The Age of Innovation: Evidence from an International Panel

by

John Sim

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

August 2012

Acknowledgements

This topic was suggested by James Poterba at the 2012 W.A. Mackintosh Visiting Lectureship in Economics at Queen's University. I would like to thank my supervisor, Gregor Smith, for his valuable guidance as well as my family and friends for their continued support.

Contents

1 Introduction

Western nations including Canada and the United States are in the middle of large population age structure changes due to the postwar baby boom. In Canada, the dependency ratio is expected to rise by nearly 50% between 2009 and 2036 (Statistics Canada, 2010) while the dependency ratio in the United States is expected to rise by nearly 60% over a similar period (United States Census Bureau, 2010). Given these changes, the effect of population age structure on economic variables has begun to receive much attention in the empirical literature. Authors have looked at how age structure affects productivity, international cash flows, and even financial crises.

However, evidence for or against a relationship between age structure and technological progress is noticeably absent. This essay attempts to take steps towards filling that gap. Macroeconomic models are sensitive to changes in the technological progress parameter, meaning a small change in its magnitude can result in drastically different forecasts. If older age groups innovate more, the burden of the higher dependency ratio will be lessened. Conversely, if younger age groups innovate more, the burden may be even greater than expected (Poterba, 2012).

It is not clear how or even if age structure should affect technological progress. Take for example two workforces: one young and one old. Can we expect one workforce to innovate more than the other? Two important inputs to innovation are investment and human capital, and the older workforce likely leads in both. If we define human capital as the sum of education and experience, the vast experience of the older workforce gives them the upper hand. Additionally, savings generally peak later in life.

Although lacking in experience, the young workforce may possess the most "up-to-date" human capital and be more prone to risk taking. This combination may make them the most suited to innovation. Of course, the final possibility is that neither workforce will innovate more than the other. Innovation may display Leontief properties whereby the the rate of innovation is constrained by the scarce input —be it investment, "up-to-date" human capital, risk taking, or experience.

I look for a relationship between age structure and technological growth using a panel of countries assembled from a variety of sources. I estimate with two different measures of technological growth: patent application counts and a computed measure. Using the computed measure as the regressand, I find little evidence of an age structure effect. However, results from the patent applications proxy give strong evidence of a life cycle hypothesis. Innovation gradually rises, reaching a peak at the 50–54 age group, before declining rapidly. These results hold after addressing issues of multicollinearity and endogeneity, and are consistent with the related studies of Feyrer (2007) and Lindh and Malmberg (1999).

The rest of this essay is organized as follows: section 2 reviews the related research, section 3 describes the data, section 4 presents the statistical methods and results, section 5 discusses, and section 6 concludes.

2 Related Research

Empirical research addressing the link between technological progress and age structure is notably sparse. However, many related studies have been conducted. Lindh and Malmberg (1999) look for a link between age structure and overall economic growth. They modify the transitional growth model of Mankiw *et al* to allow for workforce experience by interacting an index of age structure with the stock of educational capital. Measuring age structure in cohorts of fifteen years in length, they find that an increase in the proportion of workers aged 50–64 has a statistically significant positive impact on economic growth while an increase in the proportion of younger workers affects growth ambiguously. Feyrer (2007) addresses the link between productivity and age structure. He assumes Cobb-Douglas production and calculates productivity as the Solow residual. Measuring age structure as proportional cohorts and estimating in first differences, he finds that an increase in the share of workers aged 40–60 has a positive impact on productivity.

In addition to measuring productivity, the Solow residual has been used as a measure of technological progress. But this interpretation is only valid under the assumption of efficient use of all resources in all years. In the presence of competitive markets, marginal products will equal prices and resources will be employed efficiently. In reality, however, imperfections are prevalent and sometimes even built into the market; pay for seniority is a good example of this (Lindh and Prskawetz, 2006). As a result, the Solow residual actually contains two components: technological progress and efficiency; and a change in the Solow residual may be attributed to either (Basu and Fernald, 2002).

In light of the strong assumptions on the Solow residual, authors have turned to proxy variables and deterministic frontier models to measure technological growth (Growiec, 2012). Frontier models relax the efficiency assumption of growth accounting by using non-parametric, data envelopment analysis (DEA) algorithms to compute the world technology or world production frontiers. Inefficiency is measured as the distance to the frontier, technological progress as shifts in the frontier, and capital accumulation as movement along the frontier. Growiec (2012) implements DEA algorithms to construct 12 measures of technological progress. The measures differ in their underlying assumptions (*i.e.* constant returns to scale vs. variable returns to scale) and information sets. Kumar and Russell (2002) similarly employ non-parametric techniques to decompose productivity growth into technological change, efficiency change, and capital accumulation.

Proxies for technological growth also have been widely used in the literature. They include patent data, R&D expenditure, and major innovation counts, all of which have various strengths and weaknesses. Alexopoulos and Cohen (2011) generate a new proxy variable and provide a survey of many used in the past.

To my knowledge, no studies exist that have used a cross-country panel to investigate the relationship between age structure and technological progress. The only study looking for this relationship is that of Nishimura, Minetak, and Shirai (2002). They take the Japanese economy in isolation and find that the share of workers over the age of 40 had a positive effect on technological progress in the 1980s. However, the authors measure technological progress using a modified Solow residual and assume efficiency. For an advanced economy such as Japan, efficiency may be an appealing assumption. Unfortunately, Fare, Grosskopf, Norris, and Zhang (1994) present evidence that half of the increase in Japanese productivity between 1979 and 1988 is actually attributable to efficiency improvements.

3 Data

Technological growth measures for OECD countries are widely available. They include both computed measures and proxy variables. However, for my application I must expand my panel beyond OECD countries in order to obtain a sufficient number of observations to permit empirical analysis. Technological measures for non-OECD countries are much more scarce; Nevertheless, both proxies and computed measures can still be found.

Of those available, no "best practice" measure exists. For example, there is not even a consensus as to how computed measures should be generated. Different underlying assumptions or constructions of the world technology frontier will produce different results. For example, we may wish to (data permitting) differentiate between skilled and unskilled labour, or even disaggregate US data by state in order to achieve a more precise estimate of the world technology frontier. However, increasing precision tends to come at the expense of observations since it requires more data sources. I therefore opt to collect two measures of technological progress: one computed and one proxy. My computed measure is drawn from Kumar and Russell (2002, hereafter referred to as KR), while I choose patent application counts as my proxy. Table 1 presents summary statistics for both.

Table 1: Summary Statistics of Dependent Variables

	KR Growth	Patent Application Counts
Countries	54	88
Obsv. Per Country	5	6.45
Time Span	1965–1985	1960–2010

In addition to the technological progress measures, I collect population age distribution data and several additional explanatory variables. These variables and their sources will be describe in detail in section 3.3

3.1 Technological Progress: Computed Measure

KR develop the best computed variable for my application. They measure technological progress directly and with minimal underlying assumptions by decomposing productivity growth into three components: efficiency, capital accumulation, and technological progress. Just as importantly, they generate the largest number of observations of any computed measure I was able to find. Their measure is, however, sensitive to the precision with which they calculate the worldwide production frontier.

To generate their computed measure, KR begin by constructing the worldwide production frontier using a DEA algorithm. Since this approach is non-parametric, it requires no underlying assumptions on the production function. It assumes only constant returns to scale. The input variables are aggregate labour and aggregate capital, while the output variable is aggregate output. From this, the Farrell efficiency index (see Farrell, 1957) is calculated for each country in each year. It equals 1 if an observation lies on the frontier, and is less than 1 otherwise. KR then perform the following decomposition

$$\frac{y_c}{y_b} = \frac{e_c \times y_c(k_c)}{e_b \times y_b(k_b)} \tag{1}$$

where y_b is aggregate output per person in the base period, y_c is aggregate output per person in the current-period, e_b and e_c are the Farrell efficiency index values for the two periods, and k_b and k_c are capital per unit of labour in the two periods. $y_b(k_b)$ is potential output in the base-period. It is obtained by dividing aggregate output in the base-period by the Farrell efficiency index in that period. If a country is efficient, potential output will equal actual output. For an inefficient country, potential output will be greater than actual output.

Multiplying the numerator and denominator of (1) by the potential output-labour ratio at current-period capital intensity using base-period technology will complete the decomposition. However, the decomposition could also be completed by multiplying through by the potential output-labour ratio at base-period capital intensity using current-period technology. Because the choice between these alternatives is completely arbitrary and will make a difference unless technology is Hicks neutral, the authors opt to use the geometric average of the two alternatives. Thus they multiply the numerator and denominator of (1) by $(y_b(k_c)y_c(k_b))^{1/2}$, yielding:

$$\frac{y_c}{y_b} = \frac{e_c}{e_b} \times \left(\frac{\bar{y}_c(k_c)}{\bar{y}_b(k_c)} \times \frac{\bar{y}_c(k_b)}{\bar{y}_b(k_b)}\right)^{1/2} \times \left(\frac{\bar{y}_b(k_c)}{\bar{y}_b(k_b)} \times \frac{\bar{y}_c(k_c)}{\bar{y}_b(k_b)}\right)^{1/2} \tag{2}$$

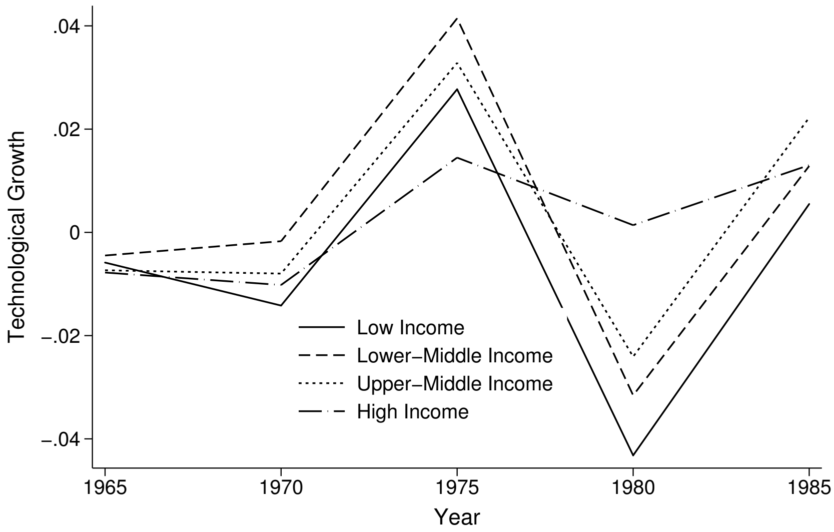
$$= \text{Efficiency} \times \text{Technological Progress} \times \text{Capital Accumulation}$$

This gives technological growth for five year intervals from 1965 through 1990. For my application, I annualise these five year intervals so that observed technological growth in period t is the geometric annualisation of total technological growth for the period t to $t + 5$. This leaves 270 observations, reported every lustrum from 1965 to 1985, for 54 countries. For the remainder of the essay, this variable will be referred to as KR Growth.

The sample includes the following countries: Argentina, Australia, Austria, Belgium, Bolivia, Canada, Chile, Colombia, Denmark, Dominican Republic, Ecuador, Finland, France, Germany, Greece, Guatemala, Honduras, Hong Kong, Iceland, India, Ireland, Israel, Italy, Jamaica, Japan, Kenya, Republic of Korea, Luxembourg, Madagascar, Malawi, Mauritius, Mexico, Morocco, Netherlands, New Zealand, Nigeria, Norway, Panama, Paraguay, Peru, Philippines, Portugal, Sierra Leone, Spain, Sri Lanka, Sweden, Switzerland, Syrian Arab Republic, Thailand, Turkey, United Kingdom, USA, Zambia, and Zimbabwe.

Figure 1 plots KR Growth across time for countries in my sample. For illustrative purposes, I grouped countries according to the World Bank income classification. KR Growth in low income, lower-middle income, and upper-middle income countries appears to move together. Additionally, KR Growth varies greatly for those countries. According to this measure, from 1970-1975, these countries experienced large technological growth before seeing substantial technological regress in the following lustrum. Conversely, technological growth in high income countries appears to remain relatively stable over time.

Figure 1: KR Growth by Income Group



3.2 Technological Progress: Patent Application Counts

I collected patent application counts from the World Bank as a proxy for technological progress. Two characteristics make patent application counts particularly suitable to my application. First, a long time series is available for many countries. Second, unlike major innovation counts, for example, patent application counts are an objective measure.

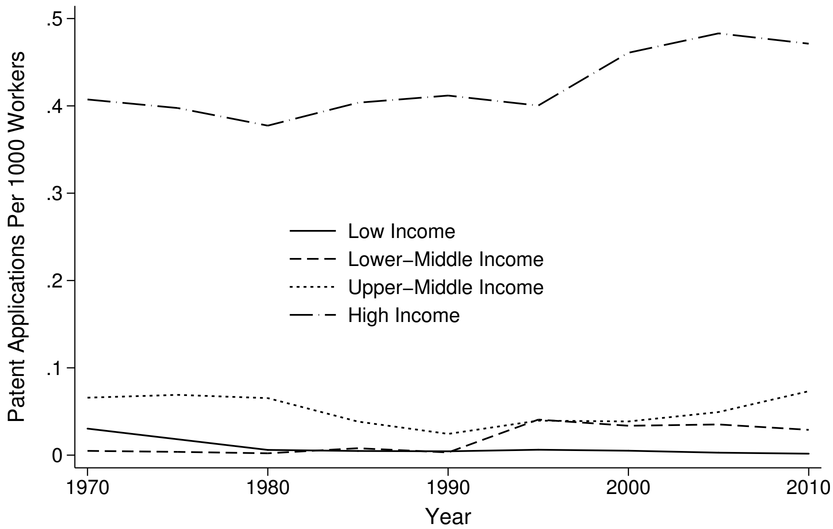
As with any proxy, patent application counts do have their drawbacks. For one, the establishment of the European Patent Office and World Intellectual Property Organization in the 1970s may have changed patent application behaviour of many european nations (Geert Boedt, personal communication, May 11, 2012). Additionally, a patent application does not necessarily imply that a valid innovation was made. The application may be denied, and even if accepted, the innovation may never make it to the market (Alexopoulos and Cohen, 2011).

I expect the number of patent application to increase with the number of workers in a population. All else equal, more people innovating should lead to more innovation. For this reason, I normalize the number of patent applications by the number of workers in each country for each year. This gives the number of patent applications per one thousand workers.

Observations are reported every 5-years with gaps from 1960 through 2010. The sample includes the following countries: Algeria, Argentina, Armenia, Australia, Austria, Bangladesh, Belarus, Belgium, Brazil, Bulgaria, Canada, Chile, China, Hong Kong, Colombia, Costa Rica, Croatia, Cuba, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Guatemala, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Latvia, Lithuania, Luxembourg, Madagascar, Malaysia, Malta, Mexico, Mongolia, Morocco, Netherlands, New Zealand, Nigeria, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Republic of Korea, Republic of Moldova, Romania, Russian Federation, Saudi Arabia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Syrian Arab Republic, Tajikistan, Thailand, Tunisia, Turkey, Ukraine, United Kingdom, USA, Uruguay, Uzbekistan, Venezuela, Viet Nam, and Zimbabwe.

Figure 2 plots patent applications per one thousand workers (hereafter referred to simply as "Patent Applications") against time for countries in my sample. Again, I grouped countries according to the World Bank income classification. High income countries file many more patents and experience a slight upward trend in patent applications not experienced by the other income groups. It is worth noting that figure 2 and figure 3 look very different despite the fact that both variables are being used to measure the same thing. This is in part caused by the fact that KR Growth allows for technological regress while Patent Applications does not.

Figure 2: Patent Applications by Income Group



3.3 Age Distribution and Controls

I collected age structure data from the United Nations (UN). The UN reports observations every 5 years for population age cohorts 5 years in length. The data is in count form meaning that for a given year and a given country, I observe the number of individuals aged 15–19, the number of individuals aged 20–24, the number of individuals aged 25–29, *etc.*. I convert this to proportions so that I observe the proportion of each population in each age cohort.

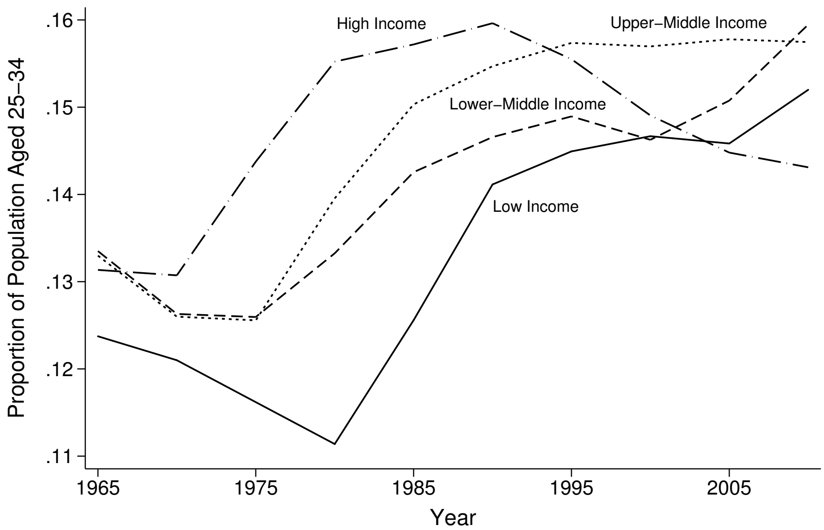
Life expectancy, total population, and real GDP per worker are also included in select specifications. Life expectancy is available from the World Bank and is defined as the number of years a newborn infant is expected to live, holding current mortality patterns constant. Real GDP per worker (hereafter referred to as RGDP) is available from PENN World Tables and total population is available from the UN.

4 Statistical Method and Results

4.1 Identification

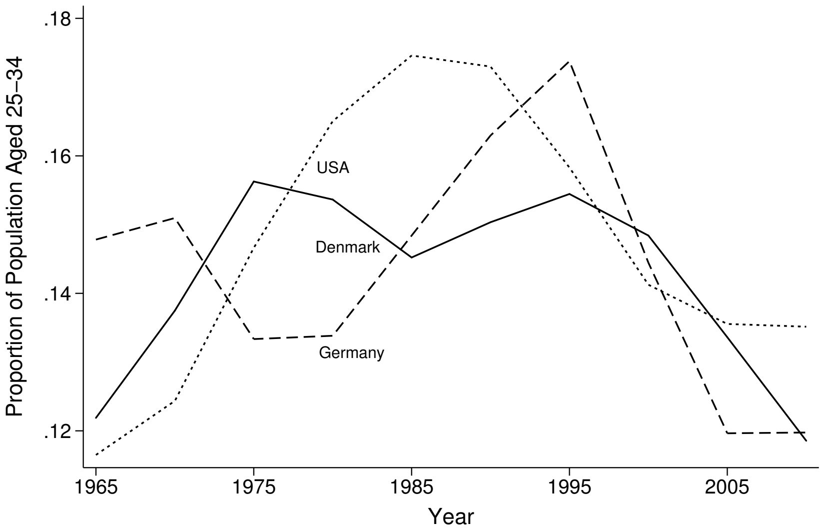
My sample contains substantial time series variation due to the postwar baby boom. Figure 3 groups countries based on income and plots the proportion of the population aged 25–34 for each. In low income countries the proportion rises from 11% in 1980 to 15% in 2010 (a 36% increase). Meanwhile, high income countries experience a 23% increase between 1970 and 1990.

Figure 3: Time Series Variation by Income Group



The variation is even more pronounced when broken down by country. Figure 4 graphs the proportions of population aged 25–34 across time in Denmark, Germany, and the USA. For the USA, the proportion rises from approximately 12% in 1965 to 18% in 1985 (a 50% increase), before falling to 14% in 2010 (a 22% decrease).

Figure 4: Time Series Variation by Country



This variation allows me to estimate the following fixed effects model:

$$h_{it} = \beta x_{it} + \phi_t + \phi_i + u_{it} \tag{3}$$

where i is the individual country and t is time. h_{it} is the technological growth variable. Depending on the specification, it will be either KR Growth or Patent Applications. x_{it} is a vector of explanatory variables that includes a set of age structure variables along with RGDP, total population, and life expectancy.[1] RGDP proxies for development factors that are not constant over time, while total population controls for population changes. ϕ_i is the set of country fixed effects that control for unobserved heterogeneity across countries that is constant over time; ϕ_t is a set of time fixed effects controlling for time variant shocks common to all countries; u_{it} is the error term. Although serial correlation of the error term should be

[1]Infant mortality rates and dependency ratios were also tried in various specifications. However, none was statistically significant at conventional levels.

minimized since observations are in 5 year intervals, I still cluster standard errors by country in all specifications.

4.2 Estimation

I estimate equation (3) with age structure data implemented as proportional cohorts. Since the proportions sum to unity, at least one cohort must be dropped to prevent singularity in $(X'X)$. I choose to drop age cohorts below 15 and above 64 years of age. The coefficient estimate for any included cohort is then interpreted as the effect of a shift in age share out of the omitted cohorts and into the included cohort.

Table 2 presents estimation results with KR Growth as the regressand. Column 1 is estimated with OLS. Only cohort 55–64 is statistically significant (with a positive coefficient) at the 5% level. The inclusion of country fixed effects (column 2) greatly affects estimates. Cohort 55–64 remains significant, and cohort 34–44 is now also significant (with a positive coefficient) at the 5% level.

Columns 3 adds in year fixed effects (not reported). They are jointly significant at the 1% level. However, none of the cohort variables in this specification is significant at conventional levels. Life expectancy is significant in column 2, but not column 3. This is likely because it is picked up in the year fixed effects. Strangely, though total population has a positive effect as we might expect in column 1, its coefficient is negative and significant in columns 2 and 3.

Table 2: KR Growth Estimation Results

	(1) KR Growth	(2) KR Growth	(3) KR Growth
total population	0.000**	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
life expect.	0.001**	0.002***	0.000
	(0.000)	(0.000)	(0.001)
cohort 15–24	-0.033	-0.005	-0.009
	(0.060)	(0.067)	(0.064)
cohort 25–34	0.071	0.085	0.057
	(0.073)	(0.075)	(0.080)
cohort 35–44	-0.013	0.213**	0.129
	(0.075)	(0.092)	(0.118)
cohort 45–54	0.033	0.183	0.016
	(0.105)	(0.121)	(0.100)
cohort 55–64	-0.188**	-0.322**	-0.067
	(0.080)	(0.127)	(0.109)
constant	-0.027**	-0.121***	-0.002
	(0.011)	(0.026)	(0.054)
country f.e..	No	Yes	Yes
year f.e..	No	No	Yes
obsv.	269	269	269

Note: Standard errors, reported in parentheses, are clustered at the country level. ***, **, * Statistically distinct from zero at the 1%, 5%, and 10% level respectively. Dependent variable is the geometric average annualized growth rate of technology for the period t to $t+5$

Table 3 reports estimation results using the natural log of Patent Applications as the regressand. Total population does not factor into these specifications because the dependent variable has already been normalized by the working-age population. Column 1 is estimated using OLS. The 55–64 age group is statistically significant (with a positive coefficient) at conventional levels and appears to have the largest impact on patent applications. Once country fixed effects are added, the magnitude decreases greatly and statistical significance is lost.

Table 3: Patent Applications Estimation Results

	(1) Log(Patent Apps)	(2) Log(Patent Apps)	(3) Log(Patent Apps)
life expect.	0.008	-0.021	0.058**
	(0.030)	(0.019)	(0.027)
RGDP	0.000***	-0.000	0.000*
	(0.000)	(0.000)	(0.000)
cohort 15–24	11.519**	3.032	3.476
	(5.209)	(3.509)	(3.063)
cohort 25–34	-8.746*	-0.371	5.009*
	(4.503)	(2.568)	(2.866)
cohort 35–44	3.187	8.934**	13.063***
	(5.335)	(4.008)	(4.251)
cohort 45–54	11.094**	14.762***	21.770***
	(4.559)	(4.353)	(5.633)
cohort 55–64	35.033***	5.431	9.606
	(5.450)	(7.907)	(6.799)
constant	-9.727***	-5.158***	-13.884***
	(1.823)	(1.586)	(2.386)
country f.e.	No	Yes	Yes
year f.e.	No	No	Yes
obsv.	568	568	568

Note: Standard errors, reported in parentheses, are clustered at the country level. ***, **, * Statistically distinct from zero at the 1%, 5%, and 10% level respectively. Dependent variable is the natural log of the number of patent application per 1000 workers. RGDP is real gross domestic product per worker

Column 3 adds year fixed effects (jointly significant at the 1% level). Cohort 45–54 and cohort 35–44 have the largest impacts and are statistically significant at the 1% level, suggesting that a shift in population proportion from age groups younger than 14 and older than 65 to the 35–54 age group causes the largest increase in technological progress.

Upon closer inspection, column 3 appears to illustrate a life cycle hypothesis. As a greater proportion of the population is shifted into older cohorts, there is an increasing positive effect on technological progress until a peak is reached at the 45–54 cohort. The positive effect is lessened in the 55–64 cohort and is presumably smaller or even negative in more elderly cohorts. A similar shape is seen in table 2 column 3.

4.3 Multicollinearity

The results in section 4.2 may be imprecise because of multicollinearity. The specifications include age structure data as proportional cohorts. Even after dropping three cohorts to avoid perfect collinearity, multicollinearity is likely still rife since the cohorts move together by definition. Severe levels of multicollinearity cause difficulty in measuring the individual contribution of each independent regressor, making it difficult to trust the estimates.

A couple of additional issues arise with the use of proportional cohorts. Firstly, there is no rule for how cohort lines should be drawn. For example, should 35-year olds be grouped in a cohort with 40-year olds or in a cohort with 30-year olds? This arbitrariness cannot be avoided and the problem is further compounded when behaviour between similar age groups varies significantly (Macunovich, 2009). Additionally, interpreting estimates as the effect of a shift in age structure out of some combination of excluded cohorts and into included cohorts only gives a vague picture of age structure effects.

For these reasons, I use Almon's (1965) distributed lag technique to accomadate age structure data in my model. This method was first applied to age structure data by Fair and Dominguez (1991) and is implemented beginning with the following specification:

$$h_{it} = \beta x_{it} + \gamma_1 p_{1it} + \gamma_2 p_{2it} + ... + \gamma_J p_{Jit} + \phi_t + \phi_i + u_{it} \tag{4}$$

J is the number of population cohorts and p_{1it} is the proportion of the population in the first age cohort in country i. Since the UN reports age structure in 5 year cohorts, there are 17 population cohorts in my sample.

I constrain the age structure coefficients, γ_j, to lie along a third-order polynomial so that:

$$\gamma_j = \eta_0 + j\eta_1 + j^2\eta_2 + j^3\eta_3$$

This allows equation (4) to be re-written as

$$h_{it} = \beta x_{it} + \eta_0 \sum_{j=1}^{17} p_{jit} + \eta_1 \sum_{j=1}^{17} j p_{jit} + \eta_2 \sum_{j=1}^{17} j^2 p_{jit} + \eta_3 \sum_{j=1}^{17} j^3 p_{jit} + \phi_t + \phi_i + u_{it} \quad (5)$$

Next, because p_{jit} sums to 1 across j, I impose $\sum_{j=1}^{17} \gamma_j = 0$ to allow estimation. This restriction yields:

$$\eta_0 = -(\eta_1/17) \sum_{j=1}^{17} j - (\eta_2/17) \sum_{j=1}^{17} j^2 - (\eta_3/17) \sum_{j=1}^{17} j^3 \quad (6)$$

which gives

$$h_{it} = \beta x_{it} + \eta_1 c_{1it} + \eta_2 c_{2it} + \eta_3 c_{3it} + \phi_t + \phi_i + u_t \quad (7)$$

where

$$c_{1it} = \sum_{j=1}^{17} j p_{jit} - (1/17) \sum_{j=1}^{17} j \quad (8)$$

$$c_{2it} = \sum_{j=1}^{17} j^2 p_{jit} - (1/17) \sum_{j=1}^{17} j^2 \quad (9)$$

$$c_{3it} = \sum_{j=1}^{17} j^3 p_{jit} - (1/17) \sum_{j=1}^{17} j^3 \quad (10)$$

Estimation now involves only three parameters, η_1, η_2, and η_3, rather than seventeen. From the estimates of η_1, η_2, and η_3, I can back out η_0 using equation (6). I can then recover the implied γ_j coefficients.

Table 4: KR Growth Estimation Results

	(1) KR Growth	(2) KR Growth	(3) KR Growth
total population	0.000	0.000^{*}	-0.000^{***}
	(0.000)	(0.000)	(0.000)
life expect.	0.000	0.000	0.001^{*}
	(0.000)	(0.000)	(0.001)
c_1	0.167^{***}	0.090^{**}	0.167^{**}
	(0.059)	(0.040)	(0.069)
c_2	-0.027^{***}	-0.015^{**}	-0.029^{**}
	(0.009)	(0.006)	(0.011)
c_3	0.001^{***}	0.001^{**}	0.001^{***}
	(0.000)	(0.000)	(0.001)
constant	0.009	0.012	0.018
	(0.022)	(0.020)	(0.035)
country f.e.	No	Yes	Yes
year f.e.	No	No	Yes
obsv.	269	269	269

Note: Standard errors, reported in parentheses, are clustered at the country level. ***, **, * Statistically distinct from zero at the 1%, 5%, and 10% level respectively. Dependent variable is the geometric average annualized growth rate of technology for the period t to t+5

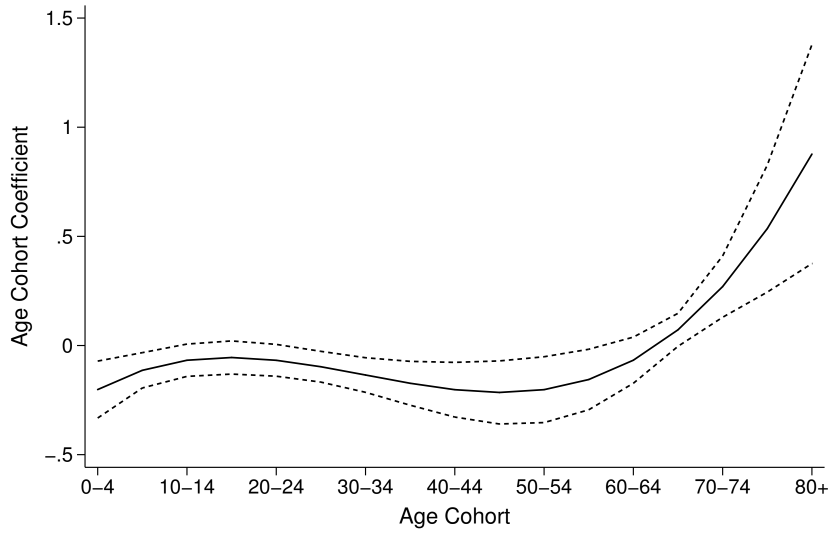
Table 4 presents results implementing Almon's method using KR Growth as the regressand. c_1, c_2, and c_3 are jointly significant at the 1% level in columns 1 and 3 and are significant at the 10% level in column 2. c_1 and c_3 are always estimated with positive signs while c_2 has a negative sign in all columns. [2] Again, year fixed effects (not reported) in column 3 are jointly significant at the 1% level.

Using estimates from column 3, I back out the implied $\hat{\gamma}_j$ coefficients along with the associated 95% confidence bands and plot them in figure 5. We see a small hump towards the beginning of the age groups, peaking at 15-19. However,

[2] I initially began estimating with a ninth-order polynomial and worked my way down, testing for joint significance. A second-order polynomial is not significant at conventional levels while a fourth-order polynomial is only marginally significant and not robust to changes in specification.

the coefficient estimate on this age group is not statistically significant. Strangely, the age cohort coefficients begin to rise again at the 50–54 age cohort.

Figure 5: KR Growth Estimation:
Implied Age Structure Coefficients



Note: The figure shows the implied $\hat{\gamma}_j$ coefficients and the 95% confidence interval generated from the results in table 4 column 3.

Table 5 reports estimates implementing Almon's method using the natural log of Patent Applications as the regressand. Column 1 is estimated with OLS. Adding country fixed effects in column 2 causes the magnitude of the estimates to fall. c_1, c_2, and c_3 are jointly significant at the 1% level in columns 1 and 2, and jointly significant at the 5% level in column 3.

Figure 6 plots the implied $\hat{\gamma}_j$ coefficients and 95% confidence bands from column 3. The graph shows a clear life cycle story and reaffirms the estimates in table 3 column 3. The peak age group for innovation is the 50–54 group, and its coefficient is significant at the 1% level. In fact, all $\hat{\gamma}_j$ coefficients estimated for cohorts falling between 35 years and 65 years are statistically significant at the 1% level.

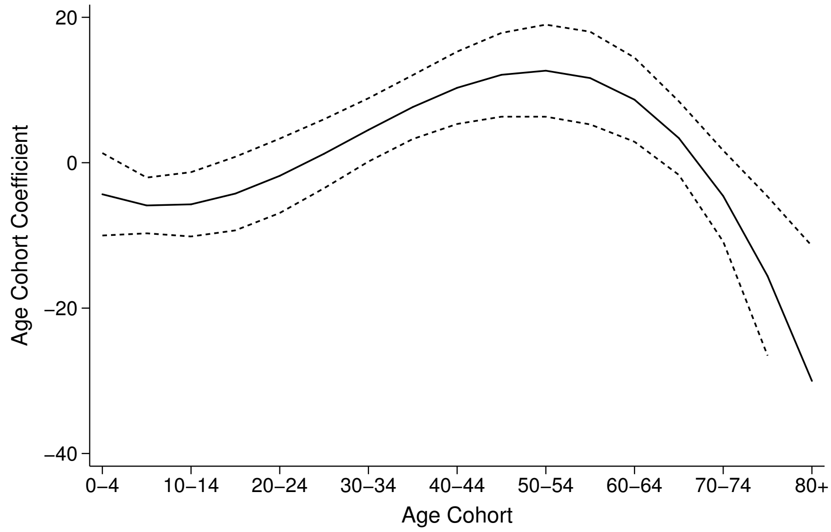
The shape of the life cycle pattern is also of interest. We see only a gradual rise in innovation with age up to the 50–54 age group. That is, there is a 25 year age gap between the first positive coefficient at the 25–30 age cohort and the innovation peak at the 50–54 cohort. However, once this peak is reached, innovation drops off drastically; within the next 15 years, the age shares become negatively associated with technological growth. These results are robust to the use of a time trend rather than year dummies and the inclusion of other (though not statistically significant) explanatory variables.

Table 5: Patent Applications Estimation Results

	(1) Log(Patent Apps)	(2) Log(Patent Apps)	(3) Log(Patent Apps)
RGDP	0.000***	0.000	0.000**
	(0.000)	(0.000)	(0.000)
c_1	-10.876***	-9.500***	-4.740
	(3.421)	(3.185)	(3.504)
c_2	1.827***	1.798***	1.209**
	(0.551)	(0.502)	(0.542)
c_3	-0.076***	-0.083***	-0.061***
	(0.025)	(0.021)	(0.022)
constant	-3.197***	-5.430***	-3.751***
	(1.170)	(1.008)	(0.654)
country f.e.	No	Yes	Yes
year f.e.	No	No	Yes
obsv.	568	568	568

Note: Standard errors, reported in parentheses, are clustered at the country level. ***, **, * Statistically distinct from zero at the 1%, 5%, and 10% level respectively. Dependent variable is the natural log of the number of patent application per 1000 workers. RGDP is real gross domestic product per worker.

Figure 6: Patent Applications:
Implied Age Structure Coefficients



Note: The figure shows the implied $\hat{\gamma}_j$ coefficients and the 95% confidence interval generated from the results in table 5 column 3

4.4 Instrumental Variables

Since populations are mobile, not only might age structure affect technological growth, but it is likely that technological growth can impact age structure. For example, a country with high technological growth may see an influx of immigration while a country with low technological growth experiences a high rate of emigration. Additionally, technological growth may decrease infant mortality rates and increase life expectancy, both of which directly affect age structure. If present, this simultaneity will cause correlation between the age structure regressors and error term, leading to biased and inconsistent estimates. Since I cannot possibly control for all paths of this feedback, I choose to instrument for age structure.

For an instrument z_{it} to be valid, I require that (a) $Cov(z_{it}, c_{pit}) \neq 0, \ p = 1, 2, 3$, and (b) $Cov(z_{it}, u_{it}) = 0$. That is, an instrument must be correlated with my age structure variables and uncorrelated with omitted determinants of technological growth (Angrist and Pischke, 2009). I therefore instrument as follows: I take cohort proportions at year *t-5* and shift them forward by five years. For example, if 5% of the population is in cohort 25–30 at *t-5*, I shift this proportion so that 5% of the population is in cohort 30–35 at time *t*. I then calculate each of z_{1it}, z_{2it}, and z_{3it} in identical fashion to c_{1it}, c_{2it}, and c_{3it} respectively. [3].

My instruments and the c_{pit} variables are highly correlated since age structure does not tend to follow a random walk. Instead, it trends in one direction or another for many years as generations make their way through successive age cohorts. My instruments also satisfy the second requirement: uncorrelated with the error term. The correlation between u_{it} and c_{pit} is generated by the simultaneity operating through mortality, immigration, and emigration rates. By shifting the cohorts as described above, I am essentially generating the age distribution that would result in the absence of immigration, emigration, and mortality; thus eliminating the source of simultaneity.

Table 6 presents instrumental variable estimation results. Both specifications are estimated using two stage least squares. Column 1 estimates with KR Growth as the dependent variable. c_1, c_2, and c_3 are jointly significant at the 5% level. The graph of implied coefficients (figure 7) closely resembles the implied coefficients in figure 5. The Cragg-Donald F-Statistic for strength of instruments has a value of 165.91, allowing me to reject the null of weak instruments at the 1% level.

Column 2 present results using the natural log of Patent Applications as the dependent variable. c_1, c_2, and c_3 are jointly significant at the 5% level and the graph of implied coefficients (figure 8) closely resembles figure 6. The magnitude of the coefficients, however, is reduced by approximately 20% from figure 6. Nevertheless, the peak is again at the 50–54 age cohort and the shape is consistent.

[3] The z_{pit} instruments are calculated using only 16 age cohorts since the 0–5 cohort is lost during of the shift

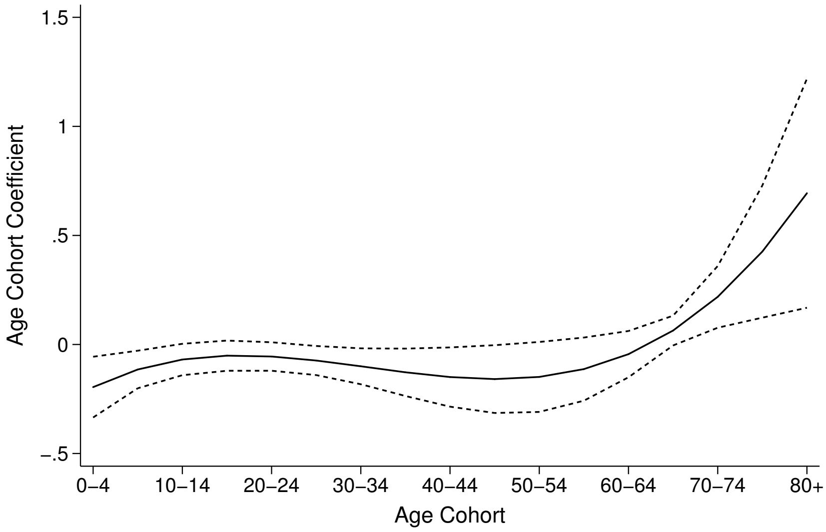
We see a gradual run-up to the peak, followed by a sharp decline. Once again I reject the null of weak instruments at the 1% level with a Cragg-Donald F-Statistic equal to 289.43.

Table 6: IV Estimation Results

	(1) KR Growth	(2) Log(Patent Apps)
total population	-0.000***	
	(0.000)	
life expect.	0.001	0.049**
	(0.001)	(0.023)
RGDP		0.000**
		(0.000)
c_1	0.146**	-3.636
	(0.071)	(2.997)
c_2	-0.024**	0.975**
	(0.012)	(0.462)
c_3	0.001**	-0.049**
	(0.001)	(0.019)
constant	0.024	-6.835***
	(0.035)	(1.732)
country f.e.	Yes	Yes
year f.e.	Yes	Yes
obsv.	269	568

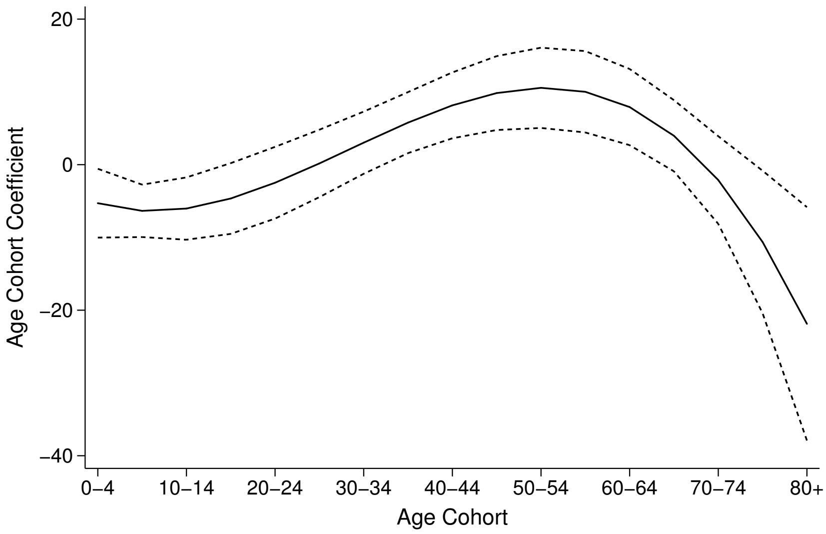
Note: Standard errors, reported in parentheses, are clustered at the country level. ***, **, * Statistically distinct from zero at the 1%, 5%, and 10% level respectively. Column 1 instruments are z_{1it}, z_{2it}, z_{3it}, total population, and life expectancy. Column 2 instruments are z_{1it}, z_{2it}, z_{3it}, RGDP, and life expectancy.

**Figure 7: KR Growth:
IV Implied Age Structure Coefficients**



Note: The figure shows the implied $\hat{\gamma}_j$ coefficients and the 95% confidence interval generated from the results in table 6 column 1.

**Figure 8: Patent Applications:
IV Implied Age Structure Coefficients**



Note: The figure shows the implied $\hat{\gamma}_j$ coefficients and the 95% confidence interval generated from the results in table 6 column 2.

5 Discussion

To summarize, my different regressands provide conflicting results. When Patent Applications is regressed on age structure, a clear life cycle hypothesis emerges. Innovation gradually rises until it reaches a peak at the 50–54 age group, then rapidly declines after that. This result remains even after addressing endogeneity concerns by instrumenting for age structure. Conversely, when KR Growth is regressed on age structure, a clear life cycle hypothesis does not emerge. This result (or lack thereof) is also robust to instrumenting and changes in specification.

The question arises of which result we should trust. Both dependent variables are imperfect measures so there is no clear answer. The KR Growth measure is sensitive to the constant returns to scale assumption; variable returns could have been assumed in its place. Additionally, the precision of the estimated world technology frontier used to measure technological growth could be increased by differentiating between skilled and unskilled labour. A more precise construction may lead to different results (Growiec, 2012).

The Patent Applications measure is also subject to valid criticisms. Patent filing behaviour can be affected by government policy changes. This means, it is possible for my results to be driven by changes in government R&D policy that are correlated with movements in age structure. This problem is likely mitigated by the size and heterogeneity of the cross sectional dimension of my panel, though it is still a concern. An additional criticism is that a patent application does not necessarily lead to a patent (a valid innovation). And even if a patent is granted, there is still no certainty that the innovation will ever make it to market or be implemented. Therefore, I am implicitly assuming that a constant fraction of patent applications each year for each country is granted into patents and implemented. The merits of this assumption are open to debate.

Despite these concerns, the Patent Applications results are particularly appealing, especially because they are consistent with results of the related studies of Feyrer (2007) and Lindh and Malmberg (1999). Feyrer finds that an increase in the share of workers aged 40–60 has a positive impact on productivity. Similarly, Lindh and Malmberg find that economic growth increases with the proportion of the population between the ages of 50 and 64. My results suggest that one of the driving forces behind the higher productivity and economic growth for these groups is an associated increase in innovation.

I have shown my results to be statistically significant. The obvious question then becomes: are they economically relevant? That is, does age structure have a large effect on technological progress? To answer this question, I follow Higgins (1998) and calculate the change in Patent Applications attributable to changes in age distribution for periods in my sample. To do this, I begin by calculating the age structure effect as

$$A_{it} = \sum_{j=1}^{17} |(p_{jit} - \bar{p}_{ji})|\gamma_j \tag{11}$$

where γ_j is the implied coefficient for cohort j from figure 8, p_{jit} is the population share in cohort j for country i at time t, and $\bar{p}_{ji}$ is the mean population share in cohort j for country i over the entire sample period. Thus, A_{it} is the age structure effect for country i in period t, and represents the deviation from country i's mean Patent Applications that is attributable to age structure.[4] On its own, the age structure effect is not terribly informative; However, taking the exponential of the difference in age structure effects for any two years gives the percentage change in Patent Applications induced by age structure. This is the age structure swing. Since I estimated in logs and am converting to levels, the formula for the age

[4]The age structure effect is measured relative only to the individual country means. This is done because the γ_j coefficients are estimated with a fixed effects model. Another valid approach would be to estimate without fixed effects then measure the age structure effect relative to the sample means.

structure swing is as follows:

$$S_{i,T} = e^{(A_{i,t_2}+\hat{\sigma}_2^2/2)-(A_{i,t_1}+\hat{\sigma}_1^2/2)} \tag{12}$$

where $S_{i,T}$ is the age structure swing for country i over period T, t_1 is the first year of period T, and t_2 is the last year of period T. $\hat{\sigma}_2^2$ and $\hat{\sigma}_1^2$ are the estimated residual variances of A_{i,t_2} and A_{i,t_1} respectively. $\hat{\sigma}_2^2$ and $\hat{\sigma}_1^2$ are likely very close in magnitude. For this reason and because of difficulty calculating their values, I instead calculate the age structure swing with the following approximation:

$$S_{i,T} = e^{A_{i,t_2}-A_{i,t_1}} \tag{13}$$

The Age structure swings are quite large. Table 7 presents the swings for Canada, USA, and UK for 1975–2000. For Canada, 13.9% of the increase in Patent Applications over that period is induced by changes in age structure. From 1975–2000 the young portion of the population fell by 32% while the middle and old portions rose by 8.6% and 36% respectively. For shorter time periods with relatively large changes in age distribution, the age structure swings for many countries (not reported) suggest that up to 50% of the change in Patent Applications is induced by age structure. These large age structure swings are consistent with the findings of Higgins for savings rates. For example, in Canada for the period 1960–1990, Higgins finds that age structure induced an increase in savings equal to 4.2% of GDP.

Table 7: Age Structure Swings for 1975–2000

	Age Structure Swing	Young	Middle	Old
Canada	13.9	-32.3	8.6	36.3
USA	10.7	-22.7	7.1	9.5
U.K.	13.8	-19.2	19.3	0.9

Note: Young: ages 0–19; Middle: ages 20–65; Old: ages 65+. Table reports percentage changes between 1975 and 2000.

6 Conclusion

Though there are several candidate hypotheses, this essay is silent as to exactly what is driving the Patent Applications results. One possibility is an experience-based human-capital explanation suggested by Lindh and Malmberg (1999). It may be that it takes until middle-age for individuals to accrue the experience necessary to effectively innovate. Another possibility is a savings behaviour explanation. Aggregate savings are high when a large portion of the population is middle-aged, meaning that more money is available for investment in innovation. However, given the openness of economies, and the fact that most innovation is generated by large enterprises, I am not sure how much traction this explanation has.

A final possibility is that when a large portion of the population is middle-aged, a relatively smaller portion of the population falls in dependent age groups. Funds that would usually be used to care for infants and the elderly can instead be directed to other activities such as R&D. However, I estimate several specifications experimenting with dependency ratios (not reported), and the results indicate that this is not the case. It is likely that a combination of the above explanations — a joint optimization of human capital and savings — drives the peak of innovation to the 50–54 age group.

The shape of the life cycle results has important implications. It suggests that as the baby boom population moved through age cohorts, economies experienced an increasing rate of technological progress (as opposed to the counterfactual in which there is no baby boom population). Throughout the last two decades, as the baby boom was near the peak innovation age, there was an innovation dividend. Unfortunately, this generation has now begun to retire. The steep decline in technological progress associated with cohorts greater that 50–54 years, and the size of the population now entering these cohorts, suggests we may soon see a decrease in technological growth. This will make the burden of the rising dependency ratio that much more difficult to bear for younger generations.

Finally, further research is required to reconcile the results of the computed growth measure with the results of the patent applications proxy. A first step may be to experiment with different constructions of the world technology frontier. Growiec's twelve different computed measures based on changes in assumptions and information sets is clear evidence that experimentation is warranted.

7 Bibliography

Alexopoulos, M., Cohen, J. (2011) Volumes of evidence: Examining technical change in the last century through a new lens. *Canadian Journal of Economics*, 44, 413–350.

Almon, S. (1965). Distributed lag between capital appropriations and expenditures. *Econometrica*. 30, 178–196

Angrist, J., Pischke, J.S., (2008). *Mostly harmless econometrics: An empiricist's companion*. New Jersey: Princeton University Press.

Basu, S., Fernald, J.G. (2002). Aggregate productivity and aggregate technology. *European Economic Review*. 46, 963–991.

Fair, R.C., Dominguez, K.M. (1991). Effects of the changing U.S. age distribution on macroeconomic equations. *The American Economic Review*, 81, 1276–1294.

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A, General*, 120, 253–282.

Fare, R., Grosskopf, S., Norris, M., Zhang, Z., (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review*, 84, 66–83.

Feyrer, J. (2007) Demographics and productivity. *The Review of Economics and Statistics*, 89, 100–109.

Growiec, J. (2012). On the measurement of technological progress across countries. Unpublished manuscript. mimeo, National Bank of Poland

Higgins, M. (1998). Demography, national savings, and international capital flows. *International Economics Review*. Vol. 39, No. 2, pp.343–369.

Kumar, S., Russell, R.R. (2002) Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *The American Economic Review*, 92, 527–548

Lindh, T., Malmberg, B. (1999) Age structure effects and growth in the OECD 1950-1990. *Journal of Population Economics*, 12, 431–449.

Lindh, T., Prskawetz, A. (Eds.). (2006) *The Impact of Population Ageing on*

Innovation and Productivity Growth in Europe: Research report 28. Vienna: Vienna Institute of Demography.

Macunovich, D.J. (2009). The role of demographics in precipitating crises in financial institutions. *IZA Discussion Paper Series.* Discussion paper No. 4436.

Nishimura, K.G., Minetaki, K., Shirai, M. (2002). Effects of information technology and aging work force on labour demand and technological progress in Japanese industries: 1980-1998. *University of Tokyo Discussion Paper F Series.* Retrieved from http://repository.dl.itc.u-tokyo.ac.jp

Poterba, J. (2012). Economic policy challenges of an aging population. *W.A. Mackintosh Visiting Lectureship in Economics.* Lecture conducted from Queen's University, Kingston, ON.

PWT 7.0. Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

Statistics Canada Demography Division. (2010). Population projections for Canada, provinces and territories 2009 to 2036. Retrieved from http://www.statscan.gc.ca

United States Census Bureau. (2010). Aging boomers will increase dependency ratio, census bureau projects. Retrieved from http://www.census.gov