DISCUSSION PAPER SERIES

The Timing of Earnings Sampling over the Life-Cycle and IV Identification of the Return to Schooling

Christian Belzil
Jörgen Hansen

The Timing of Earnings Sampling over the Life-Cycle and IV Identification of the Return to Schooling

Christian Belzil
Ecole Polytechnique,
ENSAE and IZA

Jörgen Hansen
Concordia University,
CIRANO, CIREQ and IZA

ABSTRACT

The Timing of Earnings Sampling over the Life-Cycle and IV Identification of the Return to Schooling[*]

We show that within a life-cycle skill accumulation model, IV identification of the return to schooling parameter is either achieved at any point in the life-cycle where the level of skills accumulated beyond school completion for compliers is exactly equal to the post-schooling skill level of non-compliers (the *Skill-Equality* condition), or when the skill-ratio is equal to the relative population proportions of non-compliers over compliers (the *Weighted-Skill-Ratio* condition). As a consequence, it is generally impossible to tie IV identification to any specific phase of the life-cycle and there cannot exist a generally acceptable "optimal" age to sample earnings for IV estimation. The practical example developed in the paper shows precisely how an instrument may fulfill identification at a multiplicity of ages, and how different instruments may achieve identification with specific sampling designs and fail to do so with others. Within a life-cycle skill accumulation data generating process, identification of the return to schooling requires not only implicit assumptions about the underlying model, but also assumptions about the validity of the specific age sampling distribution implied by the data.

Corresponding author:

Christian Belzil
Ecole Polytechnique
Département d'Economie
91128 Palaiseau Cedex
France
E-mail: christian.belzil@polytechnique.edu

1   Introductory Remarks

This paper addresses the following question:

What is the interaction between the timing of earnings sampling and IV identification of the return to schooling parameter in the case where the underlying data generating process is a life-cycle skill accumulation model?

The answer to this question is not trivial and has never been analyzed formally although issues surrounding the optimal sampling timing of labor market earnings over the life-cycle are gradually attracting attention.[1] Until now, the literature on IV estimation of the return to schooling remains purely cross-sectional and has ignored the potential sensitivity IV estimates to alternative age sampling strategies.[2]

To set the problem in a concrete framework, we use the common slope version of the wage equation so to eliminate ambiguity regarding which parameter should (or could) be targeted. In the model, the effect of schooling on wage is a structural policy invariant parameter that summarizes part of the skill formation technology, and it is identified by a moment condition. Consistent with the idea that agents may accumulate skills over the entire life-cycle, we assume that the econometrician's error term incorporates unobserved choices exercised by agents over the life-cycle and that he/she must therefore work with a mis-specified (or incompletely specified) outcome equation.

In the paper, we first discuss the issue at the analytical level. However, to provide a clear illustration of our results, we use a calibrated model of life-cycle skill accumulation to display the empirical relevance of our analysis. The model is used as the data generating process of the control group, and we implement IV estimation by merging data from the control group and

---

[1]Haider and Solon (2006) have investigated the correlation between life-cycle earnings (a measure of potential earnings) and current earnings with US panel data. They do this within a very specific model of earnings growth. They find that the correlation is at its highest around the mid-point of the life-cycle (toward age 40).

[2]One recent exception is Bhuller, Mogstad and Salvanes (2011) who investigate the presence of a so-called "life-cycle bias" in IV estimation with Norwegian data. However, they neither present a formal definition of the bias, nor do they discuss the dependence of the bias on the underlying data generating process. They claim that the optimal sampling age is between 30 and 40.

a large number of different treatments groups, generated by specific policy interventions. We focus our analysis on two types of instruments that are very popular in the literature on returns to schooling; one generated from unanticipated policies changing the cost of schooling, and one generated from changes in compulsory schooling.

The main results may be summarized as follows. First, to gain some economic insight about the interaction between identification and the timing of earnings sampling, we re-write the moment condition as a restriction between those affected by the policy generating the instrument (compliers) and those unaffected (non-compliers). We show that identification is either achieved at any point in the life-cycle where the level of skills accumulated beyond school completion by compliers and non-compliers are equal (the *Skill-Equality* condition), or when the skill-ratio is equal to the relative population proportions of non-compliers over compliers (the *Weighted-Skill-Ratio* condition).

In order to fulfill identification with a specific instrument, the econometrician may measure earnings at any of those specific points in the life-cycle. In cases where the econometrician can only access data in which earnings are subject to some age dispersion (as is the case in most empirical studies), the same conditions must apply with the exception that the moment condition involves not only an expectation over cross-sectional dispersion, but also one over age dispersion. The econometrician must therefore assume that the age distribution characterizing the sample data conforms to the identifying condition.

The practical example illustrated in the paper is particularly eloquent. It shows precisely how different instruments achieve identification with specific age sampling designs and fail to do so when other earnings sampling strategies are used. It also shows how a specific instrument may fulfill the *Skill-Equality* condition when earnings are measured at some age and may fulfill the *Weighted-Skill-Ratio* condition when they are measured at a different point in the life-cycle.

Our analysis of the interaction between identification and the timing of earnings sampling discloses three important implications. First, within a life-cycle skill accumulation model, there cannot exist a generally acceptable "optimal" age to sample earnings for IV estimation. Because identification is compatible with any level of differences in post-schooling skills between compliers and non-compliers, it is generally impossible to tie identification to any specific age, or any specific phase of the life-cycle. Second, the sensitivity

4

of IV estimates to the timing of earnings sampling suggests the possibility to use "age selection" as a strategy to achieve identification. Finally, our analysis clarifies the nature of implicit IV assumptions that have been discussed in the econometrics literature.[3] Within a life-cycle skill accumulation data generating process, identification of the return to schooling requires not only implicit assumptions about the underlying model, but also assumptions about the validity of the specific age sampling distribution used in a given study.

The residual parts of the paper are organized as follows. In Section 2, we present some analytical results. In Section 3 we develop a calibrated life-cycle accumulation model and use it to provide a graphical representation of the interaction between the timing of earnings sampling and identification. In Section 4, we discuss the main results emerging from IV estimation using different sampling strategies. The final section offers concluding remarks.

2 Re-Interpreting the IV Identifying Condition

As a starting point, we assume that the outcome equation is given by

$$w_{i,t_i} = \alpha + \lambda^s \cdot S_i + \varphi_{i,t_i} \tag{1}$$

where w_{i,t_i} is log earnings of individual i measured at the individual specific age denoted t_i, λ^s is the treatment effect of schooling on earnings (log), S_i denotes schooling and φ_{i,t_i} is an error term. Schooling has no time subscript since we assume that earnings are measured after all individuals have stopped formal education. As this stage, we interpret φ_{i,t_i} as the amount of post-schooling skills that has been accumulated endogenously by date t_i, without giving more details.[4]

We use the common slope version of the wage equation so to eliminate ambiguity regarding which treatment effect should (or could) be targeted.

---

[3]Heckman (1997), Vytlacil (2002), Heckman and Vytlacil (2005), and Heckman, Urzua and Vytlacil (2007) discuss implicit IV assumptions in terms of the specification of the first-stage model. In parallel, Rosenzweig and Wolpin (2000) and Keane (2010) have pointed out the behavioral implications underlying the IV identifying orthogonality condition.

[4]For instance, our definition of $\varphi(.)$ is general enough to encompass the famous Mincer model, in which all individuals accumulate post-schooling skills at an exogenous rate, and

In this model, the effect of schooling on wage is a structural policy invariant parameter that summarizes part of the skill formation technology. Because schooling has no causal effect on subsequent skill accumulation, there is no issue about defining the parameter conditionally on post-schooling choices or not.[5]

To motivate our approach, we assume that the econometrician does not observe $\varphi(.)$, and must therefore work with a mispecified (or incompletely specified) model.[6] As is commonly done in empirical work, the econometrician considers schooling as the only endogenous variable. One alternative could be to condition on age (t_i) or on potential experience $(t_i - S_i)$, but the model remains mispecified in both of these cases and our analysis of the moment conditions remains valid. We will return to this issue toward the end of the paper, but for the moment we focus on the empirical specification that seems to be favored by most empiricists (see Heckman, Lochner and Todd, 2005, for a discussion).[7]

We assume that the econometrician wants to estimate λ^s by IV methods using Z_i as an instrument. For the moment, we only assume that the instrument records an unanticipated policy change. The instrument Z is a binary indicator such that $Z_i = 1$ for treatment, $Z_i = 0$ for control. Although we have not introduced a specific data generating process for schooling and post-schooling skill accumulation, it should be understood that we consider policy interventions that are subject to a form of *Ex-Ante Randomization*. In other words, we assume that is the distribution of individual fixed endowments (the model primitives that are individual specific) is invariant to policy exposure. This will become clearer in Section 3.

---

in which
$$\varphi_{i,t_i} = \varphi_1 \cdot (t_i - S_i) + \varphi_2 \cdot (t_i - S_i)^2 + \varepsilon_{it_i}$$
where $\varphi_1 > 0$, $\varphi_2 < 0$, and where ε_{it_i} is an error term which may incorporate an individual specific effect.

[5] All the analysis presented in this section would carry through even if individuals were endowed with their own λ^s. All that is required is that a parameter of interest be identified by a moment condition.

[6] This is easily motivated by the observation that human capital is by definition unobservable. Only experience (usually defined as time spent in the market) may be observed.

[7] Because accumulated experience is by definition endogenous, many empirical labor economists are reluctant to include measures of experience (see Heckman, Lochner and Todd, 2006).

2.1 Introducing the Timing of Earnings Sampling

To proceed, we assume that the age at which earnings are sampled, denoted T_i, is a random variable with CDF $\Upsilon(t_i)$. The distribution of T reflects the fact that most cross-sections used to estimate the return to schooling by IV methods are subject to important age dispersion, thereby reflecting the age composition of the labor force at a specific point in time. In the extreme case where T is degenerate, all individuals in the cross-section are sampled at a same age. We assume that T_i is independent of Z_i, and that the lower support of T_i (denoted $t_{\min}$) is chosen so to exceed the maximum value of S_i, and therefore to avoid any censoring in schooling.

In the IV estimation procedure, λ^s is identified by an orthogonality condition. However, in the literature, dispersion in t_i across individuals is usually ignored. So, from now on, we re-interpret the IV identifying condition as one involving both cross-sectional dispersion as well as age dispersion (across individuals). As a starting point, the moment condition is stated as follows;

$$\int_{t_{\min}}^{t_{\max}} E(Z_i \varphi_i \mid t_i) d\Upsilon(t_i) = 0 \tag{2}$$

where the term $E(Z_i \varphi_i \mid t_i)$ refers to the expected value of the product of the instrument and the error term (φ) conditional on a measurement taken at fixed date t_i.

To gain some economic insight about the interaction between identification and the timing of earnings sampling, we re-write the moment condition as a restriction between those affected by the policy generating the instrument (compliers) and those unaffected (non-compliers). As will become clear later, it is particularly useful to consider such a representation of the moment condition within a life-cycle skill accumulation model, because differences in skill-accumulation progression between those affected and unaffected may easily be discussed on intuitive grounds. This condition may be viewed as an implicit assumption in any common slope classical IV model. However, and as far as we know, empiricists using IV methods never discuss identifying conditions in those terms. In most applications, the validity of an instrument is argued on the basis of ex-ante randomization.[8]

---

[8]In the LATE framework (in which slopes are individual specific), the restriction between compliers and non-compliers may be implied directly from the Independence Assumption (see Imbens and Angrist, 1994). However, in the applied literature relying on

First, we introduce a random variable, C_i which records the complying status. Ignoring the time subscript since schooling is completed by date t_i, we have

$$
\begin{aligned}
C_i &= 1 \text{ if } S_i(Z_i = 1) \neq S_i(Z_i = 0) \\
C_i &= 0, \text{if } S_i(Z_i = 1) = S_i(Z_i = 0).
\end{aligned}
$$

In large samples, the observed ratio of non-compliers over compliers is the sample analog of the following expression:

$$
\frac{\Pr(Z_i = 1, C_i = 0)}{\Pr(Z_i = 1, C_i = 1)} \equiv \frac{\Psi_0}{\Psi_1} \tag{3}
$$

From now on, we refer to $\frac{\Psi_0}{\Psi_1}$ as the *Noncomplying Odds Ratio*. This quantity, unlike $\varphi(.)$, is age invariant and has nothing to do with earnings sampling.

Because the treatment/control indicator is binary $(0,1)$, the identifying moment condition may be written as

$$
\begin{aligned}
\Pr(Z_i &= 1) \cdot \int_{t_{\min}}^{t_{\max}} E(\varphi_i \mid t_i, Z_i = 1) d\Upsilon(t_i) \\
&= \Psi_1 \cdot \int_{t_{\min}}^{t_{\max}} E(\varphi_i \mid t_i, Z_i = 1, C_i = 1) d\Upsilon(t_i) + \\
\Psi_0 \cdot \int_{t_{\min}}^{t_{\max}} E(\varphi_i & \mid t_i, Z_i = 1, C_i = 0) d\Upsilon(t_i) \\
&= 0
\end{aligned} \tag{4}
$$

When all individuals are sampled at age t (when the distribution of T is degenerate), we re-write the condition as

$$
\begin{aligned}
\Pr(Z_i &= 1) \cdot E(\varphi_i \mid Z_i = 1) \\
&= \Psi_1 \cdot E(\varphi_i \mid Z_i = 1, C_i = 1) + \Psi_0 \cdot E(\varphi_i \mid Z_i = 1, C_i = 0) \\
&= 0
\end{aligned} \tag{5}
$$

LATE, identification is also argued solely on the basis of ex-ante randomization.

2.1.1 The Skill-Equality Condition

A straight inspection of (4) or (5) reveals that the IV moment condition is met when

$$\int_{t_{\min}}^{t_{\max}} E(\varphi_i \quad | \quad t_i, Z_i = 1, C_i = 1)d\Upsilon(t_i)$$
$$= \int_{t_{\min}}^{t_{\max}} E(\varphi_i \mid t_i, Z_i = 1, C_i = 0)d\Upsilon(t_i)$$
$$= 0 \tag{6}$$

In the degenerate case, the condition is

$$E(\varphi_i \mid Z_i = 1, C_i = 1) = E(\varphi_i \mid Z_i = 1, C_i = 0) = 0 \tag{7}$$

This condition, labeled *Skill-Equality*, implies that for a given distribution of the sampling period, those affected ($C_i = 1$) and those unaffected ($C_i = 0$) have accumulated the same quantity of skills beyond schooling. This is essentially a no-selection condition.

2.1.2 The Weighted-Skill-Ratio Condition

Differences in post-schooling investment behavior between those affected and unaffected are also consistent with the fulfillment of the orthogonality condition. In such a case, identification of the return to schooling requires that

$$\frac{\int_{t_{\min}}^{t_{\max}} E(\varphi_i \mid t_i, Z_i = 1, C_i = 1)d\Upsilon(t_i)}{\int_{t_{\min}}^{t_{\max}} E(\varphi_i \mid t_i, Z_i = 1, C_i = 0)d\Upsilon(t_i)} = -\frac{\Psi_0}{\Psi_1} \tag{8}$$

Again, in a degenerate case, we rewrite (8) as

$$\frac{E(\varphi_i \mid Z_i = 1, C_i = 1)}{E(\varphi_i \mid Z_i = 1, C_i = 0)} = -\frac{\Psi_0}{\Psi_1} \tag{9}$$

We refer to this condition as the *Weighted-Skill-Ratio* condition. It implies that the ratio of skills accumulated beyond schooling by both compliers and non-compliers must be exactly equal to the relative proportion of non-compliers over compliers ($\frac{\Psi_0}{\Psi_1}$).

2.2   An Intuitive Analysis

The analysis presented above discloses the interaction between the age at which earnings are measured and IV identification of the return to schooling parameter. It may be summarized as follows.

First, any policy affecting only a subset of the population creates a heterogeneity gap between compliers and non-compliers at the level of the relevant model primitives. This heterogeneity gap, just like the population proportion of those affected, is age invariant and has nothing to do with the timing of earnings sampling.

Second, this heterogeneity gap translates into two different sequences of accumulated post-schooling skills (one for compliers and one for non-compliers). The elements of those sequences (which constitute the econometrician's error term) reflect differences in post-schooling skill accumulation between the two groups at a given point in the life-cycle. Those differences vary with age.

Third, IV identification is achieved at any point in the life-cycle where the level of skills accumulated beyond school completion for compliers is exactly equal to the post-schooling skill level of non-compliers (the *Skill-Equality* condition).[9] Alternatively, identification may be fulfilled at any point in the life-cycle when the ratio of skills accumulated beyond schooling by both compliers and non-compliers is exactly equal to the non-complying odds' ratio (the ratio of population proportions of those unaffected and those affected). This condition, which is labeled *Weighted-Skill-Ratio,* therefore implies the existence of some degree of selectivity between compliers and non-compliers.

In order to fulfill identification with a specific instrument, the econometrician may therefore measure earnings at any of those specific points in the life-cycle. In cases where the econometrician can only access data in which earnings are subject to some age dispersion (as is the case in most empirical studies), the same conditions must apply with the exception that the moment condition involves not only an expectation over cross-sectional dis-

---

[9]Note that a policy that would create no heterogeneity gap would not necessarily fulfill *Skill-Equality.* To see this, suppose that a generous tuition policy splits the population according to complying status into two identical groups (with the same heterogeneity components). Those affected (those who are more educated ex-post) would suffer a residual time horizon deficit due to schooling investment, and would not accumulate the same amount of .post-schooling skills (by any date t) than those unaffected.

persion, but also one over age dispersion. The econometrician must therefore assume (implicitly) that the age distribution characterizing the sample data conforms to the identifying condition.

One implication of our analysis is that there cannot exist a generally acceptable "optimal" age to sample earnings for IV estimation. It is generally impossible to separate the performance of an instrument from (i) the underlying model, (ii) the nature of the policy generating the instrument, (iii) and the timing of earnings sampling. Because IV identification is compatible with a wide range of differences in accumulated skills between compliers and non-compliers, it is generally impossible to tie identification to any specific age, or any specific phase of the life-cycle.

In order to provide a practical illustration of our results, we now turn to a calibrated model of life-cycle skill accumulation.

3 An Empirical Illustration

3.1 Motivation

To illustrate our results, we use a life-cycle dynamic skill accumulation model in which individuals maximize lifetime earnings by choosing the optimal time allocation between education and other activities stimulating post-schooling human capital accumulation. We analyze IV estimation of the return to schooling using a set of incentive-based instruments generated by changes in the cost of schooling (education subsidies) and by dictatorial interventions generated by changes in compulsory schooling.

To be faithful to the analysis presented in Section 2, we set the analysis in a classical model where the parameter capturing the effect of schooling is common across all individuals and in which there is no causal effect of schooling on subsequent skill accumulation. As was the case in Section 2, we assume that the econometrician has access to one instrument and is unable to condition on actual human capital investments (except for work experience, which we address in Section 4).

We now provide specific details about the model.

3.2 The Behavioral Model

Individuals allocate a total time endowment of 50 periods (years) between 3 states: schooling (s), and two distinct employment states (e and a) which differ in terms of their associated rate of skill production. State e is characterized by a low accumulation rate of human capital, while occupying state a produces human capital at a higher rate. Our model is a discrete choice pendant of the classical Ben-Porath model, and is based on the calibrated model of Belzil, Hansen and Liu (2012) which is used to evaluate compulsory schooling policies.[10] The choices are summarized in the binary indicators, d_{tk}, where $d_{tk} = 1$ when option k (s, w, a) is chosen at date t. The variables corresponding to the capitalized letters (S_t, E_t, A_t) are used to measure the number of periods accumulated in each state.[11] There is a maximum of 16 years of schooling attainable.

Individuals are risk neutral and maximize the expected value of lifetime net earnings, over the entire life-cycle. The state-specific utilities are defined below.

- **School**: The utility of individual i, at time t, who attends school (state s), denoted U_{it}^s, is

$$
\begin{aligned}
U_{it}^s &= \alpha_i^s + \alpha_1^s \cdot I(S_t \leq 4) + \alpha_2^s \cdot I(5 \leq S_t \leq 8) \ + \\
\alpha_3^s \cdot I(9 &\leq S_t \leq 12) \ + \alpha_4^s \cdot I(13 \leq S_t \leq 16) + \\
\alpha_5^s \cdot I(d_{t-1,s} &= 0) + \varepsilon_{it}^S
\end{aligned}
\tag{10}
$$

where $I(.)$ is the indicator function. The parameters $\alpha_1^s, \alpha_2^s, \alpha_3^s$ and α_4^s capture the variation in the utility of attending school across grade levels. These parameters reflect tuition costs and the like. The parameter α_5^S captures

---

[10]The version found in Belzil, Hansen and Liu (2011) incorporates dynamic effects of schooling. Other examples of dynamic skill accumulation models include the seminal piece by Ben-Porath (1967), Heckman, Lochner and Taber (1998) who consider on-the-job training, and Keane and Wolpin (1997) who consider occupation choices.

[11]In observational data, the pendant of state E could be full time employment with learning by doing, while state (A) could represent work, with on-the-job training. The distinction between Full-time employment (e) and Work and Training (a) is therefore in the intensity of human capital accumulation (a is the high intensity mode).

the psychic cost of attending school for those who would have interrupted their education. The term α_i^S represents individual heterogeneity in taste for schooling (academic ability). Finally, ε_{it}^S is a purely stochastic shock.

- **Employment and Post-Schooling Skill Accumulation**: The utility of occupying state e, U_{it}^e, depends only on the wage rate. The utility of occupying state a, U_{it}^a, is defined as the difference between the wage rate and the monetary equivalent of the psychic cost, $C_{it}^a()$. U_{it}^e, U_{it}^a, and $C_{it}^a()$, are given by the following equations;

$$U_{it}^e = W_{it} \tag{11}$$

$$U_{it}^a = W_{it} - C_{it}^a \tag{12}$$

$$C_{it}^a() = c_{0i}^a + \varepsilon_{it}^a \tag{13}$$

where W_{it} is the wage rate per unit of time, and where ε_{it}^a is a stochastic shock.

- **Earnings equation**: The earnings equation is given by the following expression:

$$\log W_{it} = w_{it} = \alpha + \lambda^s \cdot S_{it} + \lambda^e \cdot E_{it} + \lambda^a \cdot A_{it} + \varepsilon_{it}^w \tag{14}$$

where α is an intercept term, λ^s is the effect of schooling on wages, λ^e is the effect of employment on wages, λ^a is the effect of training on wages, and ε_{it}^w is a random shock.

- **The Bellman Equation:**

The problem is solved using recursive methods, and optimal choices may be characterized by a Bellman equation (Bellman, 1957). For each possible choice, there is a specific value function, $V_t^k(\Omega_t)$, equal to

$$V_t^k(\Omega_t) = U_t^k + \beta E \max\{V_{t+1}^1(\Omega_{t+1}), ..V_{t+1}^K(\Omega_{t+1}) \mid d_{kt} = 1\}$$

where β is the discount factor, and where Ω_t is the set containing all state variables known by the agent at t.

- **Individual Heterogeneity and Random Shocks:**

The heterogeneity distribution, $H_{\nu_i}(.)$, is specified as a multi-variate discrete distribution with R vectors of support points;

$$\upsilon_r = \{\alpha_r^S, \ c_{0r}^a; p_r\} \text{ for } r = 1, 2, ..20 \tag{15}$$

where p_r is the population proportion of type r. The full distribution is displayed in Table A1 (in appendix). The vector $\{\varepsilon_{it}^s, \varepsilon_{it}^a, \varepsilon_{it}^w\}$ is composed of i.i.d. mutually independent random shocks. Each one follows a Normal distribution with mean 0 and variance $\sigma(k)$ for $k = s, w, a$.

- **Model Calibration and Solution:**

To implement the model, we experimented with the parameters of the utility of attending school so to match moments similar to those observed in data such as the NLSY. In the end, we use the following values: $\alpha_1^s = -3$, $\alpha_2^s - 7$, $\alpha_3^s = -12$, $\alpha_4^s = -14$, and $\alpha_5^s = -18$. The returns to schooling (λ^s) is set at 0.06 (a value close to estimates reported in the structural literature), while the return to employment (set to 0.01) and training (set to 0.03) are chosen to reflect the fact that human capital accumulation is more intensive in state a than in state e. They also ensure that the average life-cycle earnings growth will lie between 1% and 2% per year (a well known stylized fact for the US). We set the standard deviations of all random shocks to 0.35

The discount factor is set to 0.95 As is relatively common in the literature, we solve the Bellman equations using simulated realizations of the random shocks, for each single type separately. Our solution method is exact to the extent that we solve value functions for each point in the state space (we do not use any approximation or interpolation methods).

3.3   Generating Control and Treatment Groups

To generate the control group, we simulate 50 years of choices and wage outcomes for 50,000 individuals using 2,500 realizations of the full vector of random shocks for each type. Descriptive statistics of the number of periods spent in each state is found in Appendix (Table A2). For the moment, it is sufficient to note that the average periods of schooling in the control group is between 6 and 7 years.

We now describe the policy interventions that are used to generate instruments. Each education subsidy consists of offering a reward conditional on attaining a specific grade. We consider different timings of the subsidies, as well as variations in the amount of the subsidy. To implement the subsidies, we simply need to adjust the relevant utility parameters (α_1^s, α_2^s, α_3^s or α_4^s) .We also consider the implementation of a policy that sets a minimum age for leaving school and dictates school attendance for the relevant periods.

In order to build the treatment groups, we proceed as we did for the control groups and we simulate 50 years of choices and wage outcomes under each policy intervention. We end up with 100,000 observations (50,000 in control and 50,000 in treatment). We report results for 4 different subsidies of \$2 set at Level 1, Level 2, Level 3 and Level 4. We also report results for 2 compulsory schooling policies (1 year and 2 years).

3.4   The Average Heterogeneity Gap

We start by examining the differences in heterogeneity induced by various policies. Understanding the nature of the heterogeneity gap is essential since it is the generating force behind differences in life-cycle skill accumulation behavior between compliers and non-compliers. For expositional purpouses, and without loss of generality, we focus on the average differences and examine differences in average cost of occupying state a between those affected and those unaffected. Those quantities, $E(c_0^a \mid C_i = 1)$ and $E(c_0^a \mid C_i = 0)$,are reported in Table 1.

Not surprisingly, compulsory schooling regulations tend to affect the bottom tail of the heterogeneity distribution, as the average psychic cost of occupying state a is very high for both compulsory schooling regulations (the average cost for those affected is between 10 and 8 while it is between 5 and 4 for those unaffected). Compulsory schooling policies therefore create a huge heterogeneity gap.

Education subsidies, because they can be set conditional on attending different grade levels, do not necessarily create sizeable differences in mean ability between those affected and unaffected. For instance, the average cost of those affected ranges from 7.2 (for the subsidy conditional on attending *Level 1*) to 3.8 (for the subsidy set at *Level 4)*. One of the subsidies (the one set at Level 2) raises particular interest. With an average cost equal to 5.9 for those affected and 5.3 for those unaffected, it is the only policy

15

intervention that seems to split the population into two groups that are relatively similar (on average). Effectively, compliers are only slightly less able than non-compliers. So, among all those policies considered, the Level 2 subsidy is the one that generates the smallest heterogeneity gap.

Table 1
The Heterogeneity Gap

	$E(c_0^a \mid C_i = 1)$	$E(c_0^a \mid C_i = 0)$
Subsidies		
Level 1	7.2	4.3
Level 2	5.9	5.3
Level 3	4.6	6.5
Level 4	3.8	6.8
Compulsory Schooling		
1 year	9.6	5.4
2 years	8.7	4.7

3.5 Graphical Representations of the IV Identifying Condition

One of the major advantages of representing the identifying condition as one involving differences between compliers and non-compliers is that it admits an easily interpretable graphical representation. Because one condition (*Skill-Equality*) is expressed in terms of skill levels, while the other *(Weighted Skill-Ratio)* is expressed as a ratio, we perform a separate graphical representation for each condition.

In Figures 1 to 6, we plot the full life-cycle sequences of mean accumulated skills for compliers, $E(\varphi_i \mid t, Z_i = 1, C_i = 1)$, and non-compliers $E(\varphi_i \mid t, Z_i = 1, C_i = 0)$. Each point represents the average skills accumulated beyond schooling by a specific age. For each policy intervention, we plot the sequence for the entire life-cycle (period 1 to period 50) even though the very early phase of the life-cycle (the first 10 years) is unlikely to correspond to actual sampling strategies used in applied work. The *Skill-Equality* condition is fulfilled at any point (age) at which the complier and non-complier sequences

16

cross and are both equal to 0. Any of those points represent an age at which there is no selectivity between compliers and non-compliers in terms of their average level of post-schooling skills.

Figure 1 indicates that the instrument generated by the Level 1 subsidy fulfills *Skill-Equality* when earnings are measured at age 25-26. Not surprisingly, *Skill-Equality* may also be fulfilled very early in the life-cycle (between period 1 and 4) because post-schooling skills are naturally compressed in such an early phase. The domination of those affected in the early phase of the life-cycle is easily explained by the fact that they have (on average) a lower level of ability and obtain a lower level of schooling than those unaffected. As a consequence, they start working much earlier.

In Figure 2, which is devoted to the subsidy set at Level 2, we observe something particularly interesting. Aside from a trivial early life-cycle sampling strategy, the Level 2 subsidy never fulfills *Skill-Equality*. To understand why, it is informative to examine its associated heterogeneity gap. As already noted, the subsidy set at Level 2 creates the smallest heterogeneity gap of all those 6 policies considered in the paper. As a consequence, those affected and those unaffected accumulate a comparable level of skills per period spent in the labor market. However, those affected having accumulated more schooling than those unaffected, they are also forced to sacrifice potential experience and are unable to take over those unaffected.

The subsidy set at Level 3 (Figure 3) fulfills *Skill-Equality* around age 40, while the Level 4 subsidy (Figure 4)fulfills the condition at an earlier stage (around 27). Again, this illustrates how different instruments may fulfill the identifying assumption at different ages.

Finally, Figure 5 and Figure 6 disclose both the specificities of the compulsory schooling policies, as well as some resemblance with the subsidy set at Level 1. In both cases, those affected (who obtain a very low schooling level) dominate those unaffected in the early phase of the life-cycle but are dominated beyond a certain age (around 33 or 34). This is the consequence of the huge average heterogeneity (ability) gap induced by compulsory schooling. Those affected are those who obtain a very low level of education and work much earlier than those unaffected.

In Figures 7 to 12, we plot the full sequence of accumulated skill ratios $(\frac{E(\varphi_i|t_i, Z_i=1, C_i=1)}{E(\varphi_i|t_i, Z_i=1, C_i=0)})$ for each policy intervention. In each of those figures, the horizontal line represents the non-complying odds' ratio $(\frac{\Psi_0}{\Psi_1})$. The condi-

tion labeled *Weighted-Skill-Ratio* is fulfilled at any point (age) at which the accumulated skills ratio crosses the horizontal line. In cases where the *Non-complying Odds Ratio* is large (a small fraction of the population is affected), the skill ratio must also be large in absolute value. In other words, the *Weighted-Skill-Ratio* condition is fulfilled at a particular age when the imbalances between compliers and non-compliers match perfectly their relative population proportion.

In our model, there exist a multiplicity of ages at which the skill ratio is equal the non-complying odds ratio. This is clearly the case for the subsidies set at Level 3 (Figure 9) and Level 4 (Figure 10). This is also the case with the 1-year mandatory schooling regulation. The 2-year compulsory schooling instrument is actually the only one for which the *Weighted-Skill-Ratio* condition is unambiguously fulfilled at a single age.

To summarize, each instrument fulfills identification at multiple sampling periods, and both the number of periods at which identification is achieved and the identity of those periods vary across instruments.

4   IV Estimates

In cross-sectional studies of the return to schooling, estimates obtained for a sample composed of individuals of the same age are practically never reported.[12] So, to illustrate the importance of using different sampling strategies, we also report IV estimates obtained for different age sampling distributions. In the presence of age dispersion in earnings sampling, identification is achieved when either (6) or 8) is verified. We perform IV estimation at 4 different age sampling distributions that differ only up to the first moment. For each individual, the sampling period follows a uniform distribution over a specific interval. In total, we considered 4 different intervals. Details are provided in Table 2.

One of the most logical implications of a lifecycle skill accumulation model is the dependence of the OLS estimate on the timing of earnings sampling

---

[12]There may be multiple reasons for this. One reason could be the impossibility to obtain independent variation in potential experience, in case one would like to condition on it. Another reason could be sample size. A third (and more likely) reason is that the sensitivity of IV estimates to change in age composition creates a real challenge to those who report IV estimates.

since the correlation between schooling and post-schooling skills (which generates the OLS bias) is expected to be initially negative but should increase with age. This is easily verified upon examining Table 2.

Table 2
Earnings Sampling Strategies

		Age		OLS	Sample Size
	min.	max.	mean		
Sample 1	10	20	15	0.0508	100,000
Sample 2	20	30	25	0.0582	100,000
Sample 3	30	40	35	0.0751	100,000
Sample 4	40	50	45	0.0827	100,000

Note: All OLS estimates are significant at the 0.5% level.

We now turn to IV estimates obtained from various sampling methods. Dropping the time subscript so to facilitate notation and to be consistent with cross-sectional studies, the outcome equation analyzed by the econometrician is given by the following equation:

$$w_i = \alpha + \beta \cdot S_i + \varepsilon_i \tag{16}$$

where w_i is log earnings of individual i, β (or λ^s) is the treatment effect of schooling on earnings (log) and is equal to 0.06, S_i denotes schooling, and where ε_i is the econometrician's error term. The term ε_i is the amount of post-schooling skills accumulated is generated by the individual specific solution to the dynamic programming model. The objective of the econometrician is to estimate β (or λ^s) by IV methods using Z_i as an instrument. The instrument Z is a binary indicator such that $Z_i = 1$ for treatment, $Z_i = 0$ for control. Using standard vector notation, the IV estimator is defined as

$$\hat{\beta}_{IV} = (Z'S)^{-1}Z'W \tag{17}$$

where W and S denote the relevant vectors.

All IV estimates reported are obtained for 100,000 realized wages, for a given sampling strategy. For each estimate, we compute a standard error

using 200 bootstrap replications. The estimates are found in Table 3A (age 10-age 20), Table 3B (age 20-age 30), Table 3C (age 30-age 40) and Table 3D (age 40-age 50). We also report the mean post-schooling skills values for both groups, $E(\varepsilon_i) \mid C_i = 1$, $E(\varepsilon_i) \mid C_i = 0$ (after suitable normalization), the post-schooling skill ratio $\mid \frac{E(\varepsilon_i|C_i=1)}{E(\varepsilon_i|C_i=0)} \mid$, and the *noncomplying odds ratio* $\frac{\Pr(C_i=0)}{\Pr(C_i=1)}$.

Overall, the results are consistent with the graphical analysis performed in the previous section in the sense that i) the majority of instruments fail to fulfill the identifying condition for a given sampling method, and that (ii) most sampling methods fail to fulfill identification for a given instrument. However, it should be clear that there is no guarantee that identification is obtained when the average sampling age (for the non-degenerate case) happens to correspond to a specific age at which identification is fulfilled in the degenerate case.[13]

First, we note that only one instrument (the subsidy set at Level 4) seems close to fulfill *Skill-Equality*. Its estimate, found in Table 3B and equal to 0.0618, is close to the population parameter (0.0600) when earnings are measured between age 20 and age 30. This may be verified upon examination of the quantities $E(\varepsilon_i) \mid C_i = 1$ and $E(\varepsilon_i) \mid C_i = 0$. Those numbers, equal to -0.0043 for compliers and 0.0053 for non-compliers, are practically equal to 0 and therefore imply that at that particular point in the life-cycle, there is absence of selectivity with respect to skill accumulation on the basis of the complying status.

As expected, the instrument generated by the Level 4 subsidy cannot fulfill *Skill-Equality* when earnings are measured between period 30 and 40 (Table 2C) or between period 40 and 50 (Table 3D). The results indicate that the balance between compliers and non-compliers is substantially perturbed by the change in sampling timing. In sample 3 (age 30-40), the average skills for those affected (equal to 0.0727) is now much higher than the corresponding average for those unaffected (-0.0288). As a result, the IV estimate becomes 0.0744. As already noted when commenting Figure 4, the imbalance

---

[13]This is because

$$\int_{t_{\min}}^{t_{\max}} E(Z_i\varphi_i \mid t_i)d\Upsilon(t_i) = 0 \ \nrightarrow E(Z_i\varphi_i \mid t_i) \mid_{t_i=E(T)} = 0$$

is even more important when earnings are sampled between period 40 and 50, as the average skills of those affected are equal to 0.2385 and those of the non-compliers equal to -0.1033. The consequence is an increase in the IV estimate now equal to 0.0984.

We now examine the instrument generated by the subsidy set at Level 3. This instrument performs well when earnings are measured between age 30 and 40 (unlike for the level 4 subsidy) as indicated by the IV estimate equal to 0.0626. However, it approaches identification because it is close to fulfill the *Weighted-Skill-Ratio* condition. This is revealed by the ratio (equal 3.28) which is close to the noncomplying odds ratio (3.74). In such a case, the policy generates the right level of selectivity between compliers and non-compliers. However, the IV estimate obtained when earnings are measured between age 40 and 50 is equal to 0.0838.

Finally, we turn to instruments generated from changes in compulsory schooling. None of the compulsory schooling policies are capable of identifying the population parameter. Our results indicate that instruments generated from compulsory schooling tend to perform best when earnings are measured in the early phase of the lifecycle, but are unable to achieve either of the two conditions required.

4.1   Extension to Controls for Experience

As mentioned earlier, applied econometricians may sometimes be tempted to add additional controls in the wage equation. This may be achieved by introducing an element capturing the effect of age (say a parametric function $\varpi(t)$), or a function capturing the effect of potential experience (say $\varpi(t - S_i)$). In either of those cases, our analysis remains valid as long as we reinterpret the error term as $\varphi_{it} - \varpi(.)$, with the only difference that an additional moment restriction must be introduced so to capture orthogonality between $\varpi(.)$ and the new error term. Obviously, such an approach is not advocated by all applied econometricians. Because accumulated experience is by definition endogenous, many empirical labor economists are reluctant to include measures of experience (Heckman, Lochner and Todd, 2006).

However, to remove any ambiguity, we have performed IV estimates with controls for experience and experience squared. Those estimates, found in Table A3 in Appendix, indicate clearly that controlling for experience (and its square) is not sufficient to improve the performance of IV estimates. Out

of all those estimates, only the Level 2 subsidy (with sample 1) and the Level 3 subsidy (with Sample 2) seem to perform accurately. Indeed, the overall performance is inferior, and many estimates have become negative.

5   Concluding Remarks

In conclusion, our analysis of the interaction between identification and the timing of earnings sampling discloses three important elements.

First, one obvious implication of a life-cycle skill accumulation model is that it is generally impossible to tie identification to any specific age, or any specific phase of the life-cycle. There cannot exist any general rule regarding an "optimal" age to sample earnings for IV estimation, which could be transported across data sets.[14]

Second, the sensitivity of IV identification to the timing of earnings sampling suggests the possibility to use "age selection" as a strategy to estimate the parameter of interest, at least in cases where cohort effects may be ruled out.

Finally, our analysis clarifies the nature of implicit IV assumptions that have been discussed in the econometrics literature.[15] Within a life-cycle skill accumulation data generating process, identification of the return to schooling requires not only implicit assumptions about the underlying model, but also assumptions about the reliability of the specific age sampling distribution used in a given study.

---

[14]Our results are therefore at odds with findings reported in Bhuller, Mogstad and Salvanes (2011).

[15]Heckman (1997), Vytlacil (2002), Heckman and Vytlacil (2005), and Heckman, Urzua and Vytlacil (2007) discuss implicit IV assumptions in terms of the specification of the first-stage model. In parallel, Rosenzweig and Wolpin (2000) and Keane (2010) have pointed out the behavioral implications underlying the IV identifying orthogonality condition.

<div align="center">

Table 3A
IV Estimates for Various Policy Interventions
Sample 1: 10-20

</div>

	IV estimate (st.error)	$E(\varepsilon_i)$ $C_i=1$	$C_i=0$	$\left\lvert \frac{E(\varepsilon\mid C_i=1)}{E(\varepsilon_i\mid C_i=0)} \right\rvert$	$\frac{\Pr(C_i=0)}{\Pr(C_i=1)}$
Subsidies					
Level 1	0.0500 (0.002)	0.0174	-0.0268	0.649	1.0733
Level 2	0.0491 (0.002)	-0.0351	0.0022	16.349	5.2282
Level 3	0.0475 (0.003)	-0.0398	0.0064	6.215	3.7362
Level 4	0.0400 (0.002)	-0.0410	0.0087	3.2974	2.8739
Compulsory					
1 year	0.0501 (0.024)	0.0425	-0.0049	8.6578	9.6837
2 years	0.0503 (0.005)	0.0336	-0.0133	2.5304	3.0561

Note: All IV estimates reported are obtained for 100,000 realized wages, for a given sampling strategy. For each estimate, we compute a standard error using 200 bootstrap replications.

Table 3B
IV Estimates for Various Policy Interventions
Sample 2: 20-30

	IV estimate (st.error)	$E(\varepsilon_i)$ $C_i=1$	$C_i=0$	$\left\| \frac{E(\varepsilon\|C_i=1)}{E(\varepsilon_i\|C_i=0)} \right\|$	$\frac{\Pr(C_i=0)}{\Pr(C_i=1)}$
Subsidies					
Level 1/$1	0.0495 (0.0015)	-0.0047	0.0070	0.6767	1.0733
Level 2/$1	0.0501 (0.0021)	-0.0580	0.0065	8.9414	5.2282
Level 3/$1	0.0576 (0.0029)	-0.0524	0.0132	3.9570	3.7362
Level 4/$3	0.0618 (0.0015)	-0.0043	0.0053	0.8105	1.7125
Compulsory					
1 year	0.0440 (0.0183)	0.0204	-0.0031	6.6417	9.6837
2 years	0.0484 (0.0049)	0.0115	-0.0066	1.7466	3.0561

Note: All IV estimates reported are obtained for 100,000 realized wages, for a given sampling strategy. For each estimate, we compute a standard error using 200 bootstrap replications.

<div align="center">

Table 3C
IV Estimates for Various Policy Interventions
Sample 3: 30-40

</div>

	IV estimate (st.error)	$E(\varepsilon_i)$ $C_i=1$	$C_i=0$	$\mid \frac{E(\varepsilon\mid C_i=1)}{E(\varepsilon_i\mid C_i=0)} \mid$	$\frac{\Pr(C_i=0)}{\Pr(C_i=1)}$
Subsidies					
Level 1	0.0475 (0.002)	-0.0421	0.0259	1.6231	1.0733
Level 2	0.0504 (0.002)	-0.0948	0.0137	6.9269	5.2282
Level 3	0.0626 (0.003)	-0.0385	0.0117	3.2813	3.7362
Level 4	0.0744 (0.002)	0.0727	-0.0288	2.5245	1.7125
Comp					
1 year	0.0200 (0.002)	-0.0170	0.0003	53.0000	9.6837
2 years	0.0419 (0.006)	-0.0258	0.0043	6.0093	3.0561

Note: All IV estimates reported are obtained for 100,000 realized wages, for a given sampling strategy. For each estimate, we compute a standard error using 200 bootstrap replications.

Table 3D
IV Estimates for Various Policy Interventions
Sample 4: 40-50

	IV estimate (st.error)	$E(\varepsilon_i)$ $C_i=1$	$C_i=0$	$\left\lvert \frac{E(\varepsilon\lvert C_i=1)}{E(\varepsilon_i\lvert C_i=0)} \right\rvert$	$\frac{\Pr(C_i=0)}{\Pr(C_i=1)}$
Subsidies					
Level 1	0.0475 (0.002)	-0.1153	0.0942	1.2248	1.0733
Level 2	0.0512 (0.003)	-0.1646	0.0274	6.0091	5.2282
Level 3	0.0838 (0.003)	-0.0211	0.0037	5.7890	3.7362
Level 4	0.0984 (0.002)	0.2385	-0.1033	2.3096	1.7125
Compulsory					
1 year	0.0204 (0.03)	-0.0902	0.0073	12.4441	9.6837
2 years	0.0418 (0.007)	-0.0991	0.0283	3.5051	3.0561

Note: All IV estimates reported are obtained for 100,000 realized wages, for a given sampling strategy. For each estimate, we compute a standard error using 200 bootstrap replications.

References

[1] Bellman, Richard (1957) *"Dynamic Programming"* Princeton, New-Jersey, Princeton University Press

[2] Belzil, Christian (2007) "The Return to Schooling in Structural Dynamic Models: A Survey of the Literature" *The European Economic Review,* vol. 51, no 5, 1059-1105

[3] Belzil, Christian and Hansen, Jörgen (2002a) "Unobserved Ability and the Return to Schooling" *Econometrica, 70, 575-91.*

[4] Belzil, Christian and Hansen, Jörgen (2007) "A Structural Analysis of the Correlated Random Coefficient Wage Regression Model", *Journal of Econometrics*, vol 140, 2, October, 827-848

[5] Ben Porath, Yoram (1967): "The Production of Human Capital and the Life Cycle of Earnings."*Journal of Political Economy,* 75(4), pp. 352-365.

[6] Bhuller, Manudeep, Mogstad Magne, and Salvanes Sjell (2011) "Life-cycle bias and the returns to schooling in current and lifetime earnings" Statistics Norway Discussion Paper 666

[7] Card, David (1999) "The Causal Effect of Education on Earnings" Handbook of Labor Economics, edited by David Card and Orley Ashenfelter, North-Holland Publishers.

[8] Haider, Steve. and Solon, Gary (2006) "Life-Cycle Variation in the Association between Current and Lifetime Earnings, *American Economic Review*, 96, 1308–1320.

[9] Heckman, James, Lance Lochner and Chris Taber (1998) "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents," *Review of Economic Dynamics.*

[10] Heckman, James (1997) "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources, 32 (3), 441-62.*

[11] Heckman, James and Vytlacil, Edward (2005) "Structural Equations, Treatment Effects and Econometric Policy Evaluations" *Econometrica. 73.*

[12] Heckman, James, Sergio Urzua and Vytlacil, Edward (2007) "Understanding Instrumental Variables in Models with Essential Heterogeneity", *Review of Economics and Statistics*

[13] Imbens, Guido and Angrist, Joshua (1994) "Identification and Estimation of Local Average Treatment Effects" *Econometrica, 62, 467-76.*

[14] Imbens, Guido (2009) "Better Late than Nothing" , NBER Working Paper

[15] Keane, Michael (2010) "Structural vs. Atheoretic Approaches to Econometrics, *Journal of Econometrics, 156, 3-20.*

[16] Keane, Michael and Wolpin, Kenneth (1997) "The Career Decisions of Young Men" *Journal of Political Economy, 105, 473-522.*

[17] Rosenzweig Mark and K.Wolpin (2000) "Natural Natural Experiments in Economics" *Journal of Economic Literature*, December, 827-74.

[18] Stokey, N., Lucas, R.E. (with Ed Prescott) (1989). Recursive Methods in Economic Dynamics. Harvard University Press. Cambridge, Massachusetts.

[19] Vytlacil, Edward (2002) "Independence, Monotonicity, and Latent Index Models: An Equivalence Result", *Econometrica, 70(1): 331-341*

Table A1
The Heterogeneity Distribution

type	α^S	λ_s	c_0^a	proportion
1	-2.990	0.001	9.895	0.05
2	-2.750	0.001	9.955	0.05
3	-2.750	0.001	9.955	0.05
4	-1.790	0.015	8.695	0.05
5	-1.790	0.015	8.695	0.05
6	-1.310	0.023	8.215	0.05
7	-1.310	0.023	8.215	0.05
8	-0.470	0.035	7.375	0.05
9	0.130	0.044	6.775	0.05
10	0.550	0.050	6.355	0.05
11	0.844	0.055	6.061	0.05
12	1.324	0.062	5.581	0.05
13	1.570	0.066	5.335	0.05
14	2.050	0.073	4.855	0.05
15	2.530	0.080	4.375	0.05
16	2.770	0.084	4.135	0.05
17	3.010	0.087	3.895	0.05
18	3.250	0.091	3.655	0.05
19	3.484	0.094	3.421	0.05
20	3.964	0.102	2.941	0.05
Mean	1.188	0.060	5.717	-
St Dev.	2.062	0.031	2.062	-

Table A2
Life Cycle Choices in the Control Group:
Common Slope Model

Accumulated number of periods

year	In School	Work (learning)	Work (Training)
1	0.906	0.094	0.00
10	5.475	4.525	0.010
20	6.494	12.615	0.891
30	6.494	22.006	1.500
40	6.494	29.564	3.492
50	6.494	38.119	5.387

Table A3
IV estimates of the return to schooling obtained
with controls for experience and experience squared

	IV estimate (st.error)			
	Sample 1	Sample 2	Sample 3	Sample 4
Subsidies				
Level 1	0.0527 (0.004)	0.0485 (0.005)	0.0274 (0.006)	-0.0030 (0.007)
Level 2	0.0636 (0.008)	0.0407 (0.008)	0.0287 (0.009)	-0.0168 (0.013)
Level 3	0.0728 (0.020)	0.0616 (0.013)	0.0737 (0.014)	0.0954 (0.015)
Level 4	0.0385 (0.012)	0.0815 (0.011)	0.1555 (0.014)	0.1936 (0.017)
Compulsory				
1 year	-0.0465 (0.041)	0.0392 (0.054)	-0.0441 (0.082)	-0.0627 (0.098)
2 years	-0.0479 (0.011)	0.0505 (0.013)	0.0169 (0.016)	-0.0035 (0.020)

Note: All IV estimates reported are obtained for 100,000 realized wages, for a given sampling strategy. For each estimate, we compute a standard error using 200 bootstrap replications.

Figure 1. Life-cycle sequences of mean accumulated skills using a level one subsidy experiment.
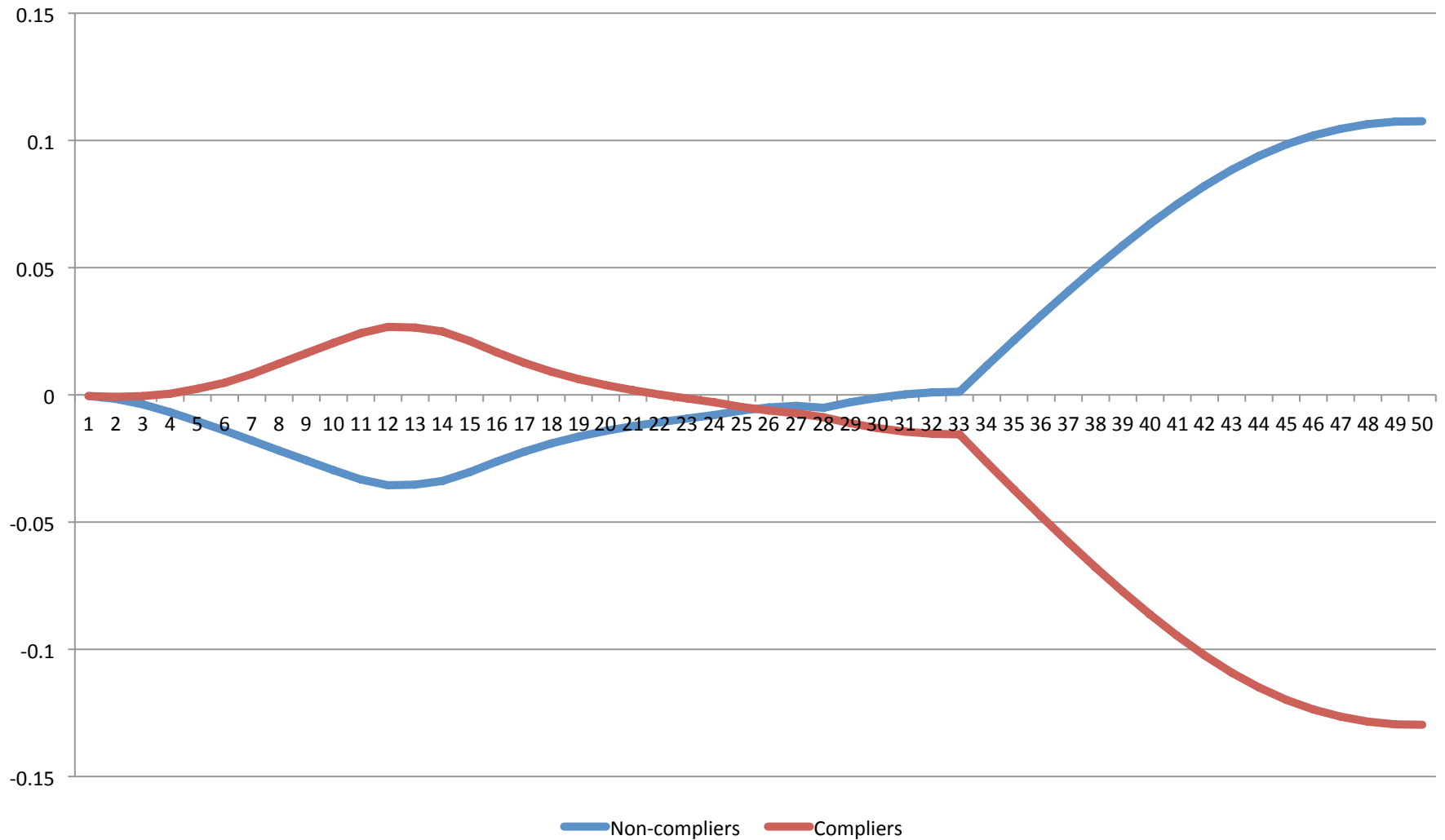
Figure 2. Life-cycle sequences of mean accumulated skills using a level two subsidy experiment.
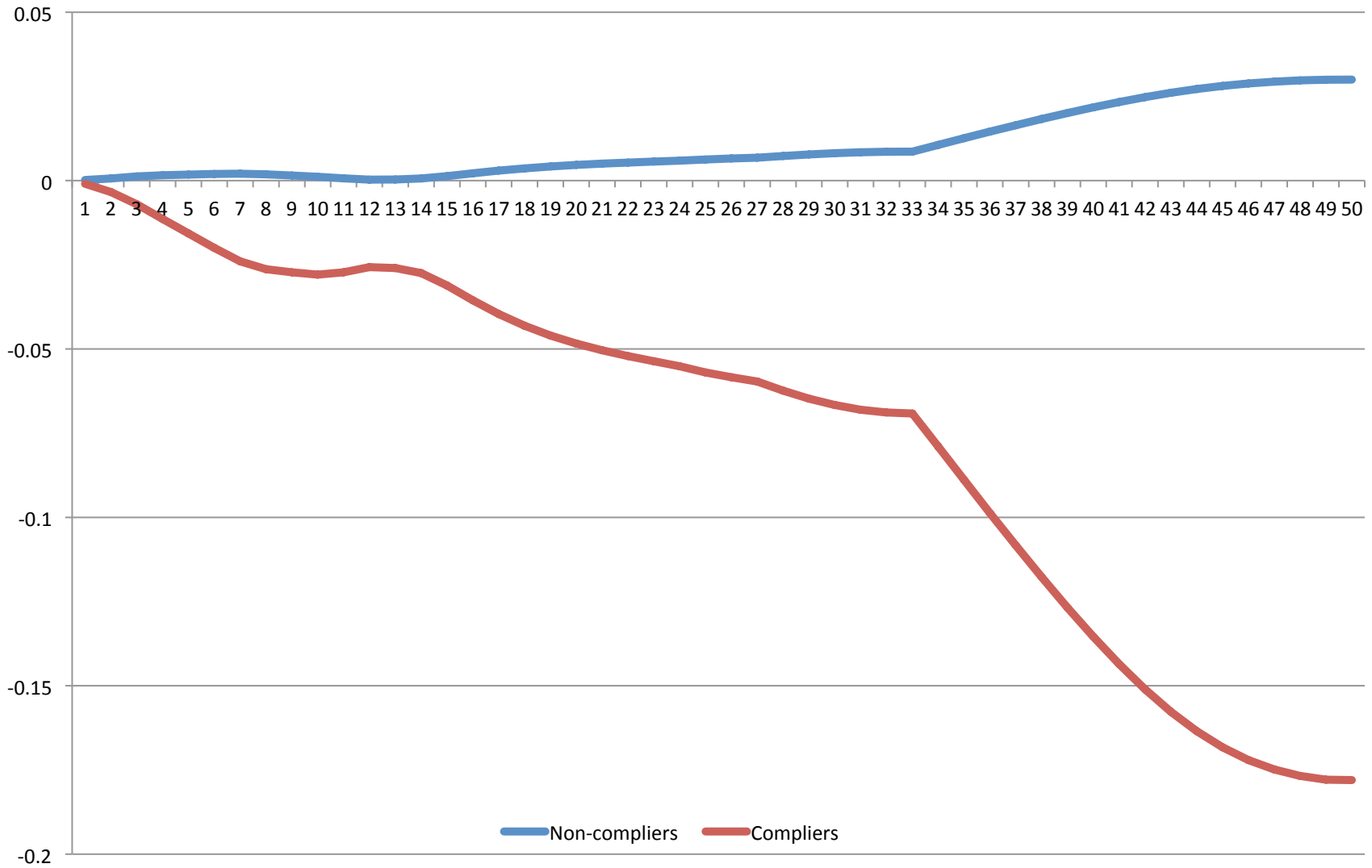
Figure 3. Life-cycle sequences of mean accumulated skills using a level three subsidy experiment.
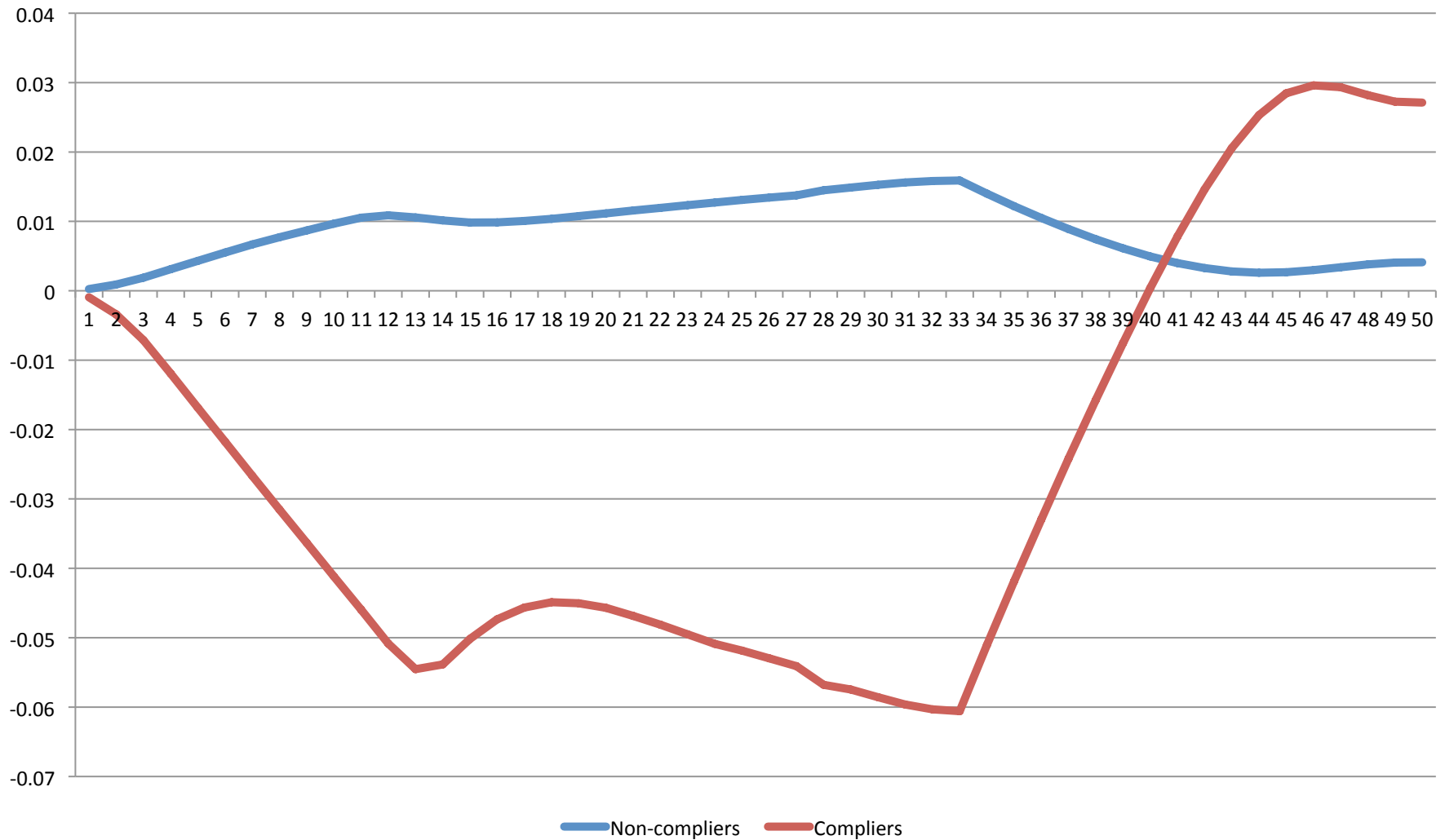
Figure 4. Life-cycle sequences of mean accumulated skills using a level four subsidy experiment.
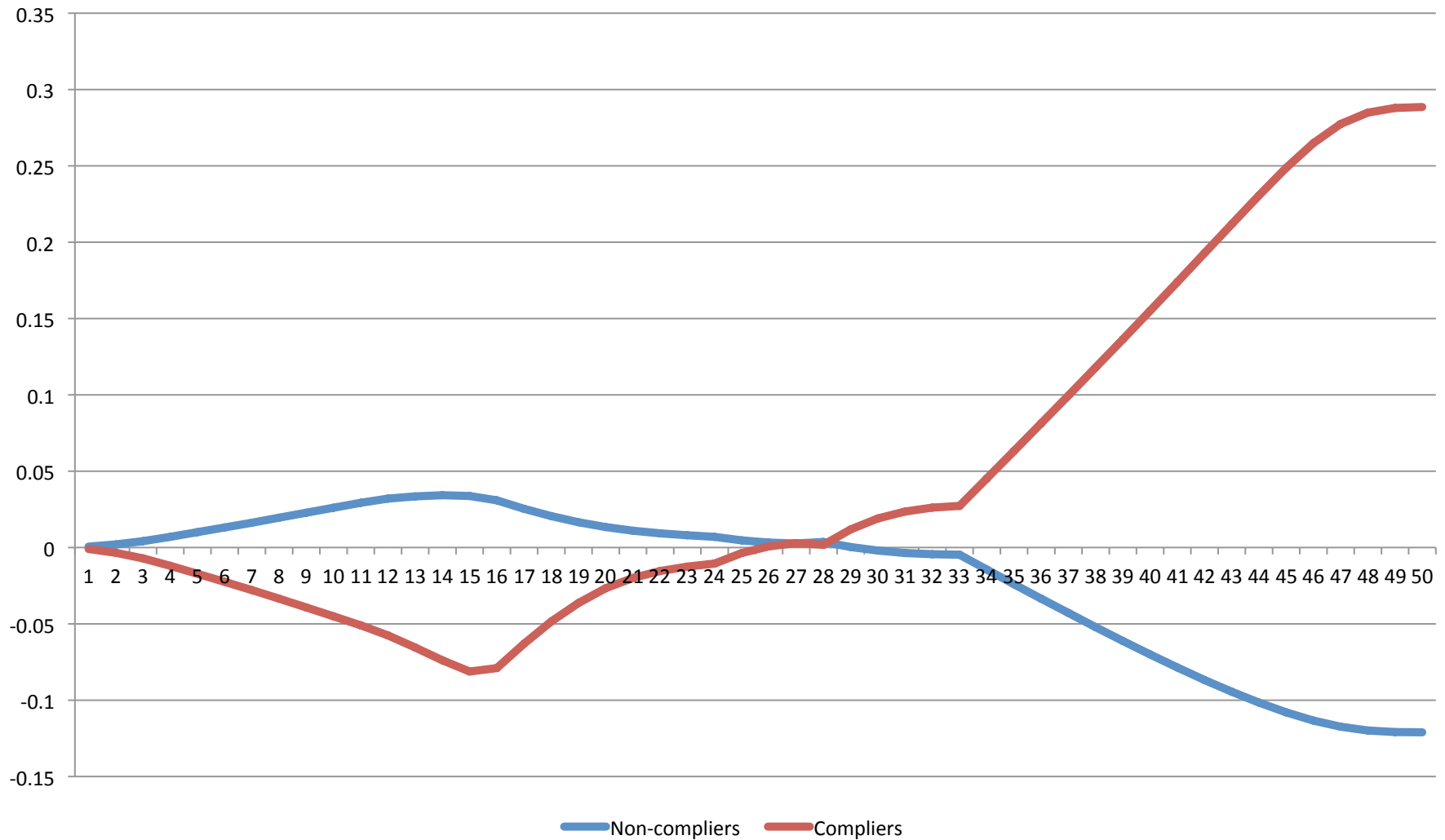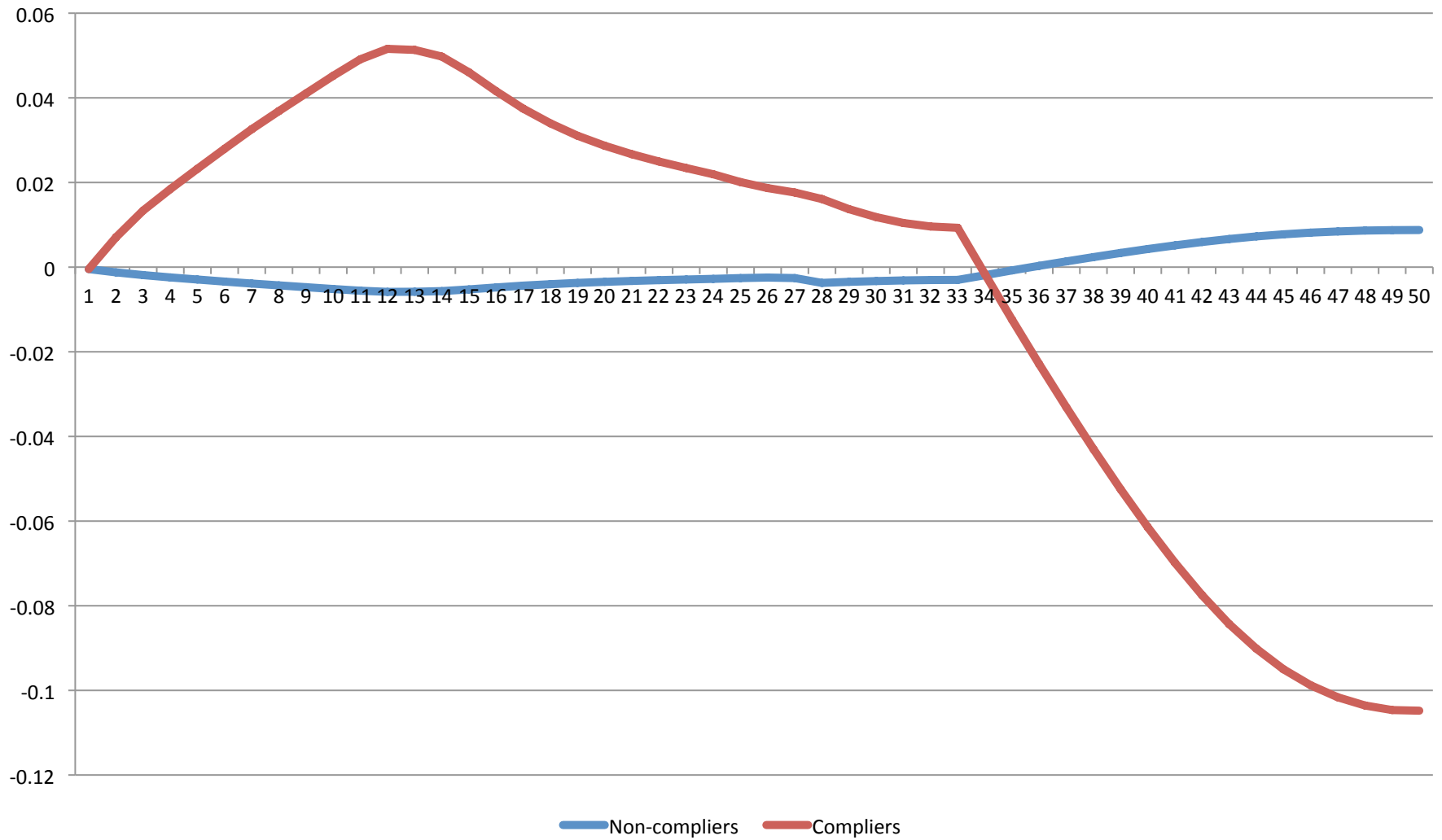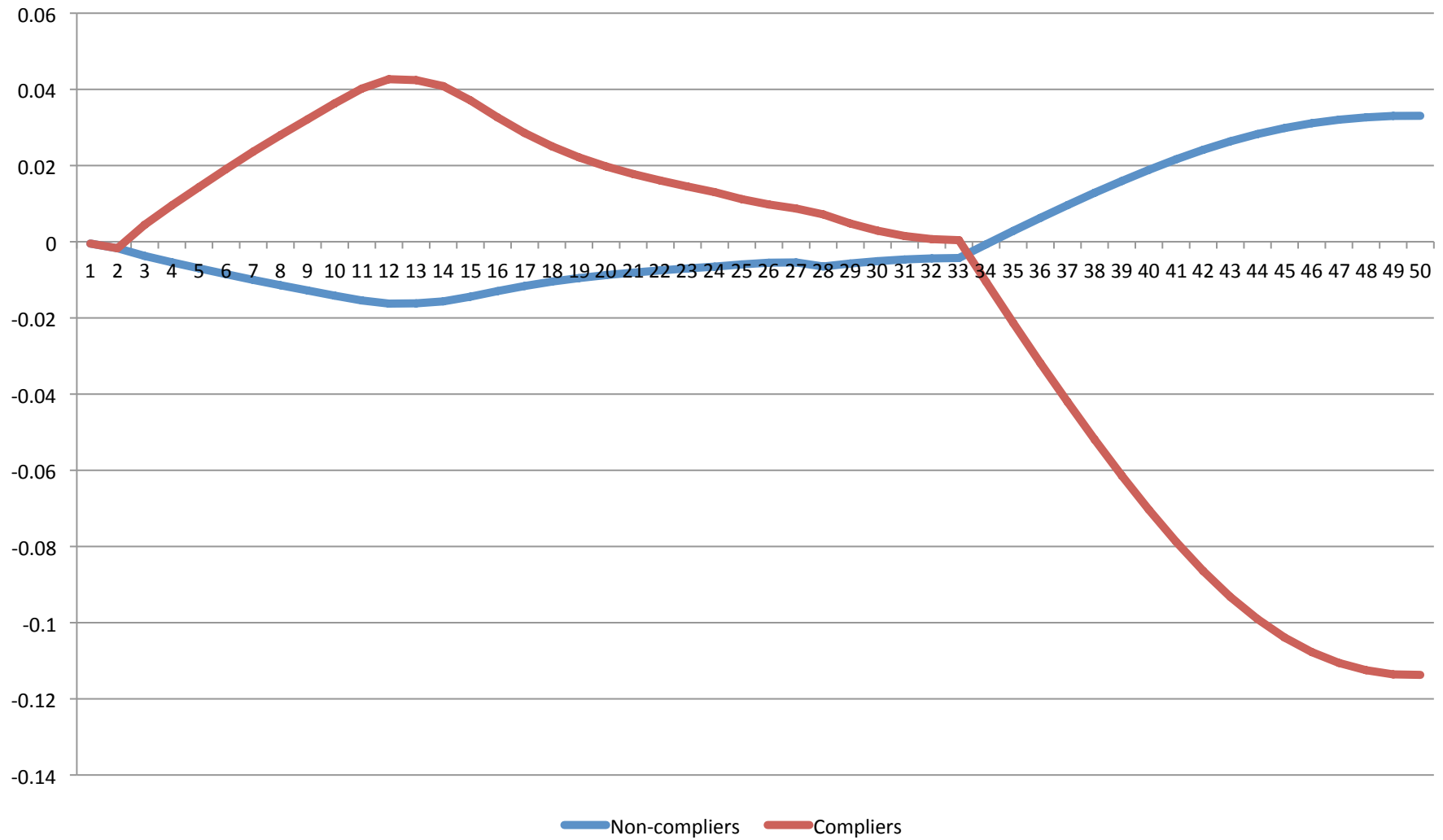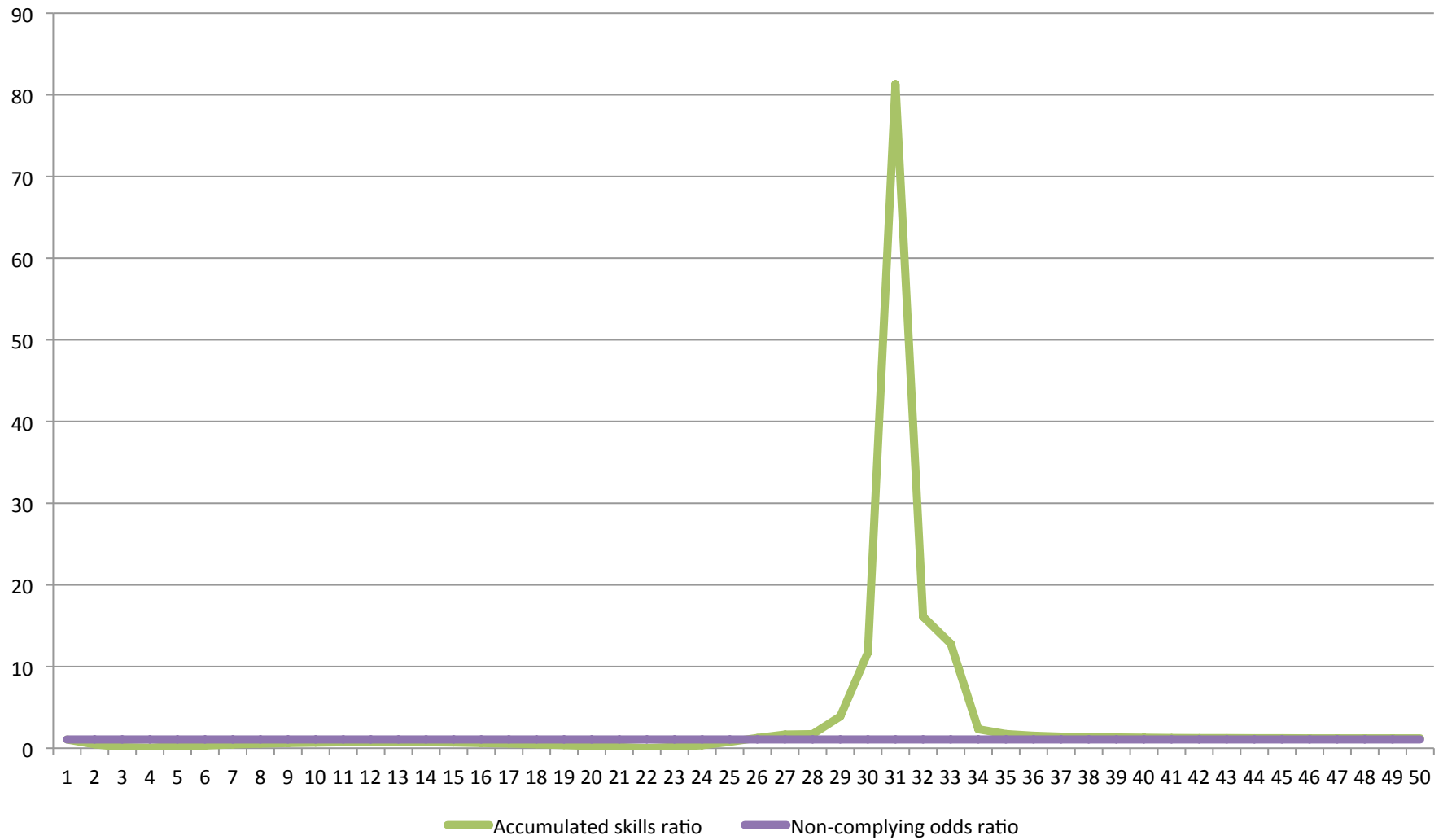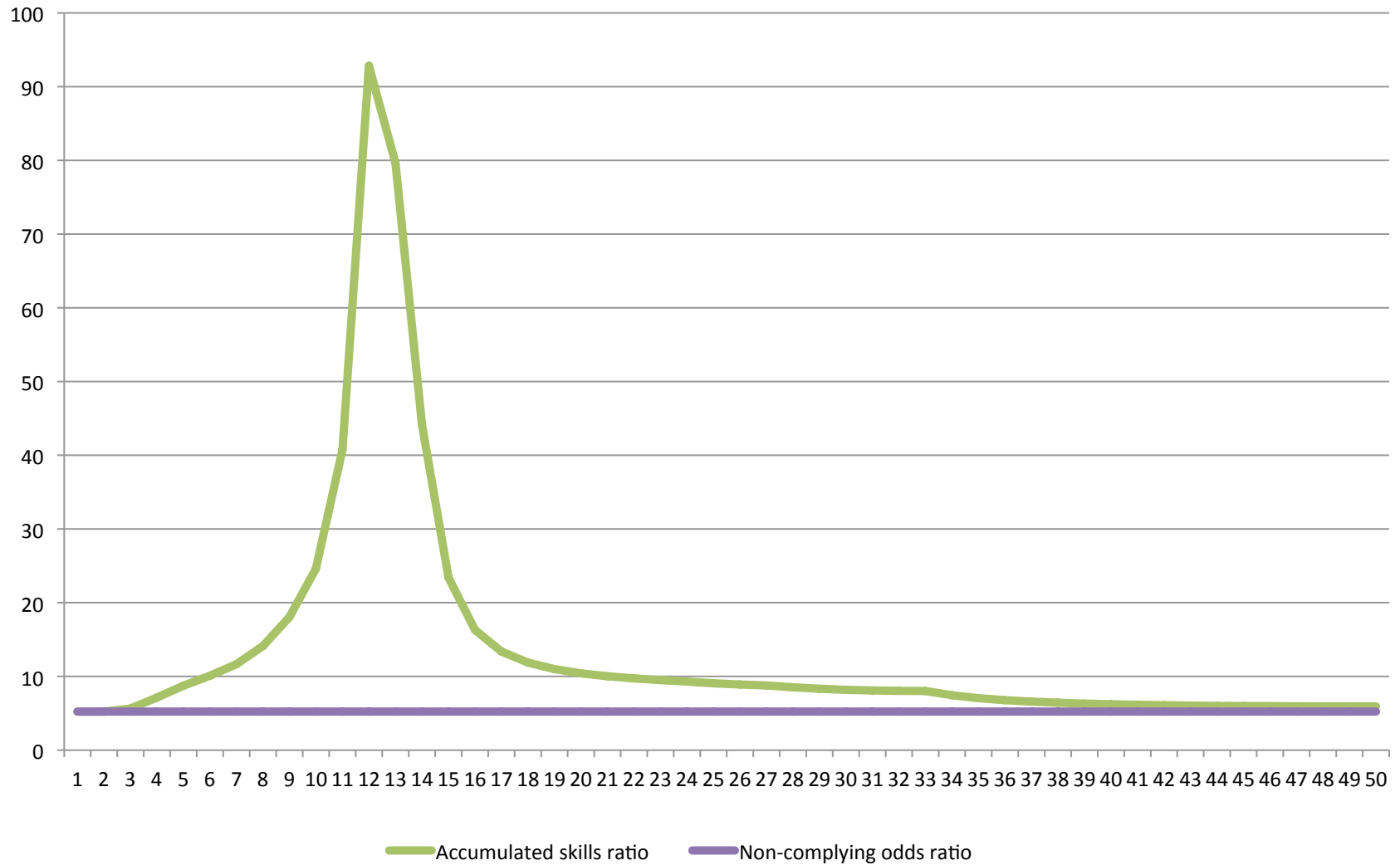
Figure 5. Life-cycle sequences of mean accumulated skills using a one-year mandatory school experiment.

Figure 6. Life-cycle sequences of mean accumulated skills using a two-year mandatory school experiment.

Figure 7. Life-cycle sequences of accumulated skills ratios using a level one subsidy experiment.

Figure 8. Life-cycle sequences of accumulated skills ratios using a level two subsidy experiment.

Figure 9. Life-cycle sequences of accumulated skills ratios using a level three subsidy experiment.
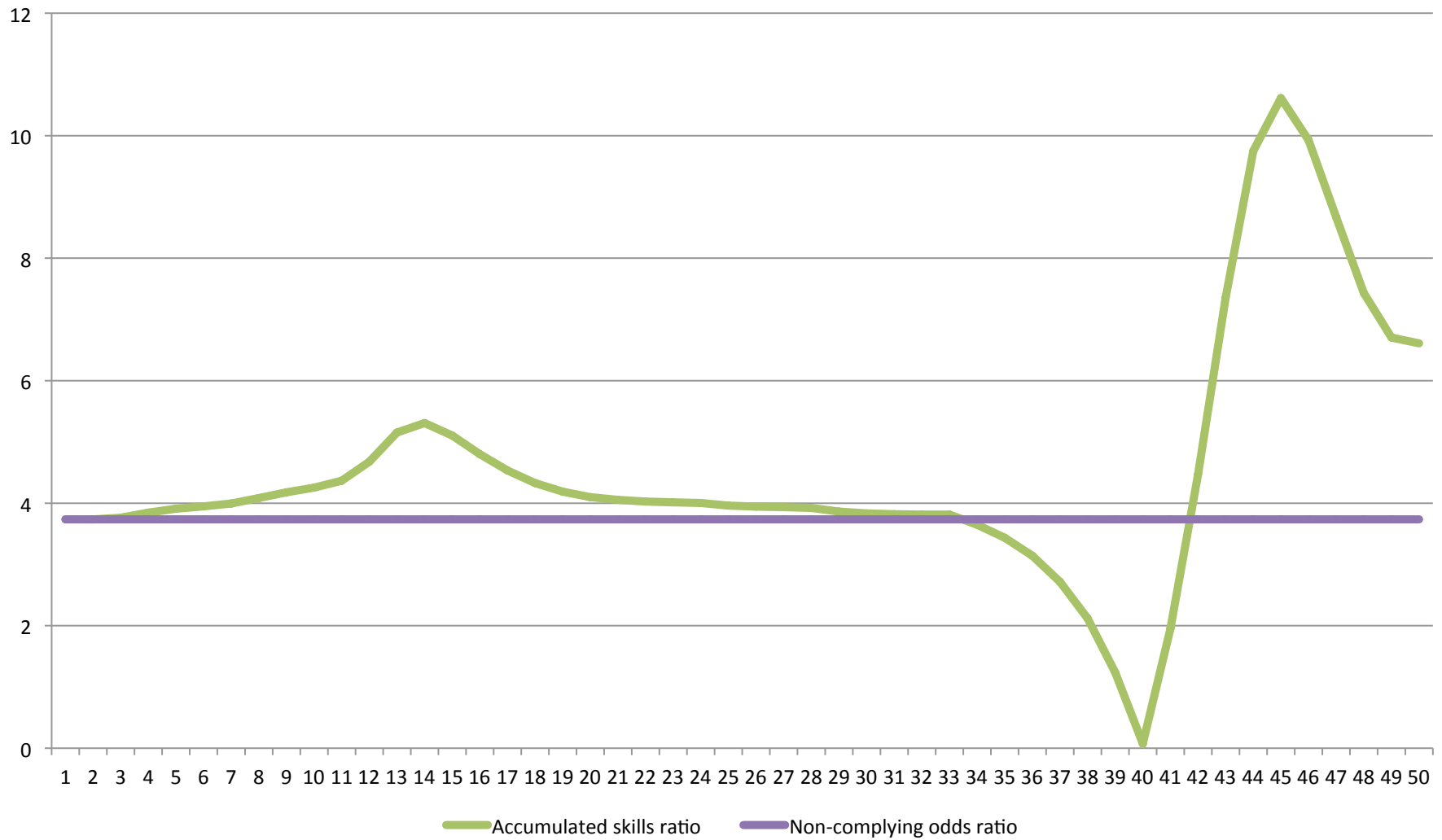
Figure 10. Life-cycle sequences of accumulated skills ratios using a level four subsidy experiment.

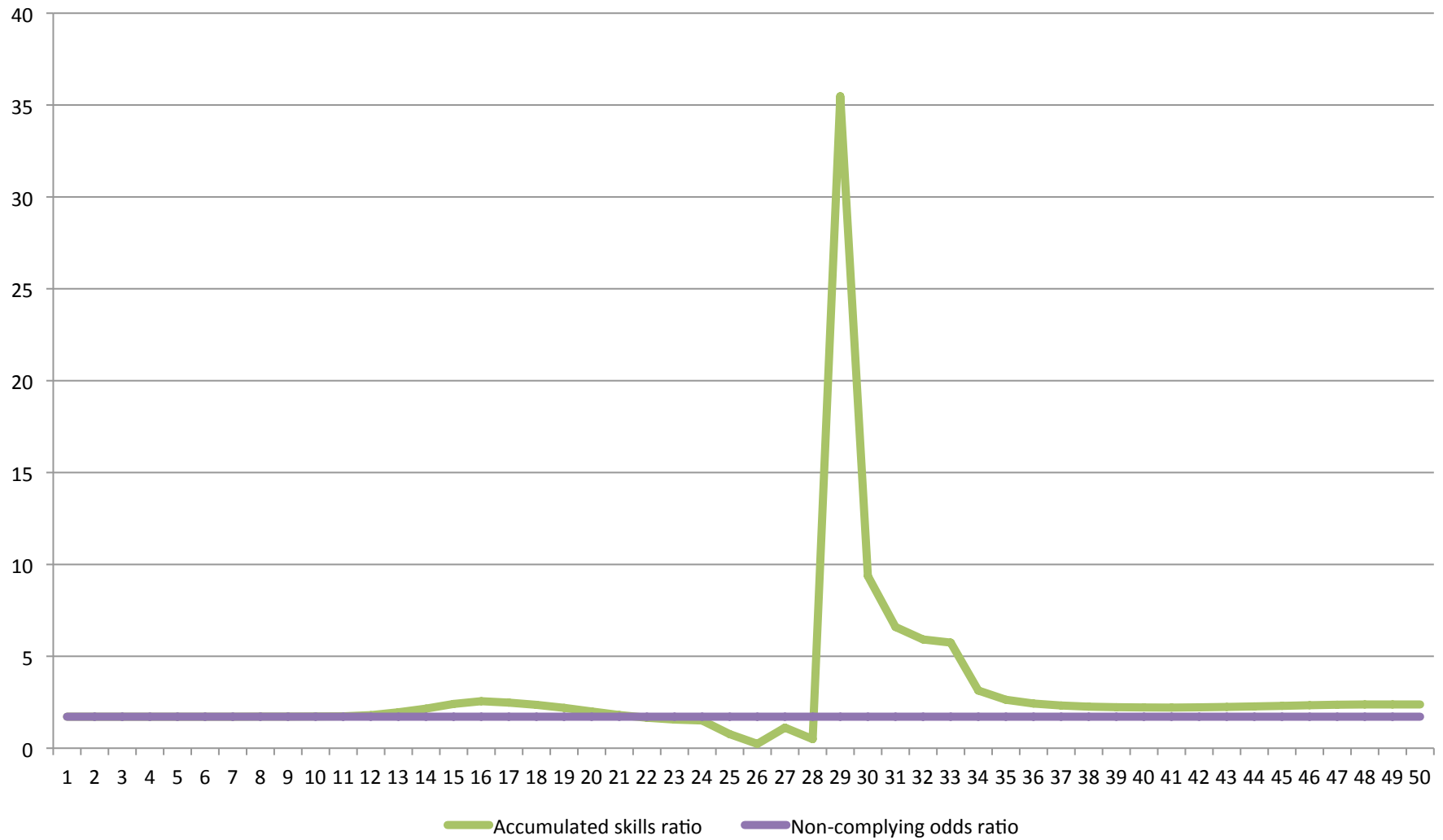Accumulated skills ratio   Non-complying odds ratio

Figure 11. Life-cycle sequences of accumulated skills ratios using a one-year mandatory school experiment.
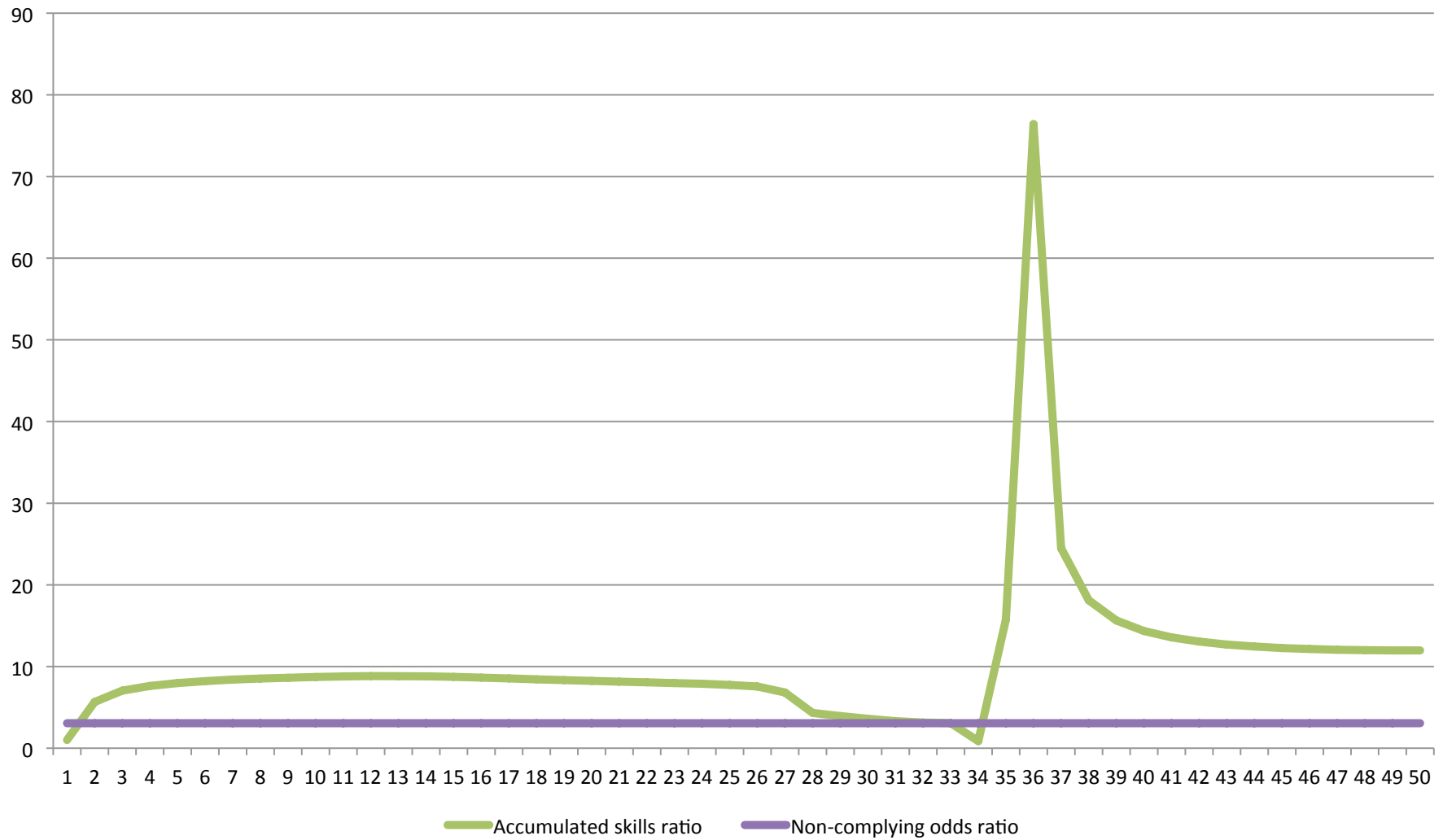
Figure 12. Life-cycle sequences of accumulated skills ratios using a two-year mandatory school experiment.