Article:
Popli, G., Gladwell, D. and Tsuchiya, A. (2012) Estimating the critical and sensitive periods of investment in early childhood: A methodological note. HEDS Discussion Paper 12/04. (Unpublished)

HEDS Discussion Paper

No. 12.04

Estimating the critical and sensitive periods of investment in early childhood: A methodological note

Gurleen Popli[a], Daniel Gladwell[b] *, Aki Tsuchiya[a b]

1. Health Economics and Decision Science, School of Health and Related Research, University of Sheffield
2. Department of Economics, University of Sheffield

For HEDS DP; March 2012

Estimating the critical and sensitive periods of investment in early childhood: A methodological note

Gurleen Popli[a], Daniel Gladwell[b *], Aki Tsuchiya[a b]

a. Department of Economics, University of Sheffield, Sheffield, S1 4DT, UK
b. School of Health and Related Research, University of Sheffield, Sheffield, S1 4DA, UK

* Corresponding author <cmp11djg@sheffield.ac.uk>

Key words

Skill formation; child development; critical period; latent variable; structural equations

Abstract

This paper provides an overview of different quantitative methods available for the statistical analysis of longitudinal data regarding child development, and in particular the identification of critical and sensitive periods for later abilities. It draws heavily on the work on human skill formation developed by the economist James Heckman, which treats ability as a latent variable and explains its formation through the simultaneous estimation of structural equations of investments and achieved abilities across time. We distinguish between two specifications of the ability formation function. One of them (the 'recursive') format explains current ability as a function of the ability and investment *at the immediately preceding period*. The other (the 'non-recursive') format explains current ability as a function of *a series of past investments*. In order to fully examine critical and sensitive periods of investments, the non-recursive formulation needs to be used. Furthermore, true abilities of an individual cannot be directly observed: what we observe are the test scores, for example, on reading and writing. We outline an approach based on structural models that treats actual test scores as measurements of the latent ability variable, and show how it can be used in the recursive and non-recursive formulation. In order to fully examine critical and sensitive periods of investments, we argue that the non-recursive formulation of this structural model is necessary. However, the non-recursive formulation requires more data than the recursive formulation, and to the best of our knowledge, has never been used in the identification of critical and sensitive periods in early childhood development.

(254wds)

Estimating the critical and sensitive periods of investment in early childhood: A methodological note

1. Introduction

Independent strands of research in neuroscience, developmental psychology, economics, and beyond have come to the shared conclusion that there are certain periods in an individual's development which are particularly influential in determining future adult outcomes such as health or socioeconomic status. Early childhood is regarded as a crucial stage in the trajectory of human development because the inputs the young child receives can have much larger and longer lasting impacts (both positive and negative) compared to the same inputs received at later stages in the child's life. This has led to the concepts of critical and sensitive periods of investment for human development.

One main methodological challenge regarding the quantitative analysis of human development is that arguably both true ability and health are not directly observable. What we can see are specific behaviours and achievements, or symptoms, from which we infer underlying ability and health. Test scores and personality evaluations are taken as indicators of latent cognitive and non-cognitive skills respectively. So for example, a reading tests score of a child is understood to reflect the child's ability, but it is only one measurement. This leads to the need for a statistical approach that treats these observed data as draws from an underlying latent variable representing the true ability of the individual. The economist James Heckman and colleagues have developed one such approach (Cunha and Heckman, 2006 and 2008; Conti, Heckman and Urzua, 2010; Conti and Heckman, 2010). The generalised framework provided by Heckman and co-authors allows for the estimation of the link between the unobserved endowments, the environmental factors, and the observed outcomes. Their models that use this approach have demonstrated that, for example, investment occurring in the early stages of childhood has a greater effect than investment occurring in adolescence (Cunha, Heckman and Schennach, 2010).

The aim of this short note is to provide an introduction to Heckman's analytical method of modelling human skill formation, and to provide a rigorous definition of the concepts of critical and sensitive periods and their operationalisation. The rest of the paper is organised as follows: in section 2 we provide a general model of ability formation over time. Section 3 illustrates how an analysis of correlations can be carried out with the general model, which allows the identification of critical and sensitive periods of investments. In section 4 we provide an alternative formulation of the general model which allows going beyond the correlation analysis, and controlling for other factors, in the identification of critical and sensitive periods. Section 5, explains the identification of the critical and sensitive periods with a linear specification of the model of ability formation. In sections 6 and 7 illustrates how to accommodate latent specifications of the ability formation function. Section 8 concludes the paper.

2. General model of ability formation

Let the ability of the child at any given time period (*t*) be defined as: . An ability formation function is given as:

$$ \quad (1) $$

where can be a vector. It can be the cognitive ability of the child, the non-cognitive ability of the child, or both. In a more general framework, it can also be the health of the child. In this paper, we refer to specification (1) as '*recursive'*, since ability at any point in time depends on the stock of abilities at the time period directly preceding it. is the initial endowment that the child is born with. is investment made in the child at time *t-1* (mainly by parents, but this could include a wider range of inputs and stimuli; can be interpreted as any investment made in the child before it is born. The function assumes that investment has an effect on the ability with a lag. is a set of control variables that may affect ability formation such as gender of the child, socioeconomic status of the parents, household composition, social environment, and so on.

Embedded in the above framework are the concepts of self-productivity and dynamic complementarity. *Self-productivity* is said to exist when higher stocks of ability in period *t-1* is associated with higher stocks of ability in period *t*; thus it implies . *Dynamic complementarity* arises when the stock of skills acquired by period *t-1* is associated with more productive investment in period *t-1* towards ; this means .

Carneiro and Heckman (2003) provide evidence of dynamic complementarities: they show that 'economic returns to job training, high school graduation, and college attendance are lower for less able persons'. Existence of dynamic complementarities makes a case for investments to be made earlier rather than later. A stronger form of dynamic complementarity holes where early investments need to be followed by later investments for the early investments to be productive

(Heckman, 2008). Currie and Thomas (1995) provide an example of this, where their findings suggests that early investments made in children have weak effects in later years if these investments are not followed by later investments.

By the very nature of how the ability formation function is defined here, longitudinal data are necessary, with the minimum of two periods of data on a given child. However, often there are data limitations. Todd and Wolpin (2003) list a series of specifications of the ability formation function that can be estimated under different data available, and the underlying assumptions involved in each (see their Table 3).

In the discussion below, for simplicity of exposition, we will assume that there are three time periods in an individual's life (*t = 1, 2, 3*), consisting of early childhood, late childhood and adulthood. Quite often the researcher is interested in the impact of investments made in childhood periods ( ) on the adult skills/abilities, given by .

Investment in a critical period by definition cannot be delayed (more generally, neither can it be made before time) as the same investment will not be productive if given at any other period of child's life. An example often given in the literature is of corrective treatment for cataract in infants. If the treatment is not provided at a specific period then the child will lose his vision; the same treatment will not be effective (productive) if given any later. In this sense there is a critical period for the investment (treatment) to be made. Similarly, investment is defined to be 'sensitive' if it is more productive at a given point in time than at a later time.

As may be expected, while certain medical interventions are known to have critical periods outside of which the intervention has no effect on health, most educational investments are more likely to have sensitive periods rather than critical periods. However, it should also be noted that the distinction between a critical and a sensitive period is one that is relative to the duration of time periods. Narrower definitions of periods (in other words, more frequent observations) will lead to fewer periods being identified as critical and more as sensitive.

3. An analysis of correlations

The starting point of any analysis in assessing critical and sensitive periods would be to look at simple correlations between the ability at different time periods and the investments made in earlier periods.

Table 1: Cross tabulation of correlations between investments and abilities at different time periods

	-	-	-	-
		-	-	-
			-	-
				-

Here, represents the correlation between ability at period *t* (θ_t) and investment at period *s* (I_s), where *t* > *s*. Since we assume investments made in period *t* do not translate into ability in the same period *t*, is not defined.

If (where and for all , then we say that is a critical period for investment in ability at period *t*. So for example given the above correlation matrix, if we find , but , then we can say that period *1* is a critical period for investment in ability at period *2*. The restrictions are placed only along the columns, and it is plausible that is non-zero and significant, since θ_2 is correlated with θ_3.

Similarly, if we find for all then is the sensitive period for investment in ability . So for example given the above correlation matrix, if we find and , then we can say period *1* is a sensitive period for investment in ability at period *3*.

4. Current ability as a function of past investment

However, the *recursive* specification given by (1) above does not allow the estimation of the critical and sensitive periods of investment. This is because the recursive formula specifies current ability θ_t as a function of ability and investment at the period immediately before (*t*-1) only. In order to examine critical and sensitive periods, we need an alternative specification. Function (1) above can be re-written as:

(2)

Ability at period *t* is now expressed as a function of a series of earlier investments, and is no longer a function of ability at the period before (with the exception of θ_0, which is independent of any investment). In this paper, we refer to this as the '*non-recursive'* specification. Using this formulation we can define the critical and sensitive periods for investment.

Formally, is defined as the critical period for ability at time *t* if , and , for .

Similarly, is defined as the sensitive period for ability at time *t* if | | , | |, for .

While the *non-recursive* specification has the benefit of allowing the identification of the critical and sensitive periods of investment, it cannot be used to estimate the extent of self-productivity and dynamic complementarities. This is because the specification is no longer a function of past ability.

5. Linear specifications

The first generalization from the correlations in the section above would be to look at the partial correlations. Consider a linear function for the *non-recursive* specification, with no control variables:

(3)

is the i.i.d. error terms. gives the correlation between and , controlling for investments made in any other period. We are not assuming any correlation between investments across time, even though empirically, there may be.

In this specification, if we find , and , for all then we can say that is the critical period for investment in ability at . Similarly, if we find that , for all then is the sensitive period for investment in ability at .

An example of this approach can be seen in Hanusheck (1998, 2003), which undertakes the analysis of correlation and partial correlations between earlier investments and later achievements for school children in the US. Both studies suggest that there is no correlation or partial correlation between inputs and achievements. The studies however, came under criticism (see Krueger, 2003) for not including the 'other' factors that might be at play.

In order to control for other things we need to include vector in equation (3).

(4)

is the i.i.d. error terms. For the simplicity of exposition we have assumed to be a single variable, but this can be generalised to being a vector. In this specification, if we find , and , for all then we can say that is the critical period for investment in ability at . Similarly, if we find that , for all then is the sensitive period for investment in ability at .

For example, Todd and Wolpin (2004) use US data on children from birth to age 14, and apply this approach to analyse the relationship between lagged investments and later reading and maths scores. Their findings suggest that the maths score at age 14 are sensitive to investment made at age 12, whereas the reading scores at age 14 are sensitive to investment made at ages 3 to 5.

Elwell-Sutton et al (2011) use a similar approach to data on Chinese adults aged ≥50 to analyse the relationship between socioeconomic disadvantage at four life stages and self-rated health, chronic obstructive pulmonary disease and metabolic syndrome. Their findings suggest that socioeconomic conditions in the early stage are not critical for determining these health outcomes later in life.

Portrait et al (2011) use Dutch data from a natural experiment to find the effects of a famine which affected those living in cities in the west of Holland during the winter of 1944-45. They find that there is a significant association between severe under nutrition in girls (age 11-14) and a probability of developing diabetes mellitus and peripheral arterial disease at the age of 60-76. Concerning nutrition this implies that for women adolescence may be a critical period in determining long term health.

6. Latent specification for the *recursive* specification

So far we have assumed that we are able to observe directly. However, arguably, cognitive and non-cognitive abilities or health per se of an individual cannot be directly observed. If so, becomes a latent variable, and what we observe are specific measurements of . For example, we cannot observe directly the latent cognitive ability of a child, but we observe the test scores of the child on reading and maths. This poses a particular challenge for the recursive formulation, since unobserved current ability becomes a function of unobserved ability from the preceding period.

6.1 Structural model

Let be the vector of a latent cognitive skill across three time periods 1, 2, and 3; be the vector of parental investment in the child acros time periods 0, 1, and 2; and be a matrix of time invariant background characteristics. Going back to the recursive specification (1) above, a structural model explaining latent ability in terms of past ability, investment, and covariates is specified as:

$$(5)$$

where is parameter matrix associated with in equation (5):

$$[\quad]$$

and

$$[\quad]$$

is a multivariate normal error term with zero mean and diagonal covariance matrix. Equation of interest is (5), but it cannot be estimated as the left hand side is unobserved. What we need is a measurment model.

6.2 Measurement model

Let be a vector with different test scores, across the three time periods. For identification we need a minimum of two measurements for each time period. Only by *normalising*, or assuming that one or the other of these measurements has perfect correlation with the latent skill, can the coeffieicnt for the other measurement be identified. This means that in total we need a minimum of . The measurement model explains the observed tests scores (Y) in terms of the latent cognitive skill (θ) of the child, and is specified as:

$$(6)$$

where is a vector of intercepts; is a multivariate normal measurement error term, with zero mean and diagonal covariance matrix; and is a parmater matrix, often referred to as the 'factor loadings'.

If then will look like:

[ ]

From equation (5) we obtain:

where is an indentity matrix, and hence:

(7) [ ]

Since from (6) we have [ ] , parameter vectors and can be estimated. Combining this with (7) gives:

(8) [ ] [ ]

Now, observed test scores (Y) are expressed as a functoin of investments (I) and covariates (X), not of unobserved past ability. Thus we can estimate equaiton (8), which amounts to simultantously estimating the abiliy function for each period, and from there recovering the parameters of original interest.

Within this specification, if we find that and , then it can be interpeted to mean that period 1 is sensitive for ability at period 3. For example, Cunha and Heckman (2006) use the recursive latent specification to estimate the ability formation in children aged 6 to 13, using US data. They model a vector of abilities – cognitive and non-cognitive. Their estimates suggest that for cognitive ability the sensitive period for investment is between age 6 to 8, and for the non-cognitive abilities the sensitive periods of investments are between ages of 10 to 12 (see their Table 12).

However, strictly speaking, it is not possible to recover all critical periods of investment within this recursive formulation. Consider sepcifcation (8) (or (5) in structural terms) above: we can identify whether or not *t-1* is crtical for *t*; if , then we know that investment in period 1 is critical for ability in period 2. Nevertheless, we cannot know whether investment in period *t-2* is critical for ability in period *t*. More specifically, to examine whether investment in period 1 is critical for period 3, would require the calculation of ―, a derivative which does not exist for specification (5).

7. Latent specification for the *non-recursive* specification

A full exploration of critical periods of investment using a latent specification of ability is possible by building a structural model based on the *non-recursive* specification of skill formation. To the best of our knowledge, this has not been used in the empirical literature, but is a natural extension of the latent skill formation model to the identification of critical periods.

The structural model takes the same form as specification (4) above, but with the exception that current ability is now treated as a latent variable:

(9)

For the measurement model, let be a vector with different test scores. As before, for identification we need a minimum of two measurements, which means . The measurement model explains the observed tests scores ( ) in terms of the latent cognitive skill ( ) of the child, and is then specified as:

(10)

where is a vector of intercepts; is a multivariate normal measurement error term, with zero mean and diagonal covariance matrix; and is a matrix of factor loadings.

From (10) we have:

[ ] ,

combining this with (9) gives us:

(11) [ ] [ ]

Observed test scores ( ) are a function of all past investments and covariates. Thus, within this specificaiton we can define the critical and the sensitive periods just as we did for specification (4). If we find $\delta_{t*} \neq 0$, and $\delta_s = 0$, for all then we can say that is the critical period for investment in ability . Similarly if we find that $|\delta_{t*}| > |\delta_s|$, for all then is the sensitive period for investment in ability .

8. Conclusion

This paper has provided an overview of different quantitative methods available for the statistical analysis of longitudinal data regarding child development, and in particular the identification of critical and sensitive periods for later abilities. It has drawn heavily on the work on human skill formation developed by Heckman, which treats ability as a latent variable and explains its formation through simultaneous estimation of structural equations of investments and achieved abilities across time.

We explicitly distinguish between two specifications of the ability formation function: the recursive and the non-recursive. The recursive format explains current ability as a function of the ability and investment *at the immediately preceding period*: this can be used to capture self-productivity and dynamic complementarity. On the other hand, the non-recursive format explains current ability as a function of *a series of past investments*, with no reference to abilities (with the exception of the endowment level of ability, θ_0): this can be used to fully capture critical and sensitive periods of investment for later ability. It is important to note that neither format can be used to examine both self-productivity and dynamic complementarity *and* critical and sensitive periods.

If actual data on ability are regarded as measurements from an unobserved latent variable, then a structural approach is necessary. Typically, the literature has based such structural models on the recursive format of ability formation. In order to fully examine critical and sensitive periods of investments, we argue that the non-recursive formulation should be used. However, one major obstacle associated with the non-recursive formulation is its need for extensive data. While in the recursive format all information on past investments prior to *t*-1 is captured in ability at *t*-1, in the non-recursive format, the analysis requires data on investment from every period that is possibly critical.

What we have outlined above can be extended in several directions. The first is to allow for richer data. For example, ability (θ) can be a multidimensional vector so that reading ability is modelled simultaneously with maths ability. Or, covariates (X) capturing household characteristics or social environment may vary across time. The second class of extensions involve the use of non-linear functions. The above assumed that the effect of investment on ability is additively separable across time, but this can be relaxed. The third class of extensions is to treat the observed investments as measurements for directly unobservable latent investments.

References

Carneiro P and Heckman JJ. 2003. “Human Capital Policy,” IZA Discussion Paper No. 821.

Conti G and Heckman JJ. 2010. “Understanding the Early Origins of the Education-Health Gradient: Framework that can also be applied to analyze Gene-Environment interactions,” *Perspectives on Psychological Science*, Vol. 5 (5), pp. 585-605.

Conti G, Heckman JJ and Urzua. 2010. “The education-health gradient,” *American Economic Review: Papers and Proceedings*, Vol. 100, pp. 234-238.

Cunha F and Heckman JJ. 2006. ‘Investing in our young people,’ Working paper, University of Chicago.

Cunha F and Heckman JJ. 2008. “Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation,” *Journal of Human Resources*, Vol. 43, pp. 738-782.

Cunha F, Heckman J and Schennach S, 2010. ‘Estimating the Technology of Cognitive and Noncognitive Skill Formation’, *Econometrica,* Vol. 78, No. 3, pp 883-931

Currie J and Thomas D. 1995. “Does Head Start make a difference?” *American Economic Review*, vol. 85, pp. 341-364.

Elwell-Sutton T, Jiang CQ, Zhang W S, Cheng KK, Lam TH, Leung GM and Chooling CM. ‘Scocioeconomic influences at different life stages on health in Guangzhou, China’, *Social Science and Medicine,* Vol 72, No. 11,pp. 1884-1892

Hanushek EA. 1998. ‘Conclusions and controversies about the effectiveness of school resources,’ *Economic Policy Review*, Vol. 4, No. 1, pp. 11-28

Hanushek EA. 2003. ‘The Failure of Input-based Schooling Policies,’ *The Economic Journal*, vol 113 (485) pp. F64-F98.

Heckman JJ. 2008. “Role of income and family influence on child outcomes,” *Annals of the New York Academy of Sciences*, vol. 1136, pp. 307-323.

Krueger AB. 2003. ‘Economic Considerations and Class Size,’ *The Economic Journal*, vol 113 (485) pp. F34–F63.

Portrait F, Teeuwiszen E and Deeg D. 2011. ‘Early life undernutrition and chronic disease at older ages: The effects of the Dutch famine on cardiovascular diseases and diabetes.’ *Social Science and Medicine,* Vol. 73, No. 5, pp. 711-718.

Todd PE and Wolpin KI. 2003. ‘On the Specification and Estimation of the Production Function for Cognitive Achievement.’ *The Economic Journal*, vol 113 (485) pp. F3-F33.

Todd PE and Wolpin KI. 2004. ‘The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps,’ PIER Working Paper, 04-019.