}} = \frac{\partial q_t}{\partial H_t} \times \frac{\partial H_t}{\partial A_t} \times \frac{\partial A_t}{\partial H_{1t}} = q_t^{1-\rho} \theta_{ht} H_t^{\rho-1} \times H_t^{1-\eta} \alpha_a A_t^{\eta-1} \times A_t^{1-\varphi} \gamma_1 H_{1t}^{\varphi-1}$$

$$= q_t^{1-\rho} \theta_{ht} H_t^{\rho-\eta} \times \alpha_a A_t^{\eta-\varphi} \times \gamma_1 H_{1t}^{\varphi-1} \qquad (15)$$

Similarly:

$$\frac{\partial q_t}{\partial H_{2t}} = q_t^{1-\rho} \theta_{ht} H_t^{\rho-\eta} \times \alpha_a A_t^{\eta-\varphi} \times \gamma_2 H_{2t}^{\varphi-1} \qquad (16)$$

$$\frac{\partial q_t}{\partial C_{1t}} = q_t^{1-\rho} \theta_{ct} C_t^{\rho-\eta} \times \beta_b B_t^{\eta-\varphi} \times \xi_1 C_{1t}^{\varphi-1} \qquad (17)$$

$$\frac{\partial q_t}{\partial C_{2t}} = q_t^{1-\rho} \theta_{ct} C_t^{\rho-\eta} \times \beta_b B_t^{\eta-\varphi} \times \xi_2 C_{2t}^{\varphi-1} \qquad (18)$$

As in Model A, I assume perfect competition and hence (15) and (17) imply that the wage of college workers in age group 1 and time period t relative to wage of high school workers in age group 1 and time period t is:

$$\frac{w_{1t}^{c}}{w_{1t}^{h}} = \frac{\frac{\partial q_t}{\partial C_{1t}}}{\frac{\partial q_t}{\partial H_{1t}}} = \left(\frac{\theta_{ct}}{\theta_{ht}}\right)\left(\frac{C_t}{H_t}\right)^{\rho-\eta}\left(\frac{\beta_b}{\alpha_a}\right)\left(\frac{B_t}{A_t}\right)^{\eta-\varphi}\left(\frac{\xi_1}{\gamma_1}\right)\left(\frac{C_{1t}}{H_{1t}}\right)^{\varphi-1} \qquad (19)$$

Taking logarithm of (19) and noticing that $\rho-\eta=\left[\frac{1}{\sigma_A}-\frac{1}{\sigma_E}\right]$; $\eta-\varphi=\left[\frac{1}{\sigma_{1,2}}-\frac{1}{\sigma_A}\right]$; *and* $\varphi-1=-\frac{1}{\sigma_{1,2}}$ imply that:

$$log\left(\frac{w_{1t}^{c}}{w_{1t}^{h}}\right) = log\left(\frac{\theta_{ct}}{\theta_{ht}}\right)+\left[\frac{1}{\sigma_A}-\frac{1}{\sigma_E}\right]log\left(\frac{C_t}{H_t}\right)+log\left(\frac{\beta_b}{\alpha_a}\right)+\left[\frac{1}{\sigma_{1,2}}-\frac{1}{\sigma_A}\right]log\left(\frac{B_t}{A_t}\right) + log\left(\frac{\xi_1}{\gamma_1}\right)-\frac{1}{\sigma_{1,2}}log\left(\frac{C_{1t}}{H_{1t}}\right) \qquad (20)$$

The same relation holds for age group 2. Hence:

$$log\left(\frac{w_{it}^{c}}{w_{it}^{h}}\right) = log\left(\frac{\theta_{ct}}{\theta_{ht}}\right)+\left[\frac{1}{\sigma_A}-\frac{1}{\sigma_E}\right]log\left(\frac{C_t}{H_t}\right)+log\left(\frac{\beta_b}{\alpha_a}\right)+\left[\frac{1}{\sigma_{1,2}}-\frac{1}{\sigma_A}\right]log\left(\frac{B_t}{A_t}\right) + log\left(\frac{\xi_i}{\gamma_i}\right)-\frac{1}{\sigma_{1,2}}log\left(\frac{C_{it}}{H_{it}}\right) \qquad (21)$$

where i=1,2 (i.e. young and old age groups).

Identification of Model B

In equation (21), $\left(log\left(\frac{\theta_{ct}}{\theta_{ht}}\right)+\left[\frac{1}{\sigma_A}-\frac{1}{\sigma_E}\right]log\left(\frac{C_t}{H_t}\right)+\left[\frac{1}{\sigma_{1,2}}-\frac{1}{\sigma_A}\right]log\left(\frac{B_t}{A_t}\right)\right)$ can be captured by a set of year dummies (δ_t) and $\left(log\left(\frac{\beta_b}{\alpha_a}\right)+log\left(\frac{\xi_i}{\gamma_i}\right)\right)$ by dummies for age groups 1 and 2 (δ_j), with residual u_{it} capturing the rest[14]. Therefore equation (21) becomes:

$$log\left(\frac{w_{it}^c}{w_{it}^h}\right)=\delta_t+\delta_j-\frac{1}{\sigma_{1,2}}log\left(\frac{C_{it}}{H_{it}}\right)+\ u_{it} \qquad (22)$$

where i=1, 2 (i.e. young and old age groups).

Using time series for College-High School relative wages and supplies for age groups 1 and 2 (young and old workers) and by assuming that relative supplies are exogenous, we can estimate equation (22) by OLS, without intercept term, in order to obtain the elasticity of substitution $(\sigma_{1,2})$ between young and old workers.

Model B shares limitations with Model A except limitation v. The whole purpose behind designing Model B was to address this single limitation.

3.3 Model C

This model is based on Ottaviano and Peri (2012) and although it does not directly enrich the model in the original study, it does however allow young and old workers to have different substitutability from the rest of the age groups[15]. One of the main differences between Model C and the previous models is that five education groups[16] are included in

[14] Card and Lemieux (2001) or Model A make a similar statement

[15] Ottaviano and Peri include immigrant and native workers in the lowest nesting level which I do not include in my model. In a way, model C picks bits and pieces from the original study in order to estimate substitutability between young and old. During robustness checks in section 7 I estimate versions of Model C that follow Ottaviano and Peri (2012) more closely.

[16] High school dropout, high school diploma, some college but no degree, college degree, and post college degree

Model C instead of two[17]. Most of the equations and assumptions are the same as in Model B. Like in the previous models, assume that the production function for aggregate output q_t in the economy is:

$$q_t = \left[\sum_e \theta_{et} X_{et}^{\rho}\right]^{1/\rho} \qquad (23)$$

where X_{et} is the labor supplied by education group e in time period t. The rest of the parameters are the same as in Models A and B.

As in model B, I use a trick that allows young and old to have different substitutability. The trick is to simply add one more nesting level to the age nest which allows young and old (age groups 1 and 2) to be studied in isolation. More precisely, I assume that:

$$X_{et} = \left[\alpha_{et} A_{et}^{\eta} + \sum_{j=3}^{N} \left(\alpha_{ejt} X_{ejt}^{\eta}\right)\right]^{1/\eta} \qquad (24)$$

$$A_{et} = \left[\gamma_{e1t} X_{e1t}^{\varphi} + \gamma_{e2t} X_{e2t}^{\varphi}\right]^{1/\varphi} \qquad (25)$$

where X_{ejt} is the labor provided by workers in education group e, age group j, and time period t; $\alpha_e, \alpha_{ejt}, \gamma_{e1t}, \gamma_{e2t}$ are technology efficiency parameters (unobserved and assumed to be time variant[18]) and the rest of the parameters are the same as in Model B.

As previously, I also assume that only age groups 1 and 2 (young and old) can be aggregated as in equations (24) and (25). If other groups are allowed to be aggregated then it would create a contradiction to my model here.

The marginal product with respect to input X_{ejt} (j=1, 2) is:

[17] Ottaviano and Peri (2012) uses 4 education groups (no high degree, high school degree, some college but no degree, college degree). The second main difference is after equation (26)

[18] In Model B, they are assumed to be time invariant because otherwise model B cannot be estimated.

$$\frac{\partial q_t}{\partial X_{ejt}} = \frac{\partial q_t}{\partial X_{et}} \times \frac{\partial X_{et}}{\partial A_{et}} \times \frac{\partial A_{et}}{\partial X_{ejt}} = q_t^{1-\rho} \theta_{et} X_{et}^{\rho-1} \times X_{et}^{1-\eta} \alpha_{et} A_{et}^{\eta-1} \times A_{et}^{1-\varphi} \gamma_{ejt} X_{ejt}^{\varphi-1}$$

$$= q_t^{1-\rho} \theta_{et} X_{et}^{\rho-\eta} \times \alpha_{et} A_{et}^{\eta-\varphi} \times \gamma_{ejt} X_{ejt}^{\varphi-1} \qquad (26)$$

As before, I equate marginal products to wages and take their ratio. One of the main differences between the models A/B and Model C is that I take the ratio of wages of workers with different levels of education but same age group in models A/B, while in Model C I take the ratio of wages of workers with same level of education but different age groups. In particular, in model C I take the ratio of wages of workers with education level e in age group 1 to the wages of workers with education level e in age group 2:

$$\frac{w_{e1t}}{w_{e2t}} = \frac{\frac{\partial q_t}{\partial X_{e1t}}}{\frac{\partial q_t}{\partial X_{e2t}}} = \left(\frac{\gamma_{e1t}}{\gamma_{e2t}}\right)\left(\frac{X_{e1t}}{X_{e2t}}\right)^{\varphi-1} \qquad (27)$$

Taking logarithm of (28) and noticing that $\varphi - 1 = -\frac{1}{\sigma_{1,2}}$ yields:

$$log\left(\frac{w_{e1t}}{w_{e2t}}\right) = log\left(\frac{\gamma_{e1t}}{\gamma_{e2t}}\right) - \frac{1}{\sigma_{1,2}} log\left(\frac{X_{e1t}}{X_{e2t}}\right) \qquad (28)$$

<u>Identification of Model C</u>

If we assume that relative supply is exogenous and also assume some functional form for $ln\left(\frac{\theta_{oit}}{\theta_{yit}}\right)$, then the elasticity of substitution between age groups 1 and 2 ($\sigma_{1,2}$) can be estimated by OLS (with no intercept term). The functional form of $ln\left(\frac{\theta_{oit}}{\theta_{yit}}\right)$ is unobserved and following Borjas et al. (2012) I will use different sets of dummies to control for it in four different specifications: no dummies, year dummies (δ_t), education group dummies $\left(\delta_j\right)$, year and education group dummies[19] $\left(\delta_t + \delta_j\right)$. This yields the following equation:

$$log\left(\frac{w_{e1t}}{w_{e2t}}\right) = \delta_t + \delta_j - \frac{1}{\sigma_{1,2}} log\left(\frac{X_{e1t}}{X_{e2t}}\right) + u_{ejt} \qquad (29)$$

Estimating equation (29) will give us the elasticity of substitution $(\sigma_{1,2})$ between age groups 1 and 2 (young and old). It also shows the effect of relative supply of workers on their relative wages. A high value of $(\sigma_{1,2})$ will imply that (a) age groups 1 and 2 are almost perfectly substitutable and (b) there is no correlation between relative wages and supplies. Note that I cannot include unrestricted year × education[20] group dummies in order to estimate (29) because then number of parameters will be greater than the number of observations.

Ottaviano and Peri (2012) – the study on which Model C is based – uses a different regression (equation 3 in their study) for estimating the elasticity of substitution between age/experience groups. My estimation method is based on another regression (equation 8 in their study). In section7, I use the method of Ottaviano and Peri (2012) to estimate Model C.

Model C has the same limitations as Model A except limitation v. One of the purposes behind designing Model C was to address this limitation.

3.4 Comparison of the Models

The strength of Model A by Card and Lemieux (2001) is that it is a standard model in the economics literature and gives an estimate for the elasticity of substitution between different age groups. But its main limitation is that it forces the substitution elasticity between any pair of age groups to be the same. The model does not help differentiate between

[19] The assumption is that $log\left(\frac{\gamma_{e1t}}{\gamma_{e2t}}\right)$ is decomposed into these dummies and the residual term u_{ejt}. The method by Ottaviano and Peri (2012) (equation 8 in their study) uses only the last specification i.e. year and education (skill) group dummies.

[20] The number of year × education group dummies is e × t. The number of year and education group dummies is e + t. The number of observations is e × t.

substitutability of workers close to and far away from each other. In particular, I am interested in the substitutability between young and old workers and the "average" elasticity calculated in this model might not serve as an accurate estimate for this.

Model B enriches Model A by allowing the substitutability between young and old workers $(\sigma_{1,2})$ to be different from the substitution elasticity between other age group pairs (σ_A). Model B achieves this by adding an extra nesting level in Model A which separates young and old workers and allows the two age groups to be studied in isolation. Comparing estimates of (σ_A) from Model A with estimates of $(\sigma_{1,2})$ for Model B will let me test whether young and old are less substitutable than other age group pairs[21].

Model C is similar to Model B but has five education groups instead of two. It is also different in its estimation equation as compared to Models A and B. Model C enriches Model A (i.e. Card and Lemieux (2001)) and Ottaviano and Peri (2012) only in the sense that I use a trick that allows substitutability of young-old workers to be different from other age groups. Comparing estimates of Model A with Model C will let me test whether young and old are less substitutable than any other age group pairs. But this comparison will be less meaningful than the comparison between Models A and B which tests the same thing[22]. This is because Model C estimates substitutability within more narrowly defined skill groups and hence may have higher values for estimates.

Comparing Models B and C can be more meaningful than comparing Models A and C. Model B estimates substitution elasticity between young and old within two

[21] An even better comparison would be to compare (σ_A) with $(\sigma_{1,2})$, where both estimates are obtained from Model B. (See section 7 where I use a three step estimation procedure to make this comparison).

[22] Perhaps a better comparison would be to compare elasticity of substitution between age groups (σ_A) obtained from replication of Ottaviano and Peri (2012) using my data with the elasticity of substitution between young and old $(\sigma_{1,2})$ obtained from Model C. An even better comparison would be between (σ_A) and $(\sigma_{1,2})$ where both estimates are obtained from Model C itself. These two exercises are performed during robustness checks in section 7.

broad/aggregated education groups[23] (i.e. within less homogenous skill groups) while Model C estimates the substitution elasticity between young and old within five disaggregated education groups (i.e. within more homogenous skill groups). Intuitively, substitutability between young and old should increase within more homogenous skill groups. Comparing estimates from Models B and C allow me to test this intuitive reasoning.

[23] There are aggregated education groups in Model B because supplies of high school and college categories contain supplies of three education groups each. See section 4 for more about this.

4 Data and Variables

I use the 1996 to 2016 March Current Population Survey (CPS) Data which provides a representative sample for workers in the U.S. economy[24]. This is good data for my thesis because my study focuses on the entire U.S. economy. The data contains variables in both nominal and real terms. I use the variables in real terms, with base year being 2016. The key variables are relative wages and relative hours supplied by workers.

Throughout the thesis, young workers are from ages 19 and 25 inclusive and old workers are from ages 61 to 65 inclusive. My objective to study substitutability of workers just at the start and right at the end of working age cycle provides the motivation for selecting this age range for old and young workers. Model A has workers from ages 26 to 60 while Models B and C include young and old age groups as well[25]. In models A and B there are two broad education categories – High School and College[26] – while in Model C there are five disaggregated education categories (high school dropout, high school diploma, some college but no degree, college degree, and post college degree).

Following the standard procedure in the literature[27], I use the data to extract separate samples for wages and hours supplied. These two samples (one sample for wages and a different sample for hours supplied) are the same for each of the three models but the construction of variables and definitions of unit of observations (or cells) are slightly different.

[24] I download the data from CEPRdata website http://ceprdata.org/cps-uniform-data-extracts/march-cps-supplement.

[25] During robustness checks I estimated Model A by including young and old and Models B and C by excluding young and old (see section 7).

[26] These are broad education categories because the supply of High School Category not only includes supply from workers with exactly high school degree but also from high school drop outs and from workers with some college but degree. Similarly, the supply of College Category contains not only the supply of workers with exactly college degree but also the supply of workers with post college degree and workers with some college but no degree.

[27] Katz and Murphy (1992), Card and Lemieux (2001), and Ottaviano and Peri (2012).

I calculate relative wages and hours supplied for each cell. These cells are the units of observation in the regression equations of the models. For Model A, I create seven age groups and seven time periods[28] which yield forty-nine age-year cells. For Model B, I create two age groups (young and old) and twenty-one time periods[29] which yield forty-two age-year cells. For Model C, I create five education groups and seven time periods[30] which result in thirty-five education-year cells. For models A and B, I construct college-high school relative wages and hours supplied for workers in each cell. For model C, I constructed old-young relative wages and hours supplied for workers in each cell.

<u>Wage Sample and Relative Wages</u>

To construct the wage sample, I follow Card and Lemieux (2001) and restrict the sample to full time male workers. Both wage/salaried and self-employed workers are included in the sample. Wages are defined as weekly earnings and are obtained by dividing annual wage and salary earnings by the number of weeks worked in last year. I also drop extreme values like workers with weekly earnings less than 95 dollars, workers with top coded income, and workers with hourly wages less than 4 and greater than 290 dollars[31].

Relative wages for Models A and B are obtained by regressing log wages on college dummy[32], a linear age term, and white ethnicity dummy (with intercept term included); the coefficient on college dummy gives the relative wages. Log relative wages for Model C are

[28] Each time period has a three years interval i.e. 21 years from 1996 to 2016 are divided into 7 time periods with a three years interval.
[29] Each time period has a one year interval i.e. 21 years from 1996 to 2016 are divided in 21 time periods with a one year interval.
[30] Each time period has a three years interval i.e. 21 years from 1996 to 2016 are divided into 7 time periods with a three years interval.
[31] These are in 2016 dollars. Card and Lemieux (2001) also drops extreme values in almost the same manner as I have done.
[32] The dummy is equal to 1 if a worker has exactly college degree and 0 if a worker has exactly high school degree.

obtained by regressing log wages on old dummy[33] and white ethnicity dummy (with intercept term included); the coefficient on old dummy gives the relative wages.

<u>Hours Supplied Sample and Relative Hours Supplied</u>

For hours supplied sample, I once again follow Card and Lemieux (2001) and use a much broader sample consisting of full time and part time workers, male and female workers, wage/salary and self-employed workers. The outliers with respect to wage that are dropped in wage sample are now included in hours supplied sample. Hours supplied is defined as the annual work hours worked by workers. This variable is constructed by multiplying weeks worked per year by 40 for full time workers and by 20 for part-time workers.

For Models A and B, hours supplied by the high school category is the sum of annual hours supplied by all workers with exactly high school degree, plus annual hours supplied by all high school drop outs (weighted by their wages relative to high school diploma workers), plus a relevant share of annual hours supplied by all workers with some college but no degree. Hours supplied by the college category is the sum of annual hours supplied by all workers with exactly college degree, plus annual hours supplied by workers with post college degrees (weighted by their wages relative to wages of workers with exactly college degree), plus a relevant share of annual hours supplied by all workers with some college but no degree (see appendix section A2 for more details). Relative hours supplied is constructed by simply taking the ratio of hours supplied by college category to hours supplied by high school category. For Model C, relative hours supplied is constructed by simply taking the ratio of annual hours supplied by all old workers with respect to annual hours supplied by all young workers.

[33] The dummy is equal to 1 if worker belongs to old group (age 61 to 65) and 0 if she belongs to young group (age 19 to 25).

Appendix section A2 contains more details about the definitions of cells and the construction of relative wages and hours supplied for the three models. Tables A1 to A6 in Section A3 of the appendix contains the relative wages and supplies obtained from the data. The models in section 3 are estimated using the data in Tables A1 to A6 and the results are presented in the next section.

5 Results

I am mainly interested in the elasticity of substitution between young and old workers so this section will primarily report estimates of the substitution elasticity from these models. The estimation of the models using OLS regression reports standard errors for the estimates of the *reciprocal* of elasticity of substitution. Using Delta method I derive the formula for the standard errors of the estimates of elasticity of substitution itself and report the standard errors and confidence intervals for the elasticity estimates calculated using this formula. See appendix section A4 for the derivation of this formula. As far as I know, none of the CES type studies in the immigration and inequality literature estimate the standard errors and confidence intervals for the elasticity of substitution itself. The present study is the only one that tries to do so.

5.1 Results for Model A

This model has two steps/stages. In the first step, equation (8) is estimated to obtain the elasticity of substitution between age groups (σ_A). In the second step, equation (7b) is estimated to obtain the elasticity of substitution between workers in college and high school categories (σ_E). In the second step, elasticity of substitution between age groups (σ_A) is also estimated which, in principle, should be close to the estimate obtained in step 1.

The estimates from the first step are shown in Table A7 in appendix section A3. Column 1 includes both age and time dummies while column 2 has age dummies and a linear time trend instead of year dummies. The coefficient estimate on log relative supply in column 1 implies that the partial elasticity of substitution between different age groups (σ_A) with the same level of education is 6.2, quite close to Card and Lemieux's (2001) estimate of around 5. Replacing year dummies with time trend in column 2 slightly decreases the estimate of

(σ_A) to 6.1. The estimates have signs consistent with theory and are statistically significant. Table 1 summarizes the point estimates from first stage along with their standard errors.

Table 1. Main Estimates of Model A

Quantity of Interest	Value
σ_A (first stage with year dummies)	6.2*** (1.7)
σ_A (first stage with linear time trend)	6.1*** (2.1)
σ_A (second stage without recession dummy)	6.5** (2.4)
σ_A (second stage with recession dummy)	6.5** (2.4)
σ_E (second stage without recession dummy)	1.8 (1.2)
σ_E (second stage with recession dummy)	1.9 (1.4)
Number of Observations (cells)	49

Notes: In this model $\sigma_A = \sigma_{1,2}$, where $\sigma_{1,2}$ is the elasticity of substitution between young and old workers. Standard errors are in parentheses and are calculated using the formula in appendix section A4.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results from the second step (estimation of (7b)) are shown in Column 1 of Table A8 in the appendix section A3. Column 2 is the same as column 1 except that it controls for the recession period by including a dummy for it in the regression. The coefficient on $log\left(\frac{C_t}{H_t}\right)$ in column 1 implies that the elasticity of substitution between college and high school educated workers (σ_E) is 1.8. Controlling for the recession period in column 2 slightly increases the estimate to 1.9. These estimates are close to the estimates obtained by Card and Lemieux (2001), which are between 2 and 2.5. The signs on my estimates are also consistent with theory. But my estimates are not significant at usual significance levels. It is also noteworthy to mention that the estimate of (σ_A) from second stage, as implied by coefficient on $\left[log\left(\frac{C_{jt}}{H_{jt}}\right) - log\left(\frac{C_t}{H_t}\right)\right]$, is 6.5. This is not too far away from the first stage estimates of 6.1

and 6.2. Table 1 summarizes the main estimates of the second stage along with their standard errors.

The estimates in Table 1 are close to Card and Lemieux (2001), which is encouraging as it shows that my understating of the basic CES method, its identification, and estimation is correct. After properly replicating the original study, I now have the authority to try to enrich Model A.

5.2 Results for Model B

The estimates for Model B (based on equation 22) are shown in appendix Table A9. Column 1 includes both age and time dummies while column 2 has age dummies and a linear time trend instead of year dummies[34]. The coefficient estimate on log relative supply in column 1 implies that the elasticity of substitution between young and old workers $(\sigma_{1,2})$ is 3.4. Replacing year dummies with time trend in column 2 slightly increase the estimate of $(\sigma_{1,2})$ to 3.5. The estimates have signs consistent with theory. Table 2 summarizes the main estimates of this model along with their standard errors. We can see that the estimates of this model are statistically significant and reasonably precise.

Table 2. Main Estimates of Model B

Quantity of Interest	Value
$\sigma_{1,2}$ (first stage with year dummies)	3.4^{**} (1.3)
$\sigma_{1,2}$ (first stage with linear time trend)	3.5^{***} (1.1)
σ_A from 2nd stage, σ_E from 3rd stage	See Section 7
Number of Observations (first stage)	42

Notes: Standard errors are in parentheses and are calculated using the formula in appendix section A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[34] There are only two dummies for age: young and old

Using a three step procedure similar to the two step procedure of Model A, I also calculate the substitution elasticity between age (σ_A) and education (σ_E) groups for Model B. Section 7 summarizes the estimates obtained from the three step procedure and discusses the implications we can draw from those estimates.

5.3 Results for Model C

The estimates for Model C (based on equation 29) are shown in the appendix Table A10. I use four different specifications which differ in the set of year and education group dummies included in the regression: Column 1 estimates the model without any dummies, column 2 adds only year dummies, column 3 adds only education group dummies, and column 4 includes both year and education group dummies. The estimate on log relative supply in column 1 implies that the elasticity of substitution between young and old workers $(\sigma_{1,2})$ is 5.7. The sign is consistent with theory. Table 3 summarizes the point estimates along with their standard errors and shows that the estimate obtained from column 1 (i.e. 5.7) is statistically significant. The substitution elasticity estimates in the remaining three columns are statistically insignificant and hence I do not consider them.

Table 3. Main Estimates of Model C

Quantity of Interest	Value
$\sigma_{1,2}$ (without dummies)	5.7^{**} (2.4)
$\sigma_{1,2}$ (with year dummies)	-42 (33)
$\sigma_{1,2}$ (with education dummies)	30 (35)
$\sigma_{1,2}$ (with year and education dummies)	15 (22)
σ_A from second stage	See Section 7
Number of Observations	35

Notes: Standard errors are in parentheses and are calculated using the formula in appendix section A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Ottaviano and Peri (2012), the study on which Model C is built, estimate different models based on different samples and get a range for elasticity of substitution between age groups. They settle on an average value of 5 for the substitution elasticity between age groups. However, if I only consider their sample and model closest to what I use here then the estimate is 7.1 (Column 1, row 3 in Table 3 of Ottaviano and Peri, 2012).

I also estimate Model C without young and old age groups using the method employed by Ottaviano and Peri. In addition, I use a two-step procedure to estimate both (σ_A) and $(\sigma_{1,2})$ for Model C. Section 7 reports the estimates from these exercises and discusses their implications.

5.4 Summary of Results

Table 4 summarizes the point estimates of the elasticity of substitution between age groups (σ_A) and between young and old workers $(\sigma_{1,2})$ for the three models. 95% confidence intervals for these point estimates are provided in the table which helps in determining the precision of estimates. All the point estimates have signs consistent with theory and are also statistically significant at usual significance levels. One thing to notice is that the 95% confidence intervals for point estimates are very large. This means that estimates are imprecise and there is a huge overlap in their confidence intervals. Therefore care should be taken when interpreting the point estimates.

One reason for imprecise estimates is that I have fewer observations and more parameters to work with for Models B and C as compared to Model A and previous studies like Card and Lemieux (2001) and Ottaviano and Peri (2012). Another reason may be that previous studies investigate time periods that are characterized by significant variations in relative wages and hours supplied, hence they obtain more precise results. Lastly, heteroskedasticity is very likely present because the units of observations are cells and each cell has different number of workers in it. The (2001) and (2012) studies both use weighted least squares which most likely yields much more efficient estimates.

Table 4. Estimates and Their Confidence Intervals

Models	**Point Estimates**	**95% Confidence Interval for Point Estimates**	**Comments**
Model A	6.2	(2.7 , 9.7)	These are the elasticity of substitution between all age groups i.e. the elasticity between young and old is assumed to be the same as between any other age group pair. The original study by Card and Lemieux (2001) yields point estimates of 3.8 and 4.9; the 95% CI for point estimates 3.8 and 4.9 are (3, 4.6) and (3.9, 5.9) respectively.
	6.1	(1.9 , 10.3)	
	6.5	(1.7 , 11.3)	
	6.5	(1.7 , 11.3)	
Model B	3.4	(0.7 , 6.1)	These are the elasticity of substitution between young and old workers $(\sigma_{1,2})$ and are allowed to be different in principle from the elasticity of substitution between any arbitrary age group pair (σ_A). See section 7 for estimates of (σ_A) for Model B.
	3.5	(1.3 , 5.9)	
Model C	5.7	(0.8 , 10.6)	This is the elasticity of substitution between young and old workers $(\sigma_{1,2})$ and is allowed to be different from the elasticity of substitution between other age group pairs (σ_A). See section 7 for estimates of (σ_A) for Model C.

6 Interpretation of Results

In this section I interpret the estimates. First, I compare the estimates of the three models and explain how these comparisons lead to two important conclusions. Second, I provide intuitive and graphical explanations for the concept of elasticity of substitution between young and old workers. Lastly, I argue how these estimates are relevant to policy.

6.1 Comparing Estimates of the Models

Model A assumes the same substitution elasticity between all age group pairs. The estimate of (σ_A) for this model ranges from 6.1 to 6.5. In the other models I estimate the elasticity of substitution between young and old $(\sigma_{1,2})$ and have explicitly allowed this estimate to be different from (σ_A). The estimates of $(\sigma_{1,2})$ for Model B are between 3.4 and 3.5. These are less than the estimates of (σ_A) for Model A. In other words, the elasticity of substitution between young and old is less than the substitution elasticity between any other age group pairs $(\sigma_{1,2} < \sigma_A)$. This leads to the first conclusion of this study: young and old workers are less substitutable than other age group pairs. The same result holds when we compare models A and C but the estimates of $(\sigma_{1,2})$ for model C are much closer to estimates of (σ_A) for model A. As mentioned in section 3.4 the comparison between Models A and B is much more meaningful than the comparison between models A and C. A more meaningful comparison would be to first obtain both $(\sigma_{1,2})$ and (σ_A) for Model B and then compare these estimates with each other and later repeat the same for Model C (see section 7 for these exercises).

We must be careful when working with point estimates. As Table 4 in section 5.4 shows, the confidence intervals of $(\sigma_{1,2})$ and (σ_A) are large which results in overlapping intervals for estimates. Although there is a large overlap in their confidence intervals, if we imagine a real number line then Model B's interval for $(\sigma_{1,2})$ is lower than Model A's interval for (σ_A).

Next I compare elasticity of substitution between young and old $(\sigma_{1,2})$ for models B and C. I do not consider Model A for this comparison because Model A does not explicitly estimates $(\sigma_{1,2})$. In Model B there are two broad education groups (High School and College) and $(\sigma_{1,2})$ is estimated for workers 'within' these education groups. In Model C there are five disaggregated education groups and $(\sigma_{1,2})$ is estimated for workers 'within' these education groups. As we move from Model B to C, the elasticity of substitution between young and old $(\sigma_{1,2})$ is calculated within more narrowly defined (or more homogenous) skill groups. Intuitively, the ease of substitutability between young and old should increase as the skill group within which substitutability is measured becomes more homogenous. The elasticity estimates from Models B and C are in line with this intuitive reasoning; as we move from Model B to C, the elasticity of substitution between young and old[35] $(\sigma_{1,2})$ increases from 3.45 to 5.7. This leads to the second conclusion of this study: the substitutability of young and old workers increases within more homogenous skill groups. Again care must be taken with comparing point estimates because of the overlaps in the confidence intervals of these point estimates (see Table 4).

6.2 Interpretation of the Elasticity of Substitution between Young and Old

Elasticity of Substitution is fundamentally a parameter of the production function (in my case, a CES production function). Intuitively it shows the ease of substituting one input with another in the production process, assuming that the output remains fixed. The higher the value for substitution elasticity, the easier it becomes for one input to be replaced by

[35] Since Models B has a range for the substitution elasticity, so I have used the average estimate in the range i.e. 3.45 is the average of 3.4 and 3.5.

another. If we have two inputs then it basically shows the shape of isoquant curves drawn in standard economic texts like Nicholson and Snyder (2008, chapter 9). For instance, if we consider only young and old inputs then the isoquant can have the shapes shown in Figure 1.

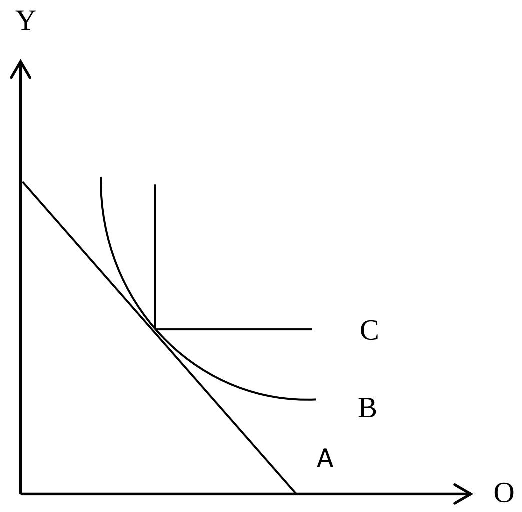

Figure 1. Possible Shapes for Isoquant with Old (O) and Young (O) labor inputs

In Figure 1, line segment A shows that young and old are perfect substitutes (infinite substitution elasticity), L shaped curve C shows that young and old are perfect complements (zero substitution elasticity), and curve B shows that young and old are imperfect substitutes (greater than zero but finite substitution elasticity). For B to be an isoquant for Cobb-Douglas production function, the substitution elasticity must be 1. As we move from C to B to A, the substitutability of young with old increases.

More precisely, the elasticity of substitution between young and old (σ_{Y-O}) defined as in Hamermesh (1986) is:

$$\sigma_{Y-O} = -\frac{\partial \ln\left(\frac{Y}{O}\right)}{\partial \ln\left(\frac{w^y}{w^o}\right)} = \frac{\partial \ln\left(\frac{Y}{O}\right)}{\partial \ln\left(\frac{w^o}{w^y}\right)} \qquad (30)^{36}$$

where Y and O stands for the amount of young and old labor used and w^y and w^o are their wages respectively. The substitution elasticity shows the percentage change in young-old relative labor used due to a one percentage change in old-young[37] relative wages, keeping output in the economy fixed. Since (σ_{Y-O}) is non-negative by the structure of the CES production function, if we ignore the $(\sigma_{Y-O} = 0)$ case then an increase (decrease) in old-young relative wages will increase (decrease) young-old relative labor used or employed.

For graphically interpreting (σ_{Y-O}), we have to assume that the supply of young workers relative to old workers is infinitely elastic (i.e. wages of young relative to old workers $\left(\frac{w^y}{w^o}\right)$ is determined exogenously). In this case, exogenous relative supply changes (wage changes) will trace out the relative demand curve and the reciprocal of elasticity of substitution is the slope of the relative demand curve. In particular, if relative wages $\left(\frac{w^y}{w^o}\right)$ decrease then relative labor employed $\left(\frac{Y}{O}\right)$ will increase, as shown in Figure 2.

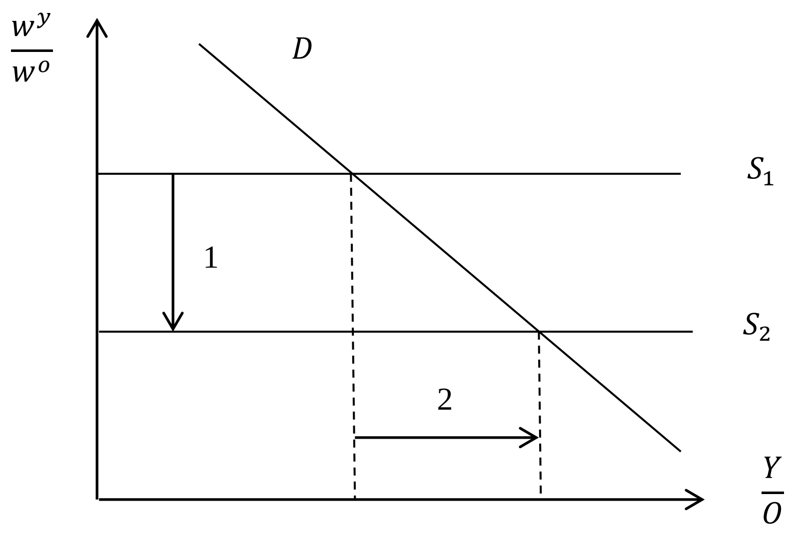

Figure 2. Young-Old Relative Demand and Substitution Elasticity

[36] By construction of CES model, σ_{Y-O} is non-negative. Therefore, in the first equality $\partial \ln\left(\frac{Y}{O}\right)$ is non-positive and in the second equality $\partial \ln\left(\frac{Y}{O}\right)$ is non-negative.

[37] When I say old-young then I am taking the ratio of old to young and when I say young-old then I am taking the ratio of young to old. It is important to note this distinction in order to avoid confusion.

Following Hamermesh (1985), I will interpret σ_{Y-O} graphically in another manner as well. Assume that the supply of young workers relative to *middle age group* is infinitely elastic i.e. wages of young relative to middle age group (w_{y-m}) is determined exogenously[38]. Similarly, assume that the supply of old workers relative to middle age group is infinitely elastic i.e. wages of old relative to middle age group (w_{o-m}) is determined exogenously. Figure 3 shows the setup for young workers with a usual downward sloping demand curve. In this setup, employers take the wages as given and the employment level is determined by the demand curve.

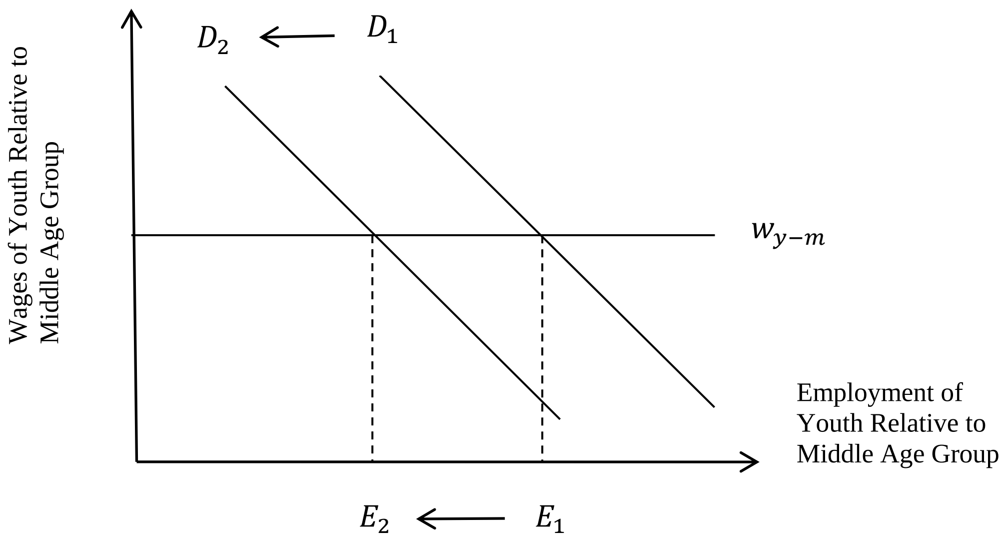

Figure 3. Graphical Interpretation of Elasticity of Substitution between young and old

Now suppose that wages of old relative to middle group (w_{o-m}) decreases exogenously while wages for young relative to middle age group (w_{y-m}) remains the same. This means that wages of old relative to young $\left(\frac{w^o}{w^y}\right)$ decreases. For convenience suppose $\left(\frac{w^o}{w^y}\right)$ decreases by 1 percent. This can make young workers less attractive as compared to old

[38] It is crucial to note that wages and employment of youth are relative to middle age group and not relative to old workers. Otherwise we would have young-old relative demand curve being traced by exogenous relative supply curve, with technological shocks shifting the demand curve – like the one shown in Figure 2 above and in Figure A1 in appendix section A1.

workers for firms and the optimal combination of young and old in the production process may change. If $\sigma_{Y-O} = x$ (x is non-negative here) then this change in relative wages would mean that if the output and other prices in the economy remain fixed, the demand for youth relative to middle age group will shift inward from D_1 to D_2. This results in a decrease in employment of youth from E_1 to E_2. The fall from E_1 to E_2 is such that the decrease in $\left(\frac{Y}{O}\right)$ is exactly by x percent. The estimates from Models B and C imply that a 1 percent decrease in old-young relative wages leads to a 3.4 to 5.7 percent decrease in young-old relative annual hours supplied. Since estimates for substitution elasticities are non-zero for all three models so young and old are p-substitutes.

In all three models, the young-old substitution elasticity is greater than 1. Therefore young and old workers are gross substitutes. This means that if old-young relative wages increases then old-young relative wage bill will decrease, where wage bill is defined as the product of wages and hours supplied of old workers. In other words, the fall in old-young relative annual hours supplied due to an increase in relative wages of old workers is large enough to not only neutralize but also penalize the relative wage advantage of old workers.

6.3 Relevance of the Elasticity of Substitution to Public Policy

One of my motivations behind estimating the elasticity of substitution between young and old workers was to critically analyze the conventional wisdom that increased labor force participation by old workers hampers the labor market opportunities of youth. This takes many forms: (1) old workers delaying retirement (due to external incentives like less generous pension and health care) and staying longer in the labor market would mean that unemployment rates of youth will increase; and (2) providing monetary incentives to firms to retain their old workers would mean that less and less young workers will be employed at firms or young workers will be working less hours. This conventional wisdom usually results in changes in policies that do more harm than good for the whole economy.

Many studies have disproved this conventional wisdim (Gruber and Wise, 2010; Munnell and Wu, 2012). These are atheoretical studies usually employing regressions and/or program evaluation tools like instrumental variables. These studies have higher internal validity because they restrict analysis to periods and countries where there are exogenous changes in variables due to policy changes or use other methods like IV to take care of endogeneity. But these studies can have lower external validity and low comparability with each other because the distribution of education levels in age groups might differ considerably between countries.

The elasticity of substitution can be an alternative measure to disprove the conventional wisdom discussed above. The elasticity of substitution provides a summary measure for the ease of substitutability of young and old workers in the production process of the economy. Intuitively, low estimates for elasticity means that young and old workers are not easily substitutable. In section 6.1, I showed that the elasticity of substitution between young and old is less than the elasticity of substitution between any other age group pair. This means firms can more easily substitute middle aged workers by old workers in the production process and hence greater employment by old workers may not harm employment of youth a lot. The advantage of elasticity of substitution used in my models is that it eases the concern about the differences in educational distribution in age groups between countries because what I am really estimating is the elasticity of substitution between young and old *within* education groups. As a result, studies using substitution elasticity can have higher external validity and greater comparability with each other. However, endogeneity can be an issue in the approach I am employing in my thesis which lowers internal validity. Future studies that combine natural experiments / program evaluation methods with CES modeling to estimate elasticity of substitution can have both higher internal and external validity.

The elasticity of substitution provides an intuitive idea about the ease of substitution between annual hours supplied by young and old workers and therefore helps in analyzing point (1) mentioned in the beginning of this section. But if I am to quantify the effect of increased supply/stock of old workers on young workers' employment and/or wages then I have to estimate other elasticities not discussed in my thesis, for instance, the elasticity of complementarity[39]. For the elasticity of complementarity I need to assume the other extreme for relative supply i.e. relative supply is completely inelastic. This means that I impose that there is a fixed stock of young and old workers for whom firms complete by fixing wages (Hamermesh, 1985). However, the effect of a one percent change in old-young relative wages on young-old relative employment/hours supplied can be quantified here because the definition of elasticity of substitution shows exactly this effect. This will help in the analysis of point (2) mentioned at the beginning of this section. The next two paragraphs will quantify this effect.

Suppose a government provides firms with an incentive that subsidizes the wages of old workers. For convenience, suppose that this results in a 1% decrease in old-young relative wages. This exogenous change in wages of old workers will change the optimal allocation of young and old workers in the production process, if we assume the output and everything else are to remain constant. More precisely, in terms of figure 2 the young-old relative supply (or wages) will shift upward and young old relative employment will fall. In terms of figure 3 this means that the demand for young workers relative to *middle age group* will shift inwards and the employment (or annual hours supplied) of young workers will fall. The estimates

[39] Hamermesh (1985) argues that if young and old are q-substitutes then an increased supply of old workers can lead to increased unemployment of youth if the relative real wages of youth are downward rigid.

from Models B and C imply that due to this 1% decrease in old-young relative wages, the young-old annual hours supplied will decrease by 3.4 to 5.7%.

The decrease in employment for youth is less than the other age groups. Model A shows a 6.1 to 6.5% decrease in annual hours supplied of other age groups relative to old due to a 1% decrease in wages of old relative to other age groups. The existing literature (Card and Raphael, 2013) would imply a fall of between 5 to 10% for the employment of other age groups relative to old due to a similar decrease in old workers' wages. These results provide support to the atheoretical studies mentioned above (Gruber and Wise, 2010; Munnell and Wu, 2012) and also address the contradictory results of atheoretical and CES type studies pointed out in the literature review section.

In a non-nested CES production function, inputs can never be complements (Card, 2016). Using a nested CES model, like model A, I get estimates of both elasticity of substitution between age (σ_A) and education (σ_E) groups. In relative terms, if (σ_A) is small as compared to (σ_E), then age groups i and j with same level of education can be Allen complements, which is not possible in a non-nested CES model (Card, 2016, p.11). Card (2016) argues that for the estimates of (σ_A) and (σ_E) found in Card and Lemieux (2001), age groups with same level of education are Allen complements. My estimates from Model A are close to Card and Lemieux (2001) and hence I draw the same conclusion. Since the elasticity of substitution between young and old ($\sigma_{1,2}$) from Models B and C is lower than elasticity of substitution between all age groups (σ_A) so this result is even more true for young and old workers i.e. young and old with same education level appear to be stronger Allen complements than arbitrary age group pair with same level of education. This is an important piece of evidence against arguments claiming that greater employment of old hurts youth.

7 Robustness Checks

In this section I summarize the results and the implications of those results for five further exercises I have pursued in this study. The first exercise is for Model A, the second and third for Model B, and the last two are for Model C. Exercises one and three are in principle doing the opposite of each other. Standard errors of estimates are in brackets after the estimates and are calculated using the formula in appendix section A4

Like in Card and Lemieux (2001), Model A restricts the analysis to workers from age 26 to 60 inclusive. Model A provides an estimate in the range of 6.1 (2.1) to 6.5 (2.4) for the 'average' elasticity of substitution between age groups (σ_A). In the first exercise, I re-estimate Model A by including young (19-25) and old (61-65) age groups in the model and obtain an estimate in the range of 5.2 (1.1) to 5.5 (1.2) for the `average' elasticity of substitution between age groups ($\sigma_A$). The reduction in `average' substitution elasticity implies that young and old workers are *less* substitutable with each other and with other age groups. This result strengthens the first conclusion of my thesis that young and old workers are less substitutable than other age groups. The data used in this exercise and the model estimates are shown in Tables A11 to A13 in appendix section A3.

In the second exercise I employ a three step procedure, which is somewhat similar to the two step procedure of Model A, to obtain estimates for the elasticity of substitution between age (σ_A) and education (σ_E) groups for Model B. Appendix section A5 contains the detail of the three step procedure I use here[40]. The first step was implemented in previous sections which gave elasticity of substitution between young and old ($\sigma_{1,2}$) in the range of 3.4 (1.3) to 3.5 (1.1) (see Table 2). In the second step I obtain estimates in the range of 4.3 (0.8) to 4.6 (0.9) for (σ_A). The third step yields estimates for (σ_E) between 3.2 (2.3) to 3.5 (2.5).

[40] The main obstacle in the three step procedure of Model B as compared to the two step procedure of Model A is the construction of relative wages and supplies for the young-old aggregator term.

The data for this exercise and the estimates for the second and third steps are shown in Tables A14 to A17 in appendix section A3.

This second exercise leads to two observations. First, the estimates of (σ_A) and (σ_E) are close to the corresponding estimates found in Model A, Card and Lemieux (2001), and other studies. This provides more credibility to my Model B and shows that the assumptions I make in Model B are more of less innocuous because they are not destroying the underlying structure of the production process of the economy. Second, the substitutability between young and old is less than the substitutability between other age group pairs $(\sigma_{1,2} < \sigma_A)$. This result further strengthens the first conclusion of my thesis.

In the third exercise, I estimate Model B without the young-old aggregator term. This in principle means transforming Model B into A. However, this exercise is not the same as estimating Model A because now I have three times more observation as compared to Model A[41]. Estimating Model B *without* the aggregator term yields estimates for the elasticity of substitution between age groups (σ_A) in the range of 6.8 (1.6) to 7 (1.5). The estimate of (σ_A) in Model B *with* the aggregator term is obtained from the second exercise which ranges from 4.3 (0.8) to 4.6 (0.9). The estimate of (σ_A) in Model A, where there are three times less observations to work with, ranges from 6.1 (2.1) to 6.5 (2.4) (see section 5). The data for this exercise is the same as in exercise 2 and can be found in Tables A14 and A15 in appendix section A3; the only difference is that I drop the observations for young-old aggregator term. The estimates from model are shown in appendix Table A18.

This third exercise leads to two observations. First, the average elasticity of substation between age groups *increases* when young and old groups are removed from the model. This means young and old are less substitutable with themselves and other age groups. This result,

[41] This is because I am using 21 one-year- interval time periods instead of 7 three-year-interval time periods.

like the exercise 1, strengthens the first conclusion of my thesis. The second observation is that the estimates of (σ_A) using fewer or more observations are very close to each other. However, the advantage of having more observations and, as a result, more precise estimates for (σ_A) in this exercise as compared to Model A will come at the expense of higher measurement errors in variables due to using smaller cells (the number of workers in age group – year cells are much lower now). Which model is preferred will depend on whether we want more precision or less measurement error.

In the fourth exercise, I use the estimation method of Ottaviano and Peri (2012) to estimate Model C without the young-old aggregator term (i.e. only workers from age 26 to 60 are included). Note that model C has five education groups and therefore the substitution between age groups is measured within more homogenous skill groups. The details of model and its identification can be found in appendix section A6. Column 1 in Table A19 in appendix section A3 shows that estimating this model yields an estimate of 29 (16) for (σ_A). This estimate implies that for most practical purposes age groups are perfect substitutes[42].

In the last exercise, I use a two-step procedure for Model C to estimate (σ_A) and $(\sigma_{1,2})$. The details of this procedure are outlined in appendix section A6. In the first step, young and old are studied in isolation. Column 2 in Table A19 in appendix shows that estimation of first stage yields an estimate of 10 (4.2) for $(\sigma_{1,2})$. And column 3 of the same tables shows[43] that the second step yields $\sigma_A = \infty$. This implies that practically age groups are perfect substitutes.

[42] I also did the fourth exercise by including young and old workers in the model (not the young-old aggregator term; only simple young and old age groups like in exercise 1 for Model A). This *increased* the estimate of (σ_A) to 40 (26). It looks as if the point estimate in this case contradicts my first conclusion because substitutability of age groups increases when young and old are included in the model. But large estimates of (σ_A) combined with large standard errors for both with and without young and old age groups lead to inconclusive results.

[43] In this study, any elasticity estimate greater than 100 is interpreted as infinite (∞) elasticity.

It was mentioned previously that comparing (σ_A) obtained from Model A with $(\sigma_{1,2})$ obtained from Model C might lead to wrong conclusions. This is because Model C estimates substitutability between more homogenous groups and as a result substitution between young and old can be relatively very high in this model. It was suggested that a better comparison would be to first obtain (σ_A) for Model C and then compare it with $(\sigma_{1,2})$ of Model C. Estimates from the fourth and fifth exercises allow us to make the latter comparison. We can clearly see that $(\sigma_{1,2})$ is less than (σ_A) for Model C. This strengthens the first conclusion of my thesis.

Finally, $(\sigma_{1,2})$ obtained for Model C from the last exercise is higher than $(\sigma_{1,2})$ obtained for Model C using my estimation method (see Table 3). This means that $(\sigma_{1,2})$ is between 5 and 10 for Model C. The estimate of $(\sigma_{1,2})$ for Model B is between 3.4 and 3.5. Therefore, no matter how we estimate the Model C, the value of $(\sigma_{1,2})$ is less in this model as compared to $(\sigma_{1,2})$ obtained from Model B. This provides more credibility to my second conclusion: young and old are more substitutable with more homogenous skill groups. Not surprisingly, comparing estimates of (σ_A) from Models A and B with estimates of (σ_A) from Model C we arrive at the conclusion that all age group pairs, not just young and old, are more substitutable within more narrowly defined (or more homogenous) skill groups.

8 Conclusions

I estimate the elasticity of substitution between young and old workers using CES type models. The existing CES models in the literature do not allow the substitution elasticity between young and old workers to be different from other age group pairs. My contribution to the literature is that I develop, compare, and estimate three CES type models, two of which allow the elasticity of substitution between young and old workers $\left(\sigma_{1,2}\right)$ to be different from the elasticity of substitution between other age groups (σ_A). Estimating these models using 1996 to 2016 March CPS data yield estimates summarized in Table 5.

Table 5. Estimates[44] From All Models of this Study

Models	Age Groups	Estimates for (σ_A)	Estimates for $\left(\sigma_{1,2}\right)$
Model A	Without Young and Old	6.1 to 6.5	-
	With Young and Old	5.2 to 5.5	-
Model B	Without Young and Old	6.8 to 7	-
	With Young and Old	4.3 to 4.6	3.4 to 3.5
Model C	Without Young and Old (Based on OP's method[45])	29	-
	With Young and Old (Based on OP's method)	∞	10
	With Young and Old (Based on my estimation method)	-	5.7

Table 5 clearly shows that the elasticity of substitution between young and old $\left(\sigma_{1,2}\right)$ is less than the elasticity of substitution between other age group pairs (σ_A). So the first conclusion I draw from the estimates of these models is that young and old workers are less

[44] Any elasticity estimate greater than 100 is interpreted as infinite (∞) elasticity. Note that Model A does not allow (σ_A) to be different from $\left(\sigma_{1,2}\right)$. Also, if we do not include young and old age groups in the model (shown by the second column) then we cannot estimate $\left(\sigma_{1,2}\right)$. These two points explain why some cells in Table 5 are left empty. Lastly, (σ_A) in the last row is not calculated because technically I cannot use a two-step procedure to find (σ_A) using my estimation method.

[45] i.e. using the estimation method of Ottaviano and Peri (2012).

substitutable than other age group pairs[46]. We can also see that the substitution elasticity between young and old $(\sigma_{1,2})$ is greater in Model C than in Model B[47]. This leads to my second conclusion: substitution elasticity is larger in more narrowly defined (or more homogenous) skill groups, implying relatively easier substitutability of young and old in more homogenous skill groups. Not surprisingly, the second conclusion holds for other age groups as well. This second conclusion makes intuitive sense.

There are atheoretical and program evaluation studies that reject the conventional wisdom that increased labor market participation by old workers will come at the expense of labor market opportunities for youth (Gruber and Wise, 2010; Munnell and Wu, 2012). However, these studies can have lower external validity and low comparability with each other because of the differences in the distribution of education between age groups in these countries. The elasticity estimates in my study show that young and old are not easily substitutable, especially when compared to the substitutability between old and middle age groups for instance and as a result more employment by old is less likely to hurt young people a lot in terms of employment, if total output in the economy is to remain fixed. Hence estimates from CES models used here strengthen atheoretical studies that reject this conventional wisdom. These CES models also have higher external validity than the atheoretical studies because substitutability is calculated *within* education groups and as a result we can be less worried about the differences in the educational distributions of age groups between countries.

[46] This conclusion is stronger when we compare estimates of Model B with estimates of Model B itself or with estimates of Model A. As mentioned previously, comparing Model C with Model A will lead to wrong conclusions. It is better to compare $(\sigma_{1,2})$ obtained from Model C with (σ_A) obtained from Model C itself.

[47] Remember that there are five disaggregated education groups in Model C as compared to only two aggregated education groups in Model B. So skill groups within which elasticity is calculated are more homogenous in Model C than in Model B.

The econometric technique used in this thesis is restricted to simple OLS. Future studies can employ more sophisticated econometric techniques to estimate the models I have developed. Generalized Least Squares, Instrumental Variables, Fixed Effects, GMM, and Simulated Maximum Likelihood methods have been used to estimate CES type models in the literature. It would be interesting to compare the magnitude and precision of estimates obtained with these methods with the ones obtained here. In addition, program evaluation methods combined with CES modeling approach used here will have both higher internal and external validity.

Another possibility for future research includes the use of occupations to define skill groups instead of or in addition to the education groups used in the models. Occupations identify the skill requirements of jobs and tasks performed by workers and hence substitutability within occupation groups might be even higher than in case of education groups.

Lastly, the literature on the elasticity of substitutability and complementarity is huge and confusing. Future studies can situate the substitution elasticity of my model in the context of the literature on elasticities and maybe even try to estimate other elasticities for my models. These elasticities may provide more quantifiable effects of increased stock of old workers on the wages and employment of youth.

Appendix

Section A1. Derivation of Elasticity of Substitution for CES production function

Card and Lemieux (2001) assume a CES aggregate production function for the economy where aggregate output in period t, y_t, is a function of college labor C_t, high school labor H_t, and the technology efficiency parameters θ_{ht} and θ_{ct}:

$$y_t = \left(\theta_{ht} H_t^{\rho} + \theta_{ct} C_t^{\rho}\right)^{1/\rho}$$

The marginal products of different education groups are:

$$\frac{\partial y_t}{\partial H_t} = \theta_{ht} \Psi_t H_t^{\rho-1}$$

$$\frac{\partial y_t}{\partial C_t} = \theta_{ct} \Psi_t C_t^{\rho-1}$$

where:

$$\Psi_t = \left(\theta_{ht} H_t^{\rho} + \theta_{ct} C_t^{\rho}\right)^{\frac{1}{\rho}-1}$$

For efficient utilization of different education groups, their marginal product should be equal to their wages. Assuming efficiency we get:

$$w_t^h = \frac{\partial y_t}{\partial H_t} = \theta_{ht} \Psi_t H_t^{\rho-1}$$

$$w_t^c = \frac{\partial y_t}{\partial C_t} = \theta_{ct} \Psi_t C_t^{\rho-1}$$

Wage of college educated workers relative to high school educated workers then becomes:

$$\frac{w_t^c}{w_t^h} = \frac{\theta_{ct}}{\theta_{ht}} \left(\frac{C_t}{H_t}\right)^{\rho-1}$$

Taking log on both sides we get:

$$ln\left(\frac{w_t^c}{w_t^h}\right) = ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + (1-\rho)ln\left(\frac{H_t}{C_t}\right) \qquad (31)$$

Alternatively we can write:

$$ln\left(\frac{H_t}{C_t}\right) = -\frac{1}{(1-\rho)}ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + \frac{1}{(1-\rho)}\ ln\left(\frac{w_t^c}{w_t^h}\right)$$

Now by the definition of elasticity of substitution (Hamermesh, 1985):

$$\sigma_E = \frac{\partial \ln\left(\frac{H_t}{C_t}\right)}{\partial \ln\left(\frac{w_t^c}{w_t^h}\right)} = \frac{1}{(1-\rho)}$$

Or alternatively:

$$\rho = 1 - \frac{1}{\sigma_E}$$

With this result, equation (31) becomes:

$$ln\left(\frac{w_t^c}{w_t^h}\right) = ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right) - \frac{1}{\sigma_E}ln\left(\frac{C_t}{H_t}\right) \qquad (32)$$

If we assume that relative supply is exogenous then (32) is the college-high school relative demand curve. Figure A1 below shows this relative demand curve. Relative supply is exogenous i.e. relative supply curve is horizontal (infinitely elastic) and when it exogenously moves up and down the relative demand curve is traced out. The slope of the relative demand curve is $-\frac{1}{\sigma_E}$. The term $ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right)$ shows the shifts in demand due to say technological shocks.

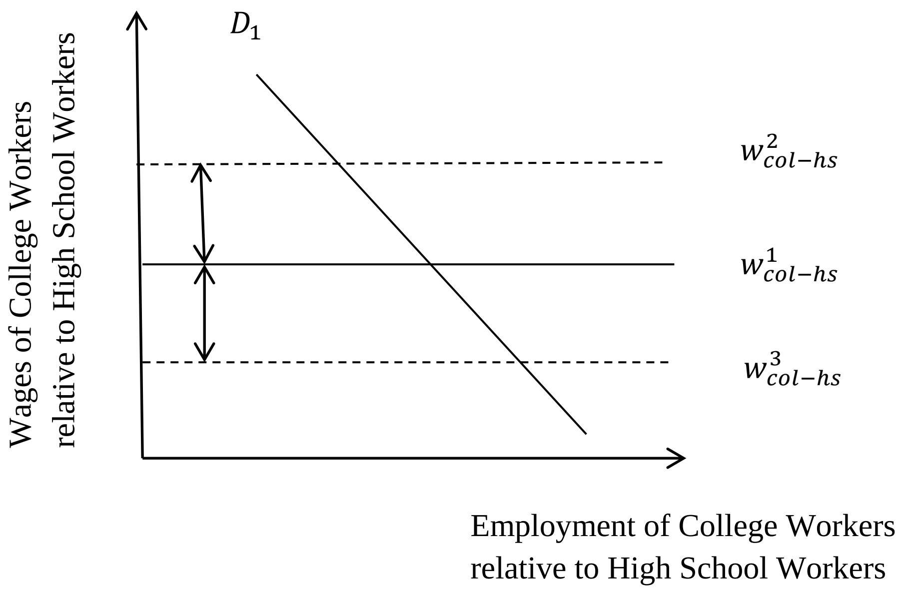

Figure A 1. College-High School Relative Demand Curve

Section A2. Data Appendix

Wage sample is restricted to full time male workers. Full time status is based on work the week before the survey. Both wage/salaried and self-employed workers were included in the sample. Hours supplied sample consists of full time and part time workers, male and female workers, wage/salary and self-employed workers. Full time and part time status is based on last year’s work.

Data for Model A

Definition of Cells

For estimating Model A (equations 8 and 7b), I construct College-HS relative wages and hours supplied for each age-year group/cell. Each cell contains an age group j for a time period t. I create seven age groups with a five years interval, and seven year groups with a three years interval. This results in forty nine age-year groups/cells for which I calculate the relative wages and supplies. I create rolling age groups. For example, the age group 26-30 for 1996-1998 year period contain workers aged 25-29 in 1996, 26-30 in 1997, and 27-31 in 1998. Each of these cells forms a unit of observation in the estimation of Model A.

Relative Wages

I create a college dummy equal to 1 if a person has exactly college degree and 0 if he has exactly high school degree. The estimates for log relative earnings are obtained by regressing log of real earnings on college dummy for each age-year group for Model A. A linear age term and a dummy for white ethnicity are included in the regression as well. The coefficient on the college dummy gives the relative earnings. The log relative earnings are shown in appendix Table A1.

Relative Hours Supplied

Hours supplied by high school category includes total annual hours supplied by all workers with exactly high school degree, plus the total annual hours supplied by all high school drop outs (weighted by their wage relative to high school workers), plus a share of total annual hours supplied by workers with some college. This share is the high school weight which can be used to express wages of workers with some college as a weighted average of high school and college wages. In my sample this share is 0.77. And the hours supplied by college category is the total annual hours supplied by college graduates, plus total annual hours supplied by workers with post college degrees (weighted by their wages relative to college graduates), plus an appropriate share (0.23) of the total annual hours supplied by workers with some college[48]. Relative hours supplied is simply constructed by dividing the annual hours supplied by college category with the annual hours supplied by high school category. The log relative annual hours supplied are shown in appendix Table A2.

Data for Model B

For estimating Model B (equation 22), I construct College-HS relative wages and hours supplied for young and old workers. Since I am interested in the substitutability of workers at the start and end of their work cycle so I define young worker to be between ages 19 and 25 and old worker to be between ages 61 and 65.

Definition of Cells

Each cell contains an age group j for a time period t. I create two age groups (young and old), and twenty one year groups with a one year interval. This results in forty two age-year groups /cells for which I calculate the relative wages and supplies. Each of these cells forms a unit of observation in the estimation of Model B.

[48] Card and Lemieux (2001) use 50-50 share (Card, 2016)

Relative Wages

Relative wages for each cell are constructed in the same manner as in Model A. The log relative earnings are shown in appendix Table A3.

Relative Hours Supplied

Relative hours supplied for each cell are constructed in the same manner as in Model A. The log relative annual hours supplied is shown in appendix Table A4.

Data for Model C

For estimating Model C (equation 29), I construct Young-Old[49] relative wages and hours supplied for education-year groups. Since I am interested in the substitutability of workers at the start and end of their work cycle so I define young worker to be between ages 19 and 25 and old worker to be between ages 61 and 65. I use five education groups.

Definition of Cells

Each cell contains an education group e for a time period t. I create five education groups [High School Dropout (HSD), High School (HS), Some College but no degree (ColND), College (Col), and College Post Graduate Degree (ColPG)] and seven year groups with a three years interval. This results in thirty five education-year groups /cells for which I calculate the relative wages and supplies. Each of these cells forms a unit of observation in the estimation of Model C.

Relative Wages

I create an old dummy equal to 1 if a worker belongs to old group (ages 61 to 65) and 0 if she belonged to young group (19-25). The estimates for young-old log relative earnings are obtained by regressing log of real earnings on old dummy for each education-year group

[49] In Model C, I have age groups 1 and 2 representing Young and Old without any order. Here I construct wages of old relative to young. But the order should not matter.

for Model C. A dummy for white ethnicity is included in the regression as well. The coefficient on the old dummy gives the relative earnings. The log relative earnings are shown in appendix Table A5.

Relative Hours Supplied

I construct "effective" hours supplied by multiplying annual hours supplied by workers in an education group with the relative wage of that education group with respect to college wages. Annual hours supplied by Old workers in an education-year cell is simply the sum of annual hours supplied by all workers in that cell. Similarly, annual hours supplied by Young workers in an education year cell is simply the sum of annual hours supplied by all workers in that cell. Relative supply is simply a ratio of these two sums. Old-Young log relative supplies are shown in appendix Table A6.

Section A3. Tables

Table A 1. Model A - College-HS log relative wages $\left(\log\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right)\right)$

		Age Groups						
		26-30	31-35	36-40	41-45	46-50	51-55	56-60
Time Periods	1996-1998	0.310	0.423	0.432	0.374	0.357	0.380	0.368
	1999-2001	0.387	0.444	0.469	0.479	0.453	0.471	0.474
	2002-2004	0.386	0.464	0.547	0.534	0.482	0.429	0.432
	2005-2007	0.361	0.501	0.495	0.550	0.522	0.454	0.432
	2008-2010	0.360	0.505	0.518	0.552	0.545	0.479	0.445
	2011-2013	0.368	0.455	0.528	0.540	0.573	0.527	0.443
	2014-2016	0.409	0.480	0.528	0.522	0.545	0.514	0.472

Table A 2. Model A - College-HS log relative hours supplied $\left(\log\left(\frac{C_{jt}}{H_{jt}}\right)\right)$

		Age Groups						
		26-30	31-35	36-40	41-45	46-50	51-55	56-60
Time Periods	1996-1998	-0.469	-0.603	-0.658	-0.554	-0.416	-0.646	-0.874
	1999-2001	-0.487	-0.510	-0.627	-0.587	-0.428	-0.438	-0.738
	2002-2004	-0.482	-0.405	-0.503	-0.495	-0.390	-0.287	-0.495
	2005-2007	-0.474	-0.354	-0.368	-0.440	-0.429	-0.293	-0.332
	2008-2010	-0.356	-0.222	-0.199	-0.349	-0.367	-0.351	-0.257
	2011-2013	-0.168	-0.124	-0.111	-0.149	-0.266	-0.318	-0.259
	2014-2016	-0.130	-0.051	-0.081	-0.074	-0.153	-0.267	-0.314

Table A 3. Model B - College-HS log relative wages $\left(\boldsymbol{log}\left(\frac{w_{it}^{c}}{w_{it}^{h}}\right)\right)$

		Age Groups	
		Young (19-25)	Old (61-65)
Time Periods	1996	0.119	0.314
	1997	0.233	0.367
	1998	0.198	0.51
	1999	0.149	0.344
	2000	0.223	0.316
	2001	0.248	0.597
	2002	0.304	0.46
	2003	0.155	0.436
	2004	0.219	0.442
	2005	0.231	0.477
	2006	0.241	0.471
	2007	0.237	0.329
	2008	0.25	0.433
	2009	0.262	0.443
	2010	0.28	0.343
	2011	0.271	0.348
	2012	0.348	0.367
	2013	0.306	0.355
	2014	0.304	0.345
	2015	0.363	0.371
	2016	0.355	0.425

Table A 4. Model B - College-HS log relative hours supplied $\left(log\left(\frac{C_{it}}{H_{it}}\right)\right)$

		Age Groups	
		Young (19-25)	Old (61-65)
Time Periods	1996	-1.139	-0.573
	1997	-1.189	-0.646
	1998	-1.210	-0.614
	1999	-1.249	-0.647
	2000	-1.269	-0.502
	2001	-1.189	-0.499
	2002	-1.253	-0.470
	2003	-1.257	-0.368
	2004	-1.225	-0.363
	2005	-1.190	-0.284
	2006	-1.181	-0.262
	2007	-1.154	-0.165
	2008	-1.107	-0.070
	2009	-1.042	0.028
	2010	-0.973	0.084
	2011	-0.935	0.074
	2012	-0.910	0.103
	2013	-0.885	0.123
	2014	-0.875	0.124
	2015	-0.916	0.082
	2016	-0.937	0.054

Table A 5. Model C - Old-Young log relative wage $\left(log\left(\frac{w_{e1t}}{w_{e2t}}\right)\right)$

		Education groups				
		HSD	HS	ColND	Col	ColPG
Time Periods	1996-1998	0.437	0.568	0.659	0.67	0.588
	1999-2001	0.399	0.482	0.6	0.579	0.76
	2002-2004	0.417	0.507	0.658	0.603	0.592
	2005-2007	0.395	0.514	0.724	0.606	0.837
	2008-2010	0.427	0.552	0.667	0.601	0.526
	2011-2013	0.386	0.565	0.731	0.523	0.657
	2014-2016	0.392	0.529	0.722	0.491	0.631

Table A 6. Model C - Old-Young log relative hours supplied $\left(log\left(\frac{X_{e1t}}{X_{e2t}}\right)\right)$

		Education groups				
		HSD	HS	ColND	Col	ColPG
Time Periods	1996-1998	-0.868	-1.247	-1.969	-1.213	1.345
	1999-2001	-1.055	-1.278	-1.807	-1.097	1.266
	2002-2004	-1.170	-1.267	-1.821	-0.910	1.539
	2005-2007	-1.187	-1.146	-1.717	-0.779	1.573
	2008-2010	-1.058	-0.993	-1.448	-0.491	1.806
	2011-2013	-0.700	-0.771	-1.158	-0.274	1.898
	2014-2016	-0.483	-0.745	-1.111	-0.217	1.761

Table A 7. Model A – Estimates from First Stage

	(1)	(2)
	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$
$log\left(\frac{C_{jt}}{H_{jt}}\right)$	-0.162*** (0.045)	-0.163*** (0.057)
age28	0.186*** (0.030)	0.219*** (0.037)
age33	0.292*** (0.028)	0.320*** (0.036)
age38	0.320*** (0.030)	0.345*** (0.039)
age43	0.323*** (0.031)	0.343*** (0.040)
age48	0.317*** (0.029)	0.333*** (0.039)
age53	0.282*** (0.030)	0.294*** (0.041)
age58	0.239*** (0.034)	0.247*** (0.047)
y2000	0.086*** (0.014)	
y2003	0.117*** (0.016)	
y2006	0.131*** (0.017)	
y2009	0.158*** (0.019)	
y2012	0.178*** (0.023)	
y2015	0.191*** (0.025)	
trend		0.004*** (0.001)
Observations	49	49
R-Squared	0.998	0.996

Notes: Model estimated using simple OLS.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 8. Model A - Estimates from Second Stage

	(1)	(2)
	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$
$log\left(\frac{C_t}{H_t}\right)$	-0.547 (0.365)	-0.522 (0.375)
$\left[log\left(\frac{C_{jt}}{H_{jt}}\right) - log\left(\frac{C_t}{H_t}\right)\right]$	-0.154** (0.057)	-0.154** (0.058)
trend	0.008** (0.004)	0.008* (0.004)
age28	0.142* (0.081)	0.147* (0.083)
age33	0.239*** (0.085)	0.244*** (0.087)
age38	0.259*** (0.089)	0.265*** (0.091)
age43	0.254*** (0.093)	0.260*** (0.095)
age48	0.239** (0.096)	0.246** (0.099)
age53	0.196* (0.101)	0.202* (0.103)
age58	0.146 (0.106)	0.153 (0.109)
y2009		0.005 (0.014)
Observations	49	49
R-Squared	0.996	0.996

Notes: Model estimated using simple OLS.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 9. Estimates for Model B

	(1)	(2)
	$log\left(\frac{w_{it}^{c}}{w_{it}^{h}}\right)$	$log\left(\frac{w_{it}^{c}}{w_{it}^{h}}\right)$
$log\left(\frac{C_{it}}{H_{it}}\right)$	-0.292** (0.110)	-0.287*** (0.088)
young	-0.237 (0.148)	-0.202 (0.130)
old	0.170** (0.064)	-0.063 (0.123)
y1997	0.066 (0.062)	
y1998	0.121* (0.062)	
y1999	0.003 (0.062)	
y2000	0.044 (0.061)	
y2001	0.210*** (0.061)	
y2002	0.164** (0.061)	
y2003	0.092 (0.061)	
y2004	0.132** (0.062)	
y2005	0.172** (0.063)	
y2006	0.179** (0.063)	
y2007	0.124* (0.065)	
y2008	0.203*** (0.068)	
y2009	0.238*** (0.072)	
y2010	0.215** (0.076)	
y2011	0.217** (0.077)	
y2012	0.273*** (0.079)	
y2013	0.253*** (0.080)	
y2014	0.248*** (0.081)	

y2015	0.279*** (0.078)	
y2016	0.295*** (0.076)	
trend		0.013*** (0.003)
Observations	42	42
R-Squared	0.986	0.972

Notes: Model estimated using simple OLS.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 10. Estimates for Model C

	(1)	(2)	(3)	(4)
	$log\left(\frac{w_{e1t}}{w_{e2t}}\right)$	$log\left(\frac{w_{e1t}}{w_{e2t}}\right)$	$log\left(\frac{w_{e1t}}{w_{e2t}}\right)$	$log\left(\frac{w_{e1t}}{w_{e2t}}\right)$
$log\left(\frac{X_{e1t}}{X_{e2t}}\right)$	-0.176** (0.073)	0.024 (0.019)	-0.033 (0.038)	-0.065 (0.094)
y1997		0.603*** (0.057)		0.756*** (0.132)
y2000		0.583*** (0.057)		0.736*** (0.131)
y2003		0.573*** (0.056)		0.732*** (0.138)
y2006		0.631*** (0.056)		0.796*** (0.144)
y2009		0.565*** (0.055)		0.750*** (0.164)
y2012		0.577*** (0.055)		0.783*** (0.186)
y2015		0.557*** (0.055)		0.766*** (0.190)
HSD			0.377*** (0.042)	-0.413 (0.241)
HS			0.496*** (0.046)	-0.298 (0.253)
ColND			0.628*** (0.064)	-0.182 (0.301)
Col			0.558*** (0.035)	-0.224 (0.220)
ColPG			0.709*** (0.064)	
Observations	35	35	35	35
R-Squared	0.145	0.966	0.991	0.992

Notes: Model estimated using simple OLS.Standard errors in parentheses* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 11. Model A Robustness Check 1 - College-HS log relative wages $\left(\boldsymbol{log}\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right)\right)$

		Age Groups								
		19-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65
Time Periods	1996-1998	0.181	0.31	0.423	0.432	0.374	0.357	0.38	0.368	0.4
	1999-2001	0.209	0.387	0.444	0.469	0.479	0.453	0.471	0.474	0.416
	2002-2004	0.226	0.386	0.464	0.547	0.534	0.482	0.429	0.432	0.446
	2005-2007	0.238	0.361	0.501	0.495	0.55	0.522	0.454	0.432	0.424
	2008-2010	0.264	0.36	0.505	0.518	0.552	0.545	0.479	0.445	0.404
	2011-2013	0.308	0.368	0.455	0.528	0.54	0.573	0.527	0.443	0.353
	2014-2016	0.34	0.409	0.48	0.528	0.522	0.545	0.514	0.472	0.381

Table A 12. Model A Robustness Check 1 - College-HS log relative hours supplied $\left(\boldsymbol{log}\left(\frac{C_{jt}}{H_{jt}}\right)\right)$

		Age Groups								
		19-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65
Time Periods	1996-1998	-1.178	-0.349	-0.433	-0.472	-0.331	-0.165	-0.361	-0.605	-0.612
	1999-2001	-1.236	-0.363	-0.340	-0.437	-0.374	-0.189	-0.172	-0.454	-0.547
	2002-2004	-1.245	-0.346	-0.218	-0.311	-0.286	-0.149	-0.022	-0.206	-0.398
	2005-2007	-1.175	-0.330	-0.151	-0.153	-0.236	-0.202	-0.031	-0.058	-0.236
	2008-2010	-1.042	-0.196	-0.001	0.029	-0.124	-0.144	-0.101	0.010	0.017
	2011-2013	-0.910	-0.008	0.110	0.129	0.094	-0.030	-0.072	0.002	0.100
	2014-2016	-0.909	0.044	0.192	0.178	0.177	0.090	-0.019	-0.058	0.087

Table A 13. Estimates for Model A – Robustness Check 1

	(1)	(2)
	$log\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right)$	$log\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right)$
$log\left(\frac{C_{jt}}{H_{jt}}\right)$	-0.181*** (0.039)	-0.194*** (0.043)
age22	-0.064 (0.054)	-0.058 (0.060)
age28	0.211*** (0.023)	0.226*** (0.025)
age33	0.328*** (0.020)	0.341*** (0.022)
age38	0.358*** (0.021)	0.367*** (0.024)
age43	0.362*** (0.021)	0.367*** (0.024)
age48	0.359*** (0.020)	0.361*** (0.023)
age53	0.327*** (0.019)	0.326*** (0.023)
age58	0.285*** (0.022)	0.280*** (0.027)
age63	0.245*** (0.023)	0.236*** (0.028)
y2000	0.072*** (0.014)	
y2003	0.107*** (0.015)	
y2006	0.122*** (0.016)	
y2009	0.153*** (0.019)	
y2012	0.175*** (0.022)	
y2015	0.193*** (0.023)	
trend		0.003*** (0.000)
Observations	63	63
R-Squared	0.997	0.996

Notes: Model estimated using simple OLS.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 14. Model B Robustness Check 2 - College-HS log relative wages $\left(log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)\right)$

		Age Groups							
		Young Old Aggregator	26-30	31-35	36-40	41-45	46-50	51-55	56-60
Time Periods	1996	0.213	0.326	0.419	0.396	0.374	0.332	0.400	0.382
	1997	0.276	0.286	0.438	0.404	0.345	0.322	0.359	0.333
	1998	0.449	0.332	0.396	0.474	0.419	0.418	0.399	0.417
	1999	0.284	0.350	0.449	0.473	0.456	0.476	0.430	0.445
	2000	0.213	0.394	0.443	0.458	0.492	0.463	0.477	0.484
	2001	0.531	0.413	0.470	0.469	0.487	0.451	0.478	0.461
	2002	0.386	0.416	0.464	0.565	0.556	0.492	0.411	0.472
	2003	0.239	0.395	0.462	0.561	0.515	0.484	0.456	0.425
	2004	0.334	0.319	0.459	0.536	0.541	0.445	0.427	0.394
	2005	0.332	0.369	0.468	0.479	0.585	0.551	0.437	0.469
	2006	0.308	0.368	0.493	0.509	0.534	0.524	0.431	0.415
	2007	0.143	0.364	0.532	0.540	0.494	0.526	0.448	0.450
	2008	0.190	0.378	0.471	0.541	0.520	0.509	0.447	0.451
	2009	0.248	0.356	0.509	0.510	0.540	0.566	0.474	0.484
	2010	0.152	0.396	0.497	0.527	0.606	0.592	0.478	0.427
	2011	0.143	0.355	0.475	0.534	0.543	0.581	0.497	0.477
	2012	0.222	0.356	0.479	0.538	0.540	0.574	0.523	0.407
	2013	0.172	0.425	0.403	0.518	0.566	0.534	0.576	0.464
	2014	0.203	0.382	0.487	0.487	0.585	0.521	0.508	0.434
	2015	0.276	0.387	0.452	0.570	0.504	0.563	0.520	0.476
	2016	0.287	0.434	0.508	0.539	0.495	0.542	0.474	0.525

Table A 15. Model B Robustness Check 2 - College-HS log relative hours supplied $\left(log\left(\frac{C_{it}}{H_{it}}\right)\right)$

		Age Groups							
		Young Old Aggregator	26-30	31-35	36-40	41-45	46-50	51-55	56-60
Time Periods	1996	-0.930	-0.372	-0.460	-0.477	-0.277	-0.191	-0.439	-0.672
	1997	-0.988	-0.348	-0.447	-0.489	-0.340	-0.159	-0.355	-0.581
	1998	-0.979	-0.320	-0.405	-0.440	-0.364	-0.157	-0.294	-0.540
	1999	-1.017	-0.356	-0.353	-0.445	-0.334	-0.174	-0.206	-0.490
	2000	-0.958	-0.360	-0.348	-0.456	-0.367	-0.184	-0.146	-0.453
	2001	-0.916	-0.348	-0.330	-0.413	-0.418	-0.223	-0.161	-0.403
	2002	-0.940	-0.338	-0.210	-0.364	-0.318	-0.106	-0.046	-0.294
	2003	-0.884	-0.351	-0.243	-0.309	-0.279	-0.151	-0.011	-0.211
	2004	-0.852	-0.346	-0.191	-0.259	-0.282	-0.152	-0.016	-0.137
	2005	-0.793	-0.328	-0.160	-0.228	-0.266	-0.206	-0.016	-0.109
	2006	-0.776	-0.344	-0.154	-0.160	-0.233	-0.217	-0.017	-0.066
	2007	-0.708	-0.304	-0.146	-0.088	-0.202	-0.196	-0.029	-0.035
	2008	-0.618	-0.243	-0.034	-0.023	-0.166	-0.148	-0.081	0.013
	2009	-0.520	-0.196	0.009	0.033	-0.112	-0.120	-0.107	0.018
	2010	-0.446	-0.116	0.039	0.067	-0.082	-0.147	-0.119	-0.026
	2011	-0.417	-0.062	0.070	0.105	0.042	-0.077	-0.093	0.018
	2012	-0.380	-0.011	0.122	0.146	0.093	-0.039	-0.081	-0.004
	2013	-0.361	0.011	0.166	0.142	0.136	0.034	-0.032	-0.023
	2014	-0.352	0.023	0.195	0.181	0.142	0.036	-0.010	-0.047
	2015	-0.391	0.063	0.165	0.146	0.173	0.082	0.004	-0.069
	2016	-0.414	0.019	0.225	0.230	0.201	0.168	-0.019	-0.093

Table A 16. Second Stage Estimates for Model B – Robustness Check 2

	(1)	(2)
	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$
$log\left(\frac{C_{jt}}{H_{jt}}\right)$	-0.216*** (0.041)	-0.234*** (0.042)
y1997	-0.007 (0.024)	
y1998	0.066*** (0.024)	
y1999	0.077*** (0.024)	
y2000	0.088*** (0.025)	
y2001	0.131*** (0.025)	
y2002	0.147*** (0.025)	
y2003	0.124*** (0.025)	
y2004	0.119*** (0.026)	
y2005	0.152*** (0.026)	
y2006	0.143*** (0.026)	
y2007	0.139*** (0.027)	
y2008	0.151*** (0.028)	
y2009	0.182*** (0.028)	
y2010	0.185*** (0.029)	
y2011	0.187*** (0.030)	
y2012	0.199*** (0.031)	
y2013	0.207*** (0.032)	
y2014	0.203*** (0.032)	
y2015	0.221*** (0.032)	
y2016	0.232*** (0.032)	
young old aggregator	-0.024 (0.044)	-0.012 (0.045)
age28	0.183*** (0.028)	0.204*** (0.026)
age33	0.299*** (0.025)	0.322*** (0.023)
age38	0.334*** (0.026)	0.356*** (0.024)
age43	0.336*** (0.026)	0.357*** (0.024)

age48	0.334*** (0.025)	0.357*** (0.022)
age53	0.296*** (0.025)	0.318*** (0.022)
age58	0.259*** (0.027)	0.280*** (0.025)
trend		0.011*** (0.001)
Observations	168	168
R-Squared	0.990	0.988

Notes: Model estimated using simple OLS.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 17. Third Stage Estimates for Model B – Robustness Check 2

	(1)	(2)
	$log\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right)$	$log\left(\frac{w_{jt}^{c}}{w_{jt}^{h}}\right)$
$log\left(\frac{C_t}{H_t}\right)$	-0.283 (0.204)	-0.302 (0.213)
$\left[log\left(\frac{C_{jt}}{H_{jt}}\right) - log\left(\frac{C_t}{H_t}\right)\right]$	-0.233*** (0.043)	-0.234*** (0.043)
young old aggregator	0.015 (0.121)	0.023 (0.125)
age28	0.231** (0.112)	0.239** (0.117)
age33	0.349*** (0.111)	0.357*** (0.115)
age38	0.383*** (0.111)	0.391*** (0.116)
age43	0.384*** (0.112)	0.393*** (0.116)
age48	0.383*** (0.111)	0.392*** (0.115)
age53	0.345*** (0.111)	0.354*** (0.115)
age58	0.307*** (0.112)	0.315*** (0.116)
trend	0.012** (0.005)	0.012** (0.006)
y2007		-0.012 (0.019)
y2008		-0.009 (0.019)
y2009		0.013 (0.019)
Observations	168	168
R-Squared	0.988	0.988

Notes: Model estimated using simple OLS. Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 18. Estimates for Model B - Robustness Check 3

	(1)	(2)
	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$	$log\left(\frac{w_{jt}^c}{w_{jt}^h}\right)$
$log\left(\frac{C_{jt}}{H_{jt}}\right)$	-0.142*** (0.030)	-0.148*** (0.034)
y1997	-0.017 (0.018)	
y1998	0.040** (0.018)	
y1999	0.075*** (0.018)	
y2000	0.095*** (0.018)	
y2001	0.098*** (0.018)	
y2002	0.131*** (0.018)	
y2003	0.123*** (0.018)	
y2004	0.101*** (0.019)	
y2005	0.136*** (0.019)	
y2006	0.127*** (0.019)	
y2007	0.142*** (0.019)	
y2008	0.143*** (0.020)	
y2009	0.165*** (0.020)	
y2010	0.178*** (0.021)	
y2011	0.178*** (0.021)	
y2012	0.176*** (0.022)	
y2013	0.190*** (0.023)	
y2014	0.180*** (0.023)	
y2015	0.190*** (0.023)	
y2016	0.200*** (0.023)	
age28	0.214*** (0.020)	0.236*** (0.020)
age33	0.322*** (0.018)	0.345*** (0.017)
age38	0.359*** (0.019)	0.381*** (0.018)
age43	0.361*** (0.019)	0.384*** (0.018)
age48	0.356*** (0.018)	0.379*** (0.017)
age53	0.318*** (0.018)	0.341*** (0.017)

age58	0.288*** (0.020)	0.310*** (0.020)
trend		0.009*** (0.001)
Observations	147	147
R-Squared	0.996	0.994

Notes: Model estimated using simple OLS.
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A 19. Estimates from Robustness Checks 4 and 5 for Model C

	(1)	(2)	(3)
	ln_wage	ln_wage	ln_wage
ln_supply	-0.034 (0.019)	-0.099* (0.041)	-0.003 (0.018)
Observations	245	70	280
R-Squared	1.000	1.000	1.000

Notes: Each column shows results from a separate OLS regression. Column 1 shows the results from the fourth robustness check. Columns 2 and 3 show results from the first and second stage of the last robustness check respectively. Each column contains (educ x year) and (age x educ) dummies. Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Section A4. Derivation of Standard Error for the Estimate of Elasticity of Substitution

The regression equations (8, 7b, 22, 29) obtained from the models estimate the *reciprocal* of the elasticity of substitution. Standard errors obtained from these OLS regressions are also for the estimates of the *reciprocal* of elasticity of substitution $\left({}^{1}\!/_{\sigma}\right)$. In this section, I derive the formula of the standard error for the estimate of the elasticity of substitution itself (σ) using the Delta method.

Let $\beta = {}^{1}\!/_{\sigma}$ and $\hat{\beta} \sim N\left(\beta, \frac{s^2}{n}\right)$ be the OLS estimate of β, where s^2 is the variance of the error term (I assume homoscedasticity). I want to estimate the standard error for $f(\hat{\beta}) = \frac{1}{\hat{\beta}} = \hat{\sigma}$. First order linear approximation of $f(\hat{\beta})$ around mean value of $\hat{\beta}$ (i.e. β) is:

$$f(\hat{\beta}) \approx f(\beta) + f'(\beta)[\hat{\beta} - \beta] = \frac{1}{\beta} - \frac{1}{\beta^2}[\hat{\beta} - \beta] = \frac{2}{\beta} - \frac{\hat{\beta}}{\beta^2}$$

where we assume that $f'(\beta) \neq 0$.

Mean, Variance, and Standard Error of $f(\hat{\beta})$ are:

$$E[f(\hat{\beta})] \approx \frac{1}{\beta}$$

$$Var[f(\hat{\beta})] = E\left[f(\hat{\beta}) - E[f(\hat{\beta})]\right]^2$$

$$\approx E\left[\frac{2}{\beta} - \frac{\hat{\beta}}{\beta^2} - \frac{1}{\beta}\right]^2$$

$$= E\left[\frac{1}{\beta} - \frac{\hat{\beta}}{\beta^2}\right]^2 = E\left[\frac{\beta - \hat{\beta}}{\beta^2}\right]^2$$

$$= \frac{1}{\beta^4}E[\hat{\beta} - \beta]^2 = \frac{Var(\hat{\beta})}{\beta^4}$$

$$SE\left(f(\hat{\beta})\right) \approx \frac{SE(\hat{\beta})}{\beta^2}$$

Therefore, the Standard Error of $(\hat{\sigma})$ is:

$$SE(\hat{\sigma}) \approx \frac{SE\left[\widehat{\frac{1}{\sigma}}\right]}{\left[\frac{1}{\sigma}\right]^2} \qquad (33)$$

where $SE\left[\widehat{\frac{1}{\sigma}}\right]$ is the standard error obtained from OLS regression equations (8, 7b, 22, 29). And for $\left[\frac{1}{\sigma}\right]$ we can use the estimate $\left[\widehat{\frac{1}{\sigma}}\right]$ obtained from OLS regressions (8, 7b, 22, 29) because of the consistency of OLS. Using (33) I calculate the standard errors and 95% confidence intervals for $(\hat{\sigma})$.

Section A5. Three Step Procedure for Model B

The three step procedure of Model B is similar to the two step procedure of Model A. In the first step once we obtain $(\sigma_{1,2})$ from equation (22), then using equations (15) to (18) we can estimate technological efficiency parameters $(\gamma_1, \gamma_2, \xi_1, \xi_2)$. I normalize $\gamma_2$and ξ_2 to 1 and take the ratios of wages of age group 1 (old) to wages of age group 2 (young) for high school and college categories separately. Equations (15) to (18) imply that:

$$\frac{w_{1t}^h}{w_{2t}^h} = \frac{\frac{\partial q_t}{\partial H_{1t}}}{\frac{\partial q_t}{\partial H_{2t}}} = \left(\frac{\gamma_1}{\gamma_2}\right)\left(\frac{H_{1t}}{H_{2t}}\right)^{\varphi-1} \qquad (34)$$

$$\frac{w_{1t}^c}{w_{2t}^c} = \frac{\frac{\partial q_t}{\partial C_{1t}}}{\frac{\partial q_t}{\partial C_{2t}}} = \left(\frac{\xi_1}{\xi_2}\right)\left(\frac{C_{1t}}{C_{2t}}\right)^{\varphi-1} \qquad (35)$$

Taking logarithm of (34) and (35) and noting that $\varphi - 1 = \frac{1}{\sigma_{1,2}}$; $\gamma_2 = 1$; and $\xi_2 = 1$ yields:

$$ln\left(\frac{w_{1t}^h}{w_{2t}^h}\right) = ln(\gamma_1) - \frac{1}{\sigma_{1,2}} ln\left(\frac{H_{1t}}{H_{2t}}\right) \qquad (36)$$

$$ln\left(\frac{w_{1t}^c}{w_{2t}^c}\right) = ln(\xi_1) - \frac{1}{\sigma_{1,2}} ln\left(\frac{C_{1t}}{C_{2t}}\right) \qquad (37)$$

Since $(\sigma_{1,2})$ is already estimated so $ln\left(\frac{w_{1t}^h}{w_{2t}^h}\right)$, $ln\left(\frac{w_{1t}^c}{w_{2t}^c}\right)$, $\frac{1}{\sigma_{1,2}} ln\left(\frac{H_{1t}}{H_{2t}}\right)$, and $\frac{1}{\sigma_{1,2}} ln\left(\frac{C_{1t}}{C_{2t}}\right)$ are known. Therefore the constant terms of the regressions based on (36) and (37) give estimates for $\gamma_1$and ξ_1. Once we know $(\sigma_{1,2}, \gamma_1, \gamma_2, \xi_1, \xi_2)$ then using (12) and (14) we can construct A_t and B_t and obtain college-high school relative supplies for the young-old aggregator term by taking the ratio of B_t to A_t. This ends the first step.

The second and third steps of Model B are simply the first and second steps of Model A respectively with an extra age group in the form of young-old aggregator. This means that we can estimate (σ_A) in the second step in Model B using on a regression similar to equation (8). After obtaining (σ_A), we can find technological efficiency parameters of age groups (including the aggregator term) and then construct $H_t$and C_t. This ends the second step. In the third step we can obtain (σ_E) for Model B using a regression similar to equation (7b). The only challenge is to construct college-high school relative wages for the young-old aggregator term, which is dealt with using some manipulations.

The marginal products of aggregate output q_t with respect to A_t and B_t are:

$$w_{At} \equiv \frac{\partial q_t}{\partial A_t} = \frac{\partial q_t}{\partial H_t} \times \frac{\partial H_t}{\partial A_t} = q_t^{1-\rho} \theta_{ht} H_t^{\rho-\eta} \times \alpha_a A_t^{\eta-1} \qquad (38)$$

$$w_{Bt} \equiv \frac{\partial q_t}{\partial B_t} = \frac{\partial q_t}{\partial H_t} \times \frac{\partial H_t}{\partial B_t} = q_t^{1-\rho} \theta_{ct} C_t^{\rho-\eta} \times \beta_b B_t^{\eta-1} \qquad (39)$$

The problem is that A_t and B_t are mathematical constructs consisting of young and old age groups assumed to be *imperfect substitutes*. So we cannot simply take the average wage of all young and old workers in high school and college category and argue that these wages are the same as (38) and (39) respectively. What we do observe are wages for young group and old group separately and somehow have to combine these two wages, using appropriate weights, to obtain (w_{At}) and (w_{Bt}). I show the derivations for (w_{At}). For (w_{Bt}) the derivations are similar.

The marginal products of aggregate output q_t with respect to H_{1t} and H_{2t} are:

$$w_{1t}^h \equiv \frac{\partial q_t}{\partial H_{1t}} = \frac{\partial q_t}{\partial H_t} \times \frac{\partial H_t}{\partial A_t} \times \frac{\partial A_t}{\partial H_{1t}} = \frac{\partial q_t}{\partial A_t} \times \frac{\partial A_t}{\partial H_{1t}} = \frac{\partial q_t}{\partial A_t} \times A_t^{1-\varphi} \gamma_1 H_{1t}^{\varphi-1} \qquad (40)$$

$$w_{2t}^h \equiv \frac{\partial q_t}{\partial H_{2t}} = \frac{\partial q_t}{\partial H_t} \times \frac{\partial H_t}{\partial A_t} \times \frac{\partial A_t}{\partial H_{2t}} = \frac{\partial q_t}{\partial A_t} \times \frac{\partial A_t}{\partial H_{2t}} = \frac{\partial q_t}{\partial A_t} \times A_t^{1-\varphi} \gamma_2 H_{2t}^{\varphi-1} \qquad (41)$$

where w_{1t}^h and w_{2t}^h are average wages for young and old groups in high school category and time period t. These are easily obtained from data.

Now using 'appropriate' weights we can write w_{At} as the sum of w_{1t}^h and w_{2t}^h:

$$\left(\frac{H_{1t}}{A_t}\right) w_{1t}^h + \left(\frac{H_{2t}}{A_t}\right) w_{2t}^h$$

$$= \frac{\partial q_t}{\partial A_t} A_t^{-\varphi} \gamma_1 H_{1t}^{\varphi} + \frac{\partial q_t}{\partial A_t} A_t^{-\varphi} \gamma_2 H_{2t}^{\varphi}$$

$$= \frac{\partial q_t}{\partial A_t} A_t^{-\varphi} [\gamma_1 H_{1t}^{\varphi} + \gamma_2 H_{2t}^{\varphi}]$$

$$= \frac{\partial q_t}{\partial A_t} A_t^{-\varphi} A_t^{\varphi}$$

$$= \frac{\partial q_t}{\partial A_t} \equiv w_{At} \qquad (42)$$

Similarly we have that:

$$\left(\frac{C_{1t}}{B_t}\right) w_{1t}^c + \left(\frac{C_{2t}}{B_t}\right) w_{2t}^c = w_{Bt} \qquad (43)$$

It is easy to note that $\left(\frac{H_{1t}}{A_t}\right), w_{1t}^h, \left(\frac{H_{2t}}{A_t}\right), w_{2t}^h, \left(\frac{C_{1t}}{B_t}\right), w_{1t}^c, \left(\frac{C_{2t}}{B_t}\right)$ $and\ w_{2t}^c$ are all observables and hence using (42) and (43) we can now constructs relative wages for the aggregator term by simply taking the ratio of w_{Bt} to w_{At}. In addition, the weights in (42) (i.e. $\left(\frac{H_{1t}}{A_t}\right)$ and $\left(\frac{H_{2t}}{A_t}\right)$) do not add up to 1. The same is true for the weights in (43). For implementing the second and third stage, which are equations similar to (8) and (7b) respectively, I first calculate average wages w_{At} and w_{Bt}, then take their ratio and finally take their logarithm.

Section A6. Robustness Checks for Model C

I do two robustness checks for Model C. In the first one, I use the estimation method of Ottaviano and Peri (2012) in order to estimate Model C without the young-old aggregator term and without young and old age groups .This yields an estimate of (σ_A) within more disaggregated (or more homogenous) skill groups as compared to Models A and B. In the second exercise, I use a two-step procedure to estimate Model C, again using the estimation method of Ottaviano and Peri (2012). The first step yields an estimate of $(\sigma_{1,2})$ within more homogenous skill groups as compared to Models A and B. The second step estimates (σ_A). Comparing (σ_A) with $(\sigma_{1,2})$ in these two exercises will lend further support to my first conclusion. Comparing $(\sigma_{1,2})$ from two step procedure of Model C with $(\sigma_{1,2})$ from three step procedure of Model B will further strengthen my second conclusion. In the following paragraphs I briefly derive regression equations for these two exercises and will also mention the data construction.

<u>Estimating Model C without Young-Old Aggregator Term (Exercise 4)</u>

The Model and its identification are based on Ottaviano and Peri (2012). Assume aggregate output in the economy has CES form:

$$q_t = \left[\sum_e \theta_{et} L_{et}^{\rho}\right]^{1/\rho} \qquad (44)$$

where L_{et} is the labor supplied by education group e in time period t; $\theta_{et}$is the efficiency parameter of this group; and ρ is a function of the elasticity of substitution between education groups (σ_E). Further assume that:

$$L_{et} = \left[\sum_a \theta_{aet} L_{aet}^{\eta}\right]^{1/\eta} \qquad (45)$$

where L_{aet} is the labor supplied by age group a and education group e in time period t; $\theta_{aet}$is the efficiency parameter of this group; and η is a function of the elasticity of substitution between age groups (σ_A). The marginal product of output with respect to L_{aet} is:

$$w_{aet} = \frac{\partial q_t}{\partial L_{aet}} = \frac{\partial q_t}{\partial L_{et}} \times \frac{\partial L_{et}}{\partial L_{aet}}$$

$$= q_t^{1-\rho}\theta_{et}L_{et}^{\rho-1} \times L_{et}^{1-\eta}\theta_{aet}L_{aet}^{\eta-1}$$

$$= q_t^{1-\rho}\theta_{et}L_{et}^{\rho-\eta} \times \theta_{aet}L_{aet}^{\eta-1}$$

Taking logarithm of both sides:

$$ln(w_{aet}) = \frac{1}{\sigma_E}ln(q_t) + ln(\theta_{et}) + \left[\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right] lnL_{et} + ln(\theta_{aet}) - \frac{1}{\sigma_A}lnL_{aet} \qquad (46)$$

Ottaviano and Peri (2012) argue that (education × year) and (age × education) dummies capture all the term on the right hand side of (46) except the last term. We can construct lnw_{aet} and lnL_{aet} from data and hence estimating (46) by OLS can give us σ_A. Note again that in (46) we only have workers from ages 26 to 60 inclusive.

Estimating Model C Using Two Step Procedure (Exercise 5)

The following is an enriched version of Ottaviano and Peri's model. Basically it is the same as Model C except the estimation method. Here I follow the estimation method by Ottaviano and Peri in order to implement the two steps[50]. As before, assume that the aggregate output has the following CES form:

$$q_t = \left[\sum_e \theta_{et}L_{et}^{\rho}\right]^{1/\rho} \qquad (47)$$

Assume further that:

[50] There are technical issues which do not allow the use of two step procedure using my estimation method.

$$L_{et} = \left[\theta_{xet}X_{et}^{\eta} + \sum_{a \neq y,o} \theta_{aet}L_{aet}^{\eta}\right]^{1/\eta} \qquad (48)$$

$$X_{et} = \left[\theta_{yet}L_{yet}^{\varphi} + \theta_{oet}L_{oet}^{\varphi}\right]^{1/\varphi} \qquad (49)$$

where φ is a function of the elasticity of substitution between young (y) and old (o) workers. Basically (48) and (49) allows young and old to be studied in isolation and to have a different elasticity than other age groups. The marginal product of output with respect to supply of young and old age groups (a=y,o) is:

$$w_{aet} = \frac{\partial q_t}{\partial L_{aet}} = \frac{\partial q_t}{\partial L_{et}} \times \frac{\partial L_{et}}{\partial X_{et}} \times \frac{\partial X_{et}}{\partial L_{aet}}$$

$$= q_t^{1-\rho}\theta_{et}L_{et}^{\rho-1} \times L_{et}^{1-\eta}\theta_{xet}X_{et}^{\eta-1} \times X_{et}^{1-\varphi}\theta_{aet}L_{aet}^{\varphi-1}$$

$$= q_t^{1-\rho}\theta_{et}L_{et}^{\rho-\eta} \times \theta_{xet}X_{et}^{\eta-\varphi} \times \theta_{aet}L_{aet}^{\varphi-1}$$

Taking logarithm of both sides:

$$ln(w_{aet}) = \frac{1}{\sigma_E} ln(q_t) + ln(\theta_{et}) + \left[\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right] lnL_{et} + ln(\theta_{xet}) + \left[\frac{1}{\sigma_{1,2}} - \frac{1}{\sigma_A}\right] lnX_{et}$$

$$+ ln(\theta_{aet}) - \frac{1}{\sigma_{1,2}} lnL_{aet} \qquad (50)$$

where a=y,o i.e. only young and old age groups are included in this equation. Similar to Ottaviano and Peri, I argue that (education × year) and (age × education) dummies capture all the term on the right hand side of (50) except the last term. Hence estimating (50) by OLS gives us $(\sigma_{1,2})$. Ottaviano and Peri also argue that (age × education) dummies in the OLS regression of (50) gives us θ_{aet}, where the assumption is that $\theta_{aet} = \theta_{ae}$. Using estimates for $(\sigma_{1,2})$ and θ_{aet} we can now construct X_{et} using (49). This is the end of the first step.

The second step is the same as estimating equation (46) with an extra age term in the form of young-old aggregator. The key challenge is to obtain wages for this aggregator term, which is resolved in exactly the same manner as in the three step procedure of Model B (see section A5 of the appendix). Estimating a regression equation similar to (46) we can estimate (σ_A) for Model C.

Note that the only difference between Model C described in this section and the Model C described in section 3 is the estimation method. In section 3 I am estimating an equation of ratios, here I am estimating levels. As a result of using levels, I have greater degrees of freedom and as a result more precise estimates.

Data and Variables

The wage and hours supplied sample is the same as before. The difference lies in the definitions of cells and construction of wages. Here I use 7 age groups, 5 education groups, and 7 three-year interval time periods for the fourth robustness check and 8 age groups (one of the age groups is the young-old aggregator) , 5 education groups, and 7 three-year interval time periods for the fifth robustness check. This yield 245 and 280 cells respectively for which I construct wages and hours supplied. Hours supplied is constructed as before. Wages on the other hand are constructed by first finding the average wage for each cell and then taking the logarithm of this average wage. Constructed data for these cells are not included here but are available on request.

References

Autor, D., Katz, L., & Kearney, M. (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *The Review of Economics and Statistics, 90*(2), 300-323.

Autor, D. (2012). Lecture 1: The Theory of Skill Premia [Lecture Notes]. Retrieved from MIT, 14.662 Labor Economics II, Spring 2012.

Borjas, G. J. (1986). The Demographic Determinants of the Demand for Black Labor. Retrieved from http://papers.nber.org/books/free86-1

Borjas, G. J. (2003). The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *The Quarterly Journal of Economics,118*(4), 1335-1374.

Borjas, G. J., Grogger, J., & Hanson, G. H. (2012). Comment: On Estimating Elasticities Of Substition. *Journal of the European Economic Association,10*(1), 198-210.

Card, D., & Lemieux, T. (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The Quarterly Journal of Economics,116*(2), 705-746.

Card, D. E., & Raphael, S. (2013). *Immigration, poverty, and socioeconomic inequality*. New York: Russell Sage Foundation.

Card, D. (2016). Lecture 7: Static Labor Demand [Lecture notes]. Retrieved from University of California Berkeley, Labor Economics, Econ 250A, Fall 2016

Center for Economic and Policy Research. 2016. March CPS Uniform Extracts, Version 1.0. Washington, DC.

Fitzenberger, B. & Kohn, K. (2006). Skill Wage Premia, Employment, and Cohort Effects: Are Workers in Germany All of the Same Type? *IZA Discussion Papers 2185, Institute for the Study of Labor (IZA).*

Grant, J. H., & Hamermesh, D. S. (1981). Labor Market Competition among Youths, White Women and Others. *The Review of Economics and Statistics,63*(3), 354.

Gruber, Jonathan and David A. Wise, eds. 2010. *Social Security Programs and Retirement around the World: The Relationship to Youth Employment.* Chicago, IL and London: The University of Chicago Press.

Guest, R., & Jensen, B. S. (2016). Applying a CRESH aggregate labour index to generate age-wage profiles. *Applied Economics Letters,23*(1), 27-33.

Hamermesh, D. S. (1985). Substitution between different categories of labor, relative wages and youth unemployment. OECD Economic Studies 5, OECD.

Hamermesh, D. S. (1986). Chapter 8 The demand for labor in the long run. *Handbook of Labor Economics,* 429-471.

Katz, L. F., & Murphy, K. M. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics,107*(1), 35-78.

Krusell, P., Ohanian, L. E., Rios-Rull, J., & Violante, G. L. (2000). Capital-skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica,68*(5), 1029-1053.

Lanot, G., & Sousounis, P. (2016). The National Minimum Wage and the Substitutability Between Young and Old Workers in Low Paid Occupations. *The Manchester School*.

Munnell, Alicia H. and April Yanyuan Wu. 2012. "Will Delayed Retirement by the Baby Boomers Lead to Higher Unemployment among Younger Workers?" Working Paper. Chestnut Hill, MA: Center for Retirement Research at Boston College. WP 2012-22.

Nicholson, W., & Snyder, C. (2008). *Microeconomic theory: basic principles and extensions*. Belmont, CA: Thomson Business and Economics.

Ottaviano, G. I., & Peri, G. (2012). Rethinking The Effect Of Immigration On Wages. *Journal of the European Economic Association,10*(1), 152-197.

Welch, F. (1979). Effects of Cohort Size on Earnings: The Baby Boom Babies Financial Bust. *Journal of Political Economy,87*(5, Part 2).