The Empirical Content of Season-of-Birth Effects: An Investigation with Turkish Data

Huzeyfe Torun
Semih Tumen

The Empirical Content of
Season-of-Birth Effects:
An Investigation with Turkish Data

Huzeyfe Torun
Central Bank of the Republic of Turkey

Semih Tumen
*Central Bank of the Republic of Turkey
and IZA*

ABSTRACT

The Empirical Content of Season-of-Birth Effects: An Investigation with Turkish Data[*]

Although the season of birth variable is often used as an instrumental variable to estimate the rate of returns to schooling in the labor economics literature, there is an emerging consensus that the season of birth is systematically associated with later outcomes in life such as the educational and labor market success; thus, it is highly likely non-random. Using a large micro-level data set from Turkey, we argue that the degree of this non-randomness can be even larger in a developing-country context. Specifically, we show that around 20 percent of all individuals in Turkey have January as their month of birth due to a combination of geographical, seasonal, institutional, and idiosyncratic factors that lead to misreporting. We further document that being January-born strongly predicts worse socio-economic outcomes in later life. We show that this can be a serious problem in evaluating policies that define eligibility based on the month of birth – such as compulsory schooling and compulsory military service laws that set the eligibility birth date cutoff as the January 1st. We confirm the validity of this concern based on a series of regression discontinuity design exercises. We conclude that, in a developing-country context, additional caution should be exercised when using the season-of-birth variable as a statistical tool.

Corresponding author:

Semih Tumen
Research and Monetary Policy Department
Central Bank of the Republic of Turkey
Istiklal Cad. No:10
06100 Ulus, Ankara
Turkey
E-mail: semih.tumen@tcmb.gov.tr

---

1   Introduction

The interaction between season of birth patterns and compulsory schooling laws are often used in the labor economics literature to construct an instrumental variable (IV) for the purpose of estimating the causal impact of education on earnings [see, e.g., Angrist and Krueger (1991)]. Several countries, including the United States, have compulsory schooling laws that keep young individuals enrolled in school until a certain age—typically until their 15th or 16th birthdays. Given this restriction, the season-of-birth differentials across individuals generate exogenous variation in the time spent in school for whom the compulsory school-leaving age is binding. This exogeneity is strictly conditional on the assumption that the season of birth is not a choice made by the parents or it is not systematically correlated with parental background.

However, there is no consensus in the literature on the validity of this exogeneity assumption. Papers including Angrist and Krueger (1992, 1995, 2001), Card (1999), and Hoogerheide, Kleibergen, and van Dijk (2007) argue that the date of birth is likely random and cannot be related to one's ability and/or parental background. There is a large body of empirical papers taking this argument as given and performing IV methods to estimate the returns to education.[1] Another set of papers—including Bound, Jaeger, and Baker (1995), Bound and Jaeger (2001), and Buckles and Hungerman (2013)—document that the exogeneity condition is typically not satisfied.[2]

Two explanations are offered in the literature to validate the latter argument. **First**, richer and/or more educated parents are more likely to have kids in the third or fourth quarters of the year rather than the first quarter. Buckles and Hungerman (2013) find that there is no seasonality in unwanted births—i.e., the births realized by teenagers and the unmarried. These women mostly have low socio-economic status. Voluntary births by women who have higher socio-economic status, on the other hand, are much less likely to be realized in winter. Therefore, seasonal patterns in fertility are mostly driven by the conception behavior of women

---

[1]See, e.g., Levin and Plug (1999), Plug (2001), Gelbach (2002), Hansen, Heckman, and Mullen (2004), Skirbekk, Kohler, and Prskawetz (2004), Lefgren and McIntyre (2006), Leigh and Ryan (2008), Maurin and Moschion (2009), Lee and Orazem (2010), and Robertsen (2011).

[2]Also see Cascio and Lewis (2006) and Dobkin and Ferreira (2010).

with higher socio-economic status. **Second**, winter births have worse outcomes due to in-utero exposure to weather and illness.[3] Moreover, if nutrition is also seasonal, this could also lead to differing fetal development patterns over different seasons [Barker (2001)]. Other studies documenting that the ones who are born early in the year are more likely to suffer from mental health problems than the ones who are born later in the year include Livingston, Adam, and Bracha (1993), Tochigi, Okazaki, Kato, and Sasaki (2004), and Doblhammer, Scholz, and Maier (2005). These arguments suggest that, although the idea is a great one, the patterns of variation in the school attainment across the quarters or months of year may not be solely driven by exogenous forces. Therefore, the causal nature of the returns to schooling estimates obtained with the season of birth IV is questionable.

In this paper, we contribute to this debate by providing evidence from Turkey using a large micro-level data set with rich content related to educational and labor market outcomes of individuals. We document that around 20 percent of the entire population in Turkey are "January-born" and, when we condition on individuals of age 60 and above, this fraction goes up to around 35 percent. We argue that this is a consequence of misreporting the birth dates. The tendency to misreport the birth date is a consequence of the combination of geographical, seasonal, and institutional factors. We list three important factors that leads to misreporting behavior as follows. **First**, especially in the past, the individuals residing in rural areas might have had difficulty in commuting to city centers in the winter season to officially report the birth of their children. These children are reported in the Spring and their birth date is usually registered as January 1st of the corresponding year. **Second**, teenage marriages are more common in some regions of Turkey and the children of teenage parents are not registered officially until the parents turn 18. These children are also generally registered as January born. **Finally**, some parents choose to report the children who are born in October, November, and December of a given year t as if they were born in January of year $t+1$. The main motivation is the perception that younger children in a given cohort are less successful in school than the older ones.

---

[3]See, e.g., Sham, O'Callaghan, Takei, Murray, Har, and Murray (1992), Gortmaker, Kagan, Caspi, and Silva (1997), and Almond (2006).

Since these reasons operate mostly in rural and less developed areas of the country, being January-born is associated with worse educational and labor market outcomes. In particular, we find that the January-born individuals receive 1 year less schooling than the children who are born in the other months of the year. Moreover, the January-born individuals earn around 5-7 percent less than the others. We conclude that the season of birth, especially being January born, predicts worse later outcomes in Turkey. Thus, we conclude that misreporting tendency is another reason why the season of birth is not orthogonal to later outcomes. We believe that this result is not specific to Turkey. In many other cultures, the birthday is registered to be January 1st if a person's real birthday is unknown for some reason and the reason is most likely associated with a low socio-economic status. Even in developed countries with heavy exposure to immigrant flows, the birth date statistics are plagued with the misreporting problem especially for the first-generation immigrants.

We also argue that our results are important for econometric policy evaluation practices. Certain important policy reforms in Turkey, such as the compulsory education and compulsory military service exemption reforms, have a cutoff birth date as January 1st.[4] Individuals born before and after this cutoff birth date are exposed to different regulations and environments. Most importantly, the official birth dates are binding for eligibility to these reforms. Since we argue that a January-born individual in Turkey is more likely to have a weaker socio-economic background, the January 1st cutoff can no more be considered as an exogenous policy reform cutoff date. In fact, the January 1st cutoff more likely separates the ones with weaker backgrounds from the ones with stronger backgrounds. Therefore, policy evaluation methods that rely on the January 1st cutoff, such as regression discontinuity designs and IV, are highly likely to produce biased results. To confirm the validity of this concern, we evaluate the effectiveness of 10 different placebo policy reforms with January 1st as the birth date cutoffs and we conclude that the ones who are born in January have worse educational and labor market outcomes than the ones born in December. In other words, any "policy" that separates January-born individuals from the December-born individuals will seem to yield results in favor of the December-born ones. Therefore, caution should be exercised in

---

[4]See, e.g., Aydemir and Kirdar (2013), Torun (2013), Cesur and Mocan (2013), and Torun and Tumen (2015).

performing these methods to evaluate the policies with January 1st as the cutoff birth date.

The plan of the paper is as follows. Section 2 describes the data and presents summary statistics. Section 3 documents the season of birth patterns in Turkey, discusses in detail the potential drivers of the season of birth patterns, and shows that these patterns may distort policy evaluation exercises performed using regression discontinuity designs and IV. Section 4 concludes.

2 Data

We use the Turkish Household Labor Force Survey (LFS) micro-level data sets for the period 2004–2013. LFS micro-level data set is compiled and published by the Turkish Statistical Institute (TURKSTAT). It is a cross-sectional, publicly-available, and nationally-representative data set. The official employment statistics in Turkey rely on the information extracted from the LFS. The LFS provides detailed information on current and past labor market outcomes accompanied by a rich set of individual characteristics of individuals including gender, age, education, marital status, urban/rural status, region of residence as well as detailed job characteristics including information on industry, occupation, part/full-time job status, public/private job status, permanent/temporary job status, and formal/informal employment status. The sample covers all private households from all settlements in Turkey. Institutional population such as the residents of schools, dormitories, kindergartens, rest homes for elderly persons, special hospitals, military barracks, and recreation quarters for officers are excluded. A two-stage stratified cluster sampling method is used. Using an address-based rotation pattern, a 50 percent overlap between two consecutive periods and in the same periods of the two consecutive years is targeted.

In addition to the standard publicly-available information in the LFS, we also use the month of birth and year of birth information that we obtain from TURKSTAT.[5] Both the month of birth and year of birth information in the survey comes from interviewees' own responses

---

[5]Note that this information is not publicly available. We thank the staff of the Labor Force Statistics Group in TURKSTAT for their help.

in a setting based on face-to-face interview. There is no cross check from administrative and other official records to test the correctness of the responses. Table (1) documents the basic summary statistics for our sample. The total number of observations in our sample is 4,624,686. In the whole sample, the fraction of January-born individuals is around 0.20. The January-born individuals are evenly distributed across gender categories, while their fraction is higher among rural, older, and less-educated individuals. We also observe a rather smaller clustering for July. In Section 3, we provide a detailed explanation of the structural reasons why the density of people that are born in January and July is higher than that born in the other months.

3 Empirical Analysis and Discussion

In this section, we provide an empirical analysis of the season-of-birth patterns in Turkey. We proceed in four steps. *First*, we find that the birth month distribution is strikingly asymmetric in Turkey. In particular, we report that, for a significant fraction of individuals in Turkey, the month of birth has been reported as January. We also show that this tendency is more pronounced for older cohorts. *Second*, we discuss the potential drivers of this asymmetry. *Third*, we investigate whether having January as the month of birth is correlated with (i) educational outcomes, (ii) labor market outcomes, and (iii) the parental background of the individuals. *Finally*, we present convincing evidence that the month-of-birth patterns in Turkey significantly reduce the reliability of certain quasi-experimental techniques—in particular, the regression discontinuity design—used to evaluate public policy.

3.1 Season-of-Birth Patterns in Turkey

We start our analysis by providing a descriptive perspective for the season of birth patterns in Turkey. We use the Turkish Household Labor Force Survey micro-level data sets for the period 2004–2013.[6] If the month of birth is totally random and there is no seasonality, then there should be an even distribution of individuals across the months, i.e., the month-of-birth probability should be $1/12 \approx 0.083$, for each month, for all individuals. But, we provide

---

[6]See Section 2 for a detailed description of this survey.

strong evidence that there is no such symmetry in the month-of-birth patterns in Turkey. In fact, there is a considerable degree of seasonality in fertility behavior. Specifically, we show that there is a strong tendency for having January as the birth month. This tendency gets stronger with age, which implies that the incidence of being born in January is stronger for older cohorts.

Next we attempt to quantify the extent of this asymmetry. Figures (1) – (2) document the birth month distribution in terms of age groups. Upper panel in Figure (1) plots the month-of-birth distribution for the youngest age group, 0–19. It says that around 14 percent of the individuals in this age interval are born in January—almost twice as the symmetric probability. For older cohorts, the degree of this asymmetry is even larger. Lower panel in Figure (1) and upper panel in Figure (2) plot the same distribution for age groups 20–39 and 40–59 respectively. The fraction of January-born individuals are 17 and 26 percent for age groups 20–39 and 40–59, respectively. Finally, lower panel in Figure (2) reports the month-of-birth patterns for the oldest age group, i.e., 60 and above. The probability of being January-born is around 35 percent for older individuals. These figures highlight four interesting observations: (1) the probability of being January-born is much larger than the probability of getting born in the other months, (2) the extent of this asymmetry is much larger for older cohorts than the younger cohorts, (3) July is another month we see clustering (much smaller than the January effect though), and (4) it is almost invariably observed across age groups that the birth probabilities decline steadily over the year. Next we discuss the potential factors that can explain these season-of-birth patterns and the associated asymmetries.

3.2 Potential Drivers of the Season-of-Birth Patterns

The underlying force driving these asymmetries is the tendency to report the birth date as the January 1st. There are several factors that lead to misreporting. The **first** is related to the institutional setup, which is a combination of the regulations, geographical factors, and seasonal factors. The fraction of rural population is typically high in Turkey and it was even higher in the past.[7] Until late 1980s, the transportation services, health services, and the

---

[7]The fraction of rural population is around 0.20–0.25 based on the most recent figures. It was approximately 0.50 in 1960s.

related infrastructure services were not highly developed in the rural regions of the country. The connection between the city centers and the villages were mostly broken during winter due to weather conditions.[8] Anecdotal evidence suggests that, in some regions, the duration of disconnection was around 6–7 months—i.e., from November to April. Children born in these regions during winter season were altogether reported to the Civil Registration Unit by the village headmen after the winter season. The convention was to report the birthday of this children as the January 1st of the corresponding year. The birth registration system were allowing for such flexibility in the past, although the current rules are strict and require official documentation issued by a hospital or health agency. As a result, the tendency to report January 1st as the birth date is due to a combination of institutional, geographical, and seasonal factors. The exact birth dates of the individuals in these category are mostly unknown.

The **second** mechanism is related to teenage marriages, which are currently rare in Turkey bu was quite common in the past. Teenage marriages are mostly performed through religious ceremonies, since the laws do not permit legal marriages below age 18.[9] Although teenage marriages are unlawful, they are mostly socially approved since, in general, a religious ceremony is publicly held. Children born out of these marriages are not officially registered until their parents turn age 18. After age 18, parents get married legally and the children are officially registered. Their birth dates are recorded as January 1st of the year after the parents get legally married. The two factors listed above account for a big chunk of the "January effect" in birth records. Based on a fact documented by the Anadolu Agency (the government-operated news agency in Turkey) in January 2012, around 20 percent of the individuals reside in the Siverek province of Urfa (a large city in the Southeastern Turkey) have January 1st as their birthdays—40,000 individuals out of a total of 200,000.

Third, another type of selection might also be operating in the background, which is not necessarily driven by the mechanisms explained above. This factor is related to parents'

---

[8]This was a particular concern for the Eastern and Northern parts of the country, which have been surrounded by high mountains.

[9]More precisely, individuals can legally get married only if they have completed age 17.

perception about the school starting age of children. Specifically, some parents believe that children who start school early attain worse educational outcomes relative to the ones who start late. There are also several empirical papers in the literature supporting the validity of this belief.[10] Based on these concerns, parents of the children who are born later in a given year (i.e., in October, November, or December) may choose to report the birth dates of their children as the January of the following year. As a consequence of this choice, the children will become one of the oldest students in his class, while s/he would be one of the youngest students had the parents reported correctly. There is some anecdotal evidence that this might also be a relevant mechanism. But, it is definitely a minor factor relative to the other two factors explained above. On a similar note, some parents may also tend to misreport to delay the compulsory military service for their male children.

These three factors explain the "January effect" documented by Figures (1) – (2). Clearly, the tendency for misreporting has declined over time. There are three main reasons for this decline. First, the health system, transportation services, and the related infrastructure have significantly improved over time. Second, several reforms have been performed to improve the accuracy of the Civil Registration system including a new census system as well as a fully computerized population monitoring. And, third, enforcement have become more strict in the form of increased fines and punishment for misreporting the birth date and/or reporting the birth late. Although the misreporting tendency have significantly declined over time, the data suggest that it is still widespread for the new generation [see the upper panel of Figure (1)].

There is also a smaller, but non-negligible, "July effect" in the data. For all age groups, the probability of being July-born is significantly larger than the neutral probability of 0.083. To be precise, the fraction of the individuals who're born in July is approximately 0.11, 0.10, 0.11, and 0.17 for age groups 0–19, 20–39, 40–59, and above 60, respectively. Again, the likelihood of being July-born is larger for older cohorts. Unlike the January effect, which was an artifact of misreporting, the July effect is more likely due to data imputation. There is a certain fraction

---

[10]See, e.g., Bedard and Dhuey (2006), McEwan and Shapiro (2008), and Elder and Lubotsky (2009). We should note, however, that there is no consensus in the related literature about the impact of school starting age on educational attainment and that there are serious identification problems mostly due to the difficulty of separating age, maturity, and education effects.

of individuals (mostly old individuals), for whom the month of birth is missing and only the year of birth is known. For some waves of the survey, the month of birth is entered as "July" by the interviewers for those whose month of birth is missing. This is mainly performed for the purpose of approximating the age of the individual, which is a key variable in the data. This imputation is the main force driving the July effect.

The factors listed above are consistent with the observed patterns in the season of birth. These factors also suggest that certain type of individuals tend to misreport the birth dates of their children. These are most likely the poor, uneducated individuals living in rural areas. In the rest of the paper, we document that the season of birth variable is systematically correlated with educational outcomes of both the parents and children. Lastly, we demonstrate that these correlations can severely distort empirical policy evaluation exercises.

3.3   Education and Labor Market Outcomes

In this section, we investigate the correlations between an individual's season of birth and the outcomes s/he faces in school and the labor market. The outcomes that we focus on include years of schooling, employment, and wages. It should be noted that the mechanic forces related to misreporting and imputation are most likely the main reasons underlying the differences in outcomes. However, the outcomes may also differ with respect to season of birth due to selective seasonality in actual fertility behavior—see, e.g., Buckles and Hungerman (2013). Unfortunately, the data do not allow us to distinguish between the misreporting effect from the effect coming from actual fertility behavior. However, we believe that the findings we document below are still quite valuable, because they suggest that season of birth records are far from being random even in the absence of seasonality in actual fertility behavior.

3.3.1   Labor Market Outcomes

Figure (3) summarizes employment outcomes with respect to month of birth. We restrict our sample to the individuals above age 30 for the purpose of concentrating on the ones who have completed their schooling. We focus on two outcomes: employment-to-population

ratio and the labor force participation ratio. The figure suggests that both the employment to population ratio and the labor force participation ratio are approximately 5–6 percentage points lower in January than the average of the other months. In other words, the ones who have misreported birth dates are the ones who are less likely to be employed and who have low labor market attachment levels. Note that the July effect on employment outcomes is also significant—slightly less prominent than the January effect.

Figure (4) documents the differences in real wages with respect to the month of birth. The nominal wages are deflated with the CPI, taking 2004 as the base year, to obtain real wages. A similar story is observed for real wages; that is, the real wages for those workers whose month of birth is reported as January earn around 10–12 percent lower wages than the average earnings for those who were born in other months. While a July effect is also observed, it is not as strong as the January effect. One important observation regarding the month of birth effects on labor market outcomes is that the labor market outcomes for those who were born later in the year are better than the ones who were born early. This pattern is not directly related to the January effect. In fact, there might also be seasonality in fertility behavior for the ones with better socio-economic background—similar to what Buckles and Hungerman (2013) document. The data set we analyze, however, does not allow us to distinguish the misreporting effect from seasonality in the fertility behavior.

We also perform regressions to check whether controlling for personal characteristics—in particular, the education level—can explain the month of birth differences in labor market outcomes. Table (2) reports the regression results. The first two columns regresses a dummy variable taking 1 if the individual is employed and 0 otherwise on a set of observables including gender, age, marital status, urban/rural indicator, region dummies, and survey-year dummies. Education controls are included in column [2], but excluded in column [1]. In the last two columns, the dependent variable is a dummy variable taking 1 if the individual participates the labor market and 0 otherwise. Again, education controls are included in column [4], but excluded in column [3]. In all regressions, we find that controlling for observed characteristics explains the January effect. We even document that the employment and labor

11

force participation probabilities are slightly higher for those who were born in January. This is not surprising, because the January-born individuals are more likely to be old, reside in rural areas, be employed as an unpaid family worker, and be located in the Eastern part of the country. The employment and labor force participation tendencies are already low for this type of individuals regardless of their month of birth. So, the employment and labor force participation gaps disappear once we control for observables.

Table (3) reports the regressions results for real wages and formal/informal employment. The dependent variable is the real wage in the first two columns and informal employment dummy variable taking 1 if the worker is employed formally and 0 if s/he is employed informally in the last two columns. Columns [1] and [3] exclude the education controls, while [2] and [4] include. The control variables include gender, age, marital status, an urban/rural indicator, part/full-time job status, industry dummies, region dummies, and survey-year dummies. Unlike the results documented in Table (2), we find that part of the January effect is unexplained for real wages and formal/informal employment. For real wages, the unexplained part is approximately 5 percentage points before controlling for education and it goes down to around 2.5 percentage points after including the education controls. So, January-born individuals are less likely to earn high wages even after for controlling for the observed determinants of wages. Same is true for formal employment. On average, January-born individuals are more likely to be employed in informal jobs. The results presented in Tables (2) and (3) complement each other. In terms of rough employment and labor force participation outcomes, there are no important differences between January-born individuals and others.[11] However, the composition of employment differs substantially. The ones who are born in January are more likely to be employed in "bad" jobs, i.e., low-pay and/or informal jobs.

3.3.2 Educational Outcomes

Figures (5) – (7) reveal that there are huge differences between month of births with respect to educational outcomes. Figure (5) plots the high school and college graduation rates. Clearly, there is a huge January effect for both outcomes. To be specific, the high school and college

---

[11]We also observe that the July-effect becomes smaller after controlling for the observed covariates.

graduation rates for January-born adults are 0.15 and 0.6, respectively, while these rates are 0.31 and 0.13 for an average October-born adult, respectively. We also observe that there is also a significant July-effect in educational outcomes; that is, the high school and college graduation rates for an average July-born individual are 0.16 and 0.07, respectively. Figures (6) and (7) say that these strong month of birth effects in education hold invariably across different birth cohorts and also across gender categories.

We also check whether controlling for observed characteristics explains these gaps. Table (4) reports the regression results. We focus on three outcomes: years of completed schooling (column [1]), high school graduation (column [2]), and college graduation (column [3]). The control variables include gender, age, marital status, an urban/rural indicator, region dummies, and survey-year dummies. We find that the observed variables can explain only a small fraction of the raw January effect. To be concrete, we find that, on average, the January-born individuals stop schooling approximately 1 year early and they are around 10 percent less likely complete high school education or above.

3.4 Parental Background

Next we focus on the relationship between one's month of birth and parental background. For this purpose, we focus on children of age 15 and younger. Figures (8) and (9) summarize the differences in maternal and paternal education, respectively, in terms of month of birth. Although the incidence of misreporting appears to diminish for younger individuals, we still observe a systematic relationship between being January-born and having parents with worse educational backgrounds. We see that this tendency holds for both maternal and paternal education. Specifically, we find that the mothers of January-born children have, on average, 0.3 years less completed education than the mothers of children born in the other months. The magnitude for the paternal education gap is slightly higher, around 0.35 years, on average.

3.5 Implications for Policy Evaluation

Quasi-experimental designs have become increasingly popular tools used by economists in evaluating the effectiveness of public policies. An unexpected exogenous change in the economic environment is the main requirement of a quasi-experimental design. Most of the time, such an unexpected change comes from a policy reform. For some reforms, there is a sharp cutoff date describing eligibility. There are several examples for such policy reforms in Turkey. For example, the 1997 education reform extended the compulsory schooling from 8 years to 12 years. The eligibility is defined in terms of the date of birth. Specifically, the ones who were born on or after 1987 were eligible, while those who were born before 1987 were not.[12] Another example is the 1999 compulsory military service (paid) exemption reform. The ones who were born on or before December 31st, 1972 were eligible for the paid exemption, while those who were born on or after January 1st, 1973 were not eligible.[13] In an ideal setting, a regression discontinuity design (RDD) will identify the impact of the reform. The prerequisite for having an "ideal setting" in an RDD is that the reform should resemble a randomized experiment in such a way that the cutoff date randomly separates the ones who receive the treatment from the ones who do not.

To be more concrete, suppose that we have a policy reform at hand that separates the treatment group from the control group based on the birth dates of the subjects. Suppose further that, just as the compulsory military service exemption reform mentioned above, the ones who were born on or before December 31st, 1972 are in the treatment group, while the those who were born on or after January 1st, 1973 are placed into the control group. The main problem here is that, although the policy reform constitute an unexpected and exogenous change in the economic environment, the distribution of individuals around the reform date prevents the formation of an artificial randomization. The evidence we provide above suggests that the January-born individuals are systematically different from the December-born individuals in the sense that, on average, they have worse educational and labor market outcomes; moreover,

---

[12]See Aydemir and Kirdar (2013) and Torun (2013) for the details of the reform.

[13]See Torun and Tumen (2015) for the causal impact of this reform on educational and labor market outcomes of the eligible males.

their parents have worse socio-economic characteristics. In such a setup, whether there is an actual treatment or not, we expect to see that the outcomes of the January-born individuals should be worse than the outcomes of the December-born individuals for any birth year.

To confirm the validity of this conjecture, we perform RDD exercises for the purpose of comparing the real earnings of individuals born right before and after January 1st, just as if January 1st were the reform cutoff date. We execute this exercise for 10 different birth cohorts, from 1965 to 1975. We set a two-month window, from December in year t to January in year $t+1$. The observations for December in year t represent the treated outcomes, while the observations for January in year $t+1$ represent the control outcomes. Note that, unlike the US, the compulsory schooling laws in Turkey do not have age cutoffs—i.e., the children in the US are required to enrolled in school until their 15th or 16th birthdays (depending on the state of residence), while, in Turkey, the laws require children to get the degree. This means that there is no reason to expect variation in outcomes between December and January due to compulsory schooling laws.

Table (5) documents the results of the 10 different RDD exercises for two different outcomes. Each of these exercises is performed based on the placebo assumption that there is a policy reform in the corresponding year putting the December-born individuals in year t into the treatment group and the January-born individuals in year $t+1$ into the control group. The outcomes of interest are the monthly labor market earnings (including the performance pay, bonuses, etc) and the years of completed education. The covariates include gender, age dummies, urban/rural dummy, region dummies, and survey-year dummies. The education dummies are included into the earnings regressions. The first two columns report the results of the earnings regressions. Column [1] excludes the education dummies, while column [2] includes them. Column [1] suggests that for almost all birth-year cohorts, the treatment outcome is positive and statistically significant. In other words, for each exercise, we find a positive effect as if there were an actual policy reform. The magnitude of the effect on earnings is around 10 percentage points. Based on the discussion presented above, we conclude that this result should be fully attributed to the January effect. Column [2] suggests that the effect

15

almost disappears when we control for education. As a complementary exercise, Column [3] suggests that the January effect on earnings is fully driven by the January effect on education. This result is consistent with the results presented in Sections 3.2 and 3.3.

4   Concluding Remarks

The season of birth is often used in labor economics as an instrumental variable (combined with compulsory schooling laws) that generates exogenous variation in years of completed education. However, there is also a large set of papers showing that there is a systematic relationship between the season of birth and later outcomes—such as labor market outcomes, education, health, etc. The explanations for this association usually focus on factors that determine seasonality in fertility behavior or seasonal exposure to nutrition and disease. In this paper, we use a large micro-level data set from Turkey to show that a combination of geographical, seasonal, and institutional factors can also lead to a strong systematic association between season of birth and later outcomes. In particular, we show that there is a strong tendency to be registered as "January-born" in Turkey and this tendency is higher for people with worse socio-economic backgrounds. As a result, there is a sharp difference between the later outcomes of individuals who are born in December of year t and those who were born in January of year $t + 1$. We also argue that these differences also lead to problems that contaminate policy evaluation exercises that rely on sharp discontinuities (such as regression discontinuity designs and IV) determined by cutoff birth dates.

Although we use data from Turkey to document the extent of the "January effect," the results we document are likely not specific to Turkey. For example, it is well known that the January 1st is a quite common birth day among immigrants in the United States. In many other cultures, the birthday is considered to be January 1st if a person's real birthday is unknown for some reason. Overall, this paper documents a new explanation for the systematic association between the season of birth of individuals and their later outcomes. Our analysis suggests that empirical work using season of birth as a key econometric variable should also account for the geographical, seasonal, and institutional forces that might lead to misreporting of the

birth dates.

References

ALMOND, D. (2006): "Is the 1918 Influenza Pandemic is Over? Long-term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population," *Journal of Political Economy*, 114, 672–712.

ANGRIST, J. D. AND A. B. KRUEGER (1991): "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106, 976–1014.

——— (1992): "The Effect of Age at School Entry on Educational Attainment: An Application on Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 87, 328–336.

——— (1995): "Split Sample Instrumental Variables Estimates for the Return to Schooling," *Journal of Business and Economic Statistics*, 13, 225–235.

——— (2001): "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 15, 69–85.

AYDEMIR, A. AND M. G. KIRDAR (2013): "Estimates of the Return to Schooling in a Developing Country: Evidence from a Major Policy Reform in Turkey," Unpublished manuscript, Middle East Technical University.

BARKER, D. (2001): "Fetal and Infant Origins of Adult Disease," *Monatsschrift fur Kinderheilkunde*, 149, 2–6.

BEDARD, K. AND E. DHUEY (2006): "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects," *Quarterly Journal of Economics*, 121, 1437–1472.

BOUND, J. AND D. A. JAEGER (2001): "Do Compulsory School Attendance Laws Alone Explain the Association between Quarter of Birth and Earnings?" *Research in Labor Economics*, 19, 83–108.

BOUND, J., D. A. JAEGER, AND R. M. BAKER (1995): "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous

Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, 443–450.

BUCKLES, K. S. AND D. M. HUNGERMAN (2013): "Season of Birth and Later Outcomes: Old Questions, New Answers," *Review of Economics and Statistics*, 95, 711–724.

CARD, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, New York, NY: Elsevier, vol. 5, 1801–1863.

CASCIO, E. AND E. LEWIS (2006): "Schooling and the Armed Forces Qualifying Test," *Journal of Human Resources*, 41, 294–318.

CESUR, R. AND N. MOCAN (2013): "Does Secular Education Impact Religiosity, Electoral Participation and the Propensity to Vote for Islamic Parties? Evidence from an Education Reform in a Muslim Country," NBER WP# 19769.

DOBKIN, C. AND F. FERREIRA (2010): "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?" *Economics of Education Review*, 29, 40–54.

DOBLHAMMER, G., R. SCHOLZ, AND H. MAIER (2005): "Month of Birth and Survival to Age 105+: Evidence from the Age Validation Study of German Semi-supercentenarians," *Experimental Gerontology*, 40, 829–835.

ELDER, T. E. AND D. H. LUBOTSKY (2009): "Kindergarten Entrance Age and Children's Achievement Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources*, 44, 641–683.

GELBACH, J. (2002): "Public Schooling for Young Children and Maternal Labor Supply," *American Economic Review*, 92, 307–322.

GORTMAKER, S., J. KAGAN, A. CASPI, AND P. SILVA (1997): "Daylength during Pregnancy and Shyness of Children: Results from Northern and Southern Hemisphere," *Developmental Psychobiology*, 31, 107–114.

HANSEN, K. T., J. J. HECKMAN, AND K. MULLEN (2004): "The Effect of Schooling and Ability on Achievement Test Scores," *Journal of Econometrics*, 121, 39–98.

HOOGERHEIDE, L., F. KLEIBERGEN, AND H. VAN DIJK (2007): "Natural Conjugate Priors for the Instrumental Variables Regression Model Applied to the Angrist-Krueger Data," *Journal of Econometrics*, 138, 63–103.

LEE, C. AND P. ORAZEM (2010): "High School Employment, School Performance, and College Entry," *Economics of Education Review*, 29, 29–39.

LEFGREN, L. AND F. MCINTYRE (2006): "The Relationship between Women's Education and Marriage Outcomes," *Journal of Labor Economics*, 24, 787–830.

LEIGH, A. AND C. RYAN (2008): "Estimating Returns to Education Using Different Natural Experiment Techniques," *Economics of Education Review*, 27, 149–160.

LEVIN, J. AND E. PLUG (1999): "Instrument Education and the Returns to Schooling in the Netherlands," *Labour Economics*, 6, 521–534.

LIVINGSTON, R., B. ADAM, AND H. S. BRACHA (1993): "Season of Birth and Neurodevelopmental Disorders: Summer Birth is Associated with Dyslexia," *American Academy of Child and Adolescent Psychiatry*, 32, 612–616.

MAURIN, E. AND J. MOSCHION (2009): "The Social Multiplier and Labor Market Participation of Mothers," *American Economic Journal: Applied Economics*, 1, 251–272.

MCEWAN, P. J. AND J. S. SHAPIRO (2008): "The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates," *Journal of Human Resources*, 43, 1–29.

PLUG, E. (2001): "Season of Birth, Schooling and Earnings," *Journal of Economic Psychology*, 22, 641–660.

ROBERTSEN, E. (2011): "The Effects of Quarter of Birth on Academic Outcomes at the Elementary School Level," *Economics of Education Review*, 30, 300–311.

SHAM, P., E. O'CALLAGHAN, N. TAKEI, G. MURRAY, E. HAR, AND R. MURRAY (1992): "Schizophrenia Following Pre-natal Exposure to Influenza Epidemics between 1939 and 1960," *British Journal of Psychiatry*, 160, 461–466.

SKIRBEKK, V., H.-P. KOHLER, AND A. PRSKAWETZ (2004): "Birth Month, School Graduation, and the Timing of Births and Marriages," *Demography*, 41, 547–568.

TOCHIGI, M., Y. OKAZAKI, N. KATO, AND T. SASAKI (2004): "What Causes Seasonality of Birth in Schizophrenia?" *Neuroscience Research*, 48, 1–11.

TORUN, H. (2013): "Compulsory Schooling Laws and Early Labor Market Outcomes," Unpublished manuscript, Central Bank of the Republic of Turkey.

TORUN, H. AND S. TUMEN (2015): "The Impact of Compulsory Military Service Exemption on Education and Labor Market Outcomes: Evidence from a Natural Experiment," Unpublished manuscript, Central Bank of the Republic of Turkey.
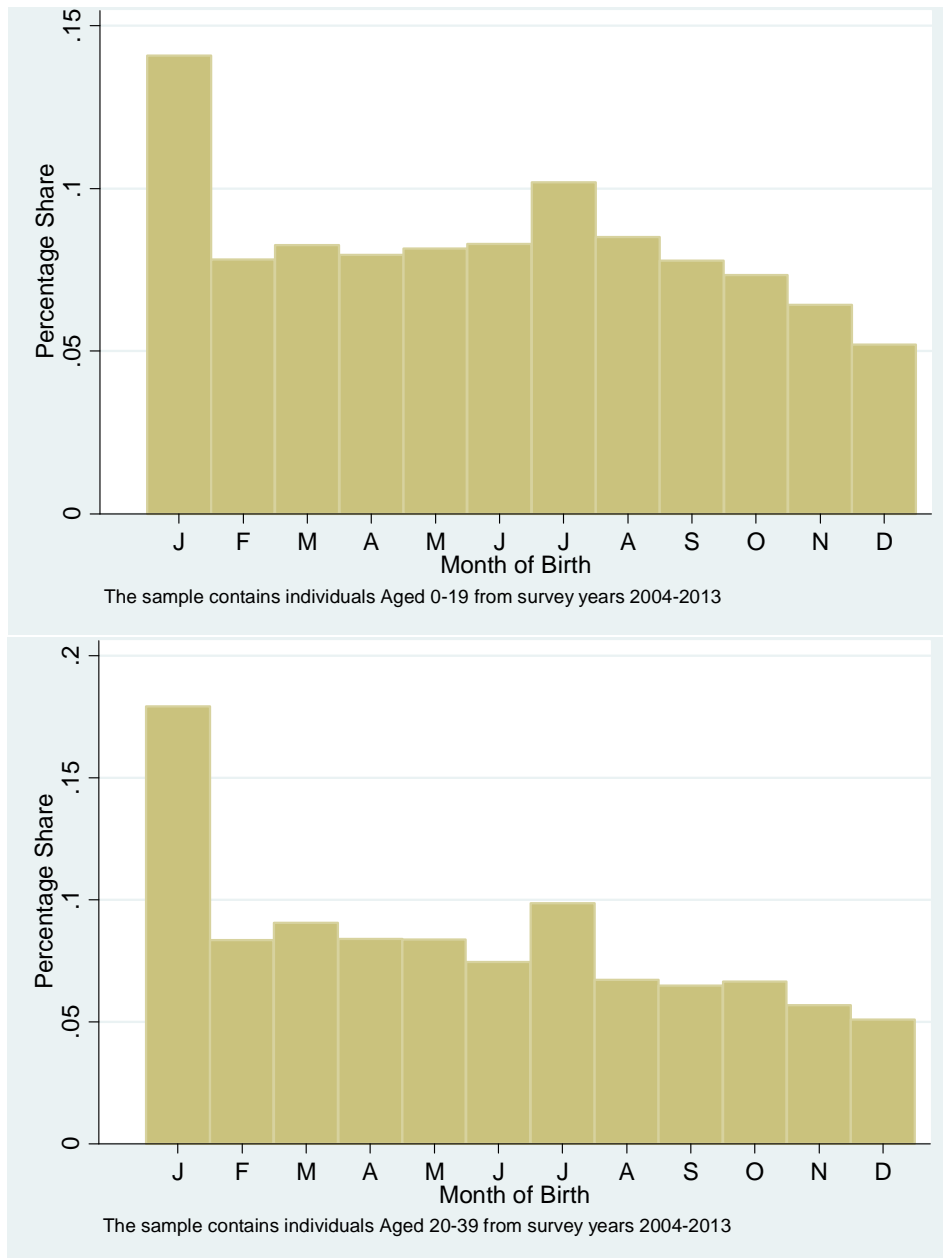
Figure 1: **Month of birth distribution:** Age 0–19 (upper panel) and 20–39 (lower panel).
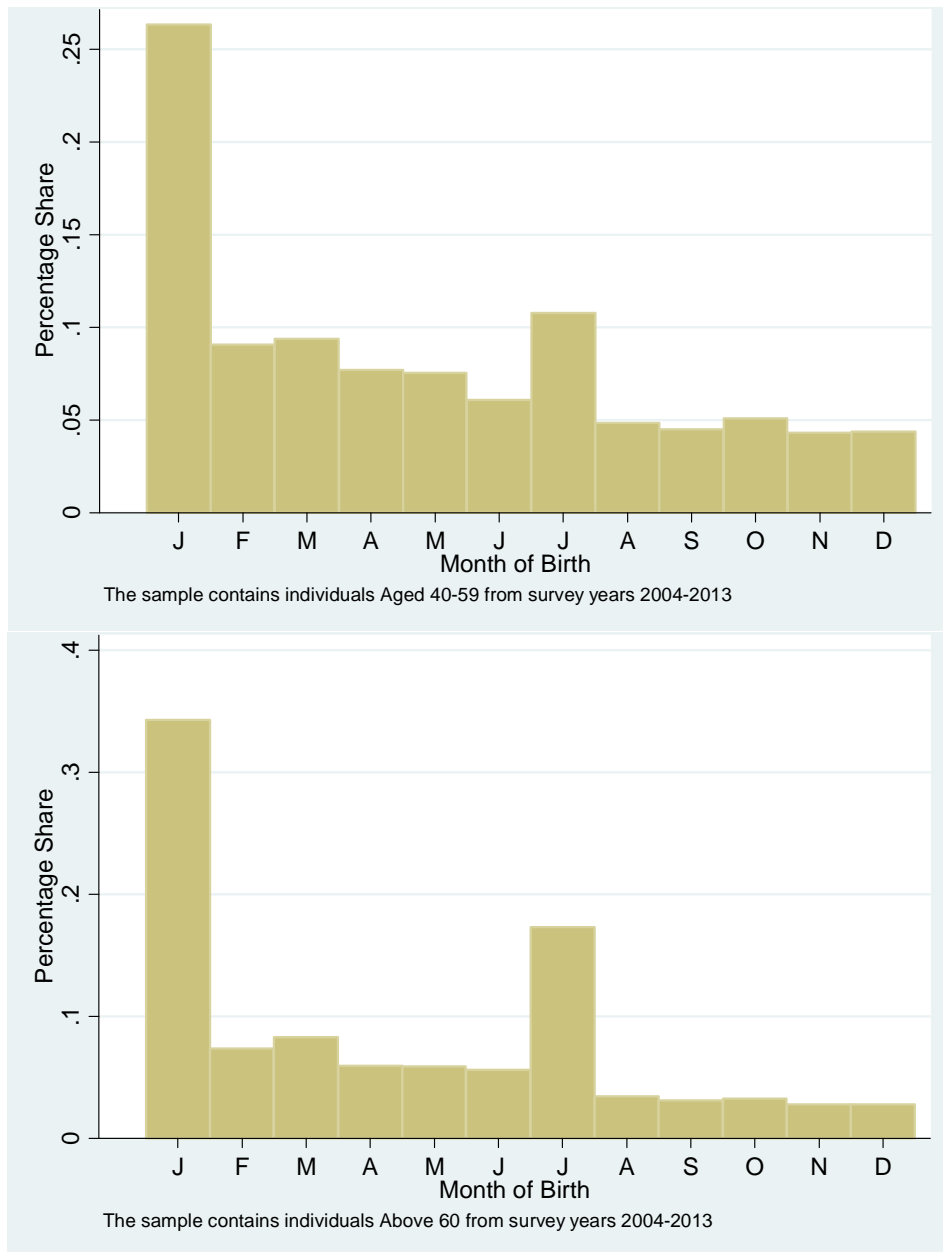
Figure 2: **Month of birth distribution:** Age 40–59 (upper panel) and above 60 (lower panel).

The sample contains individuals above 30 from all survey years 2004-2013
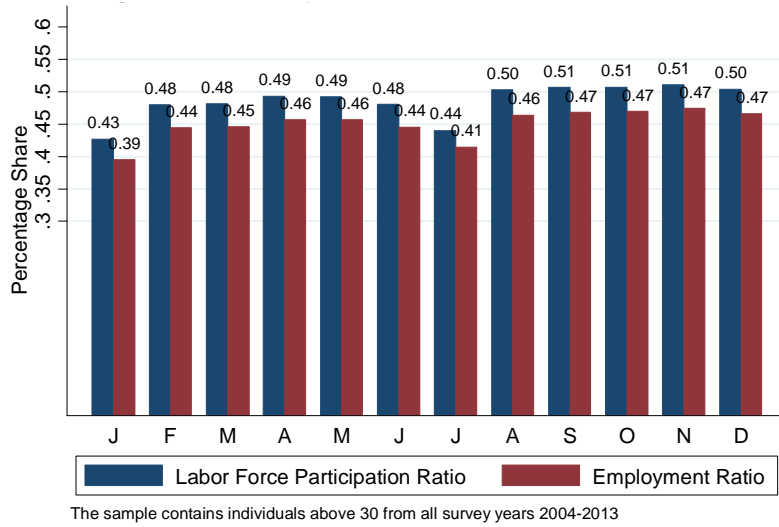
Figure 3: **Employment and labor force participation across months of birth.**

The sample contains salary employees from all survey years 2004-2013
The monthly wages were deflated using CPI index.
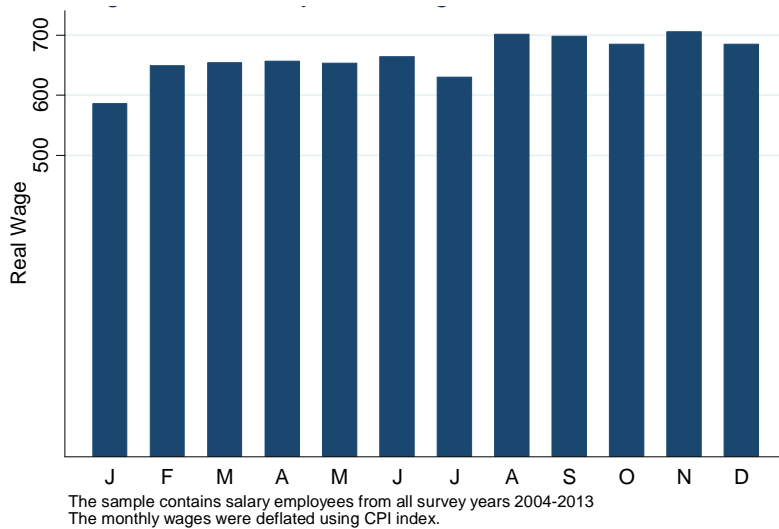
Figure 4: **Real wages across months of birth.**

The sample contains individuals above 30 from all survey years 2004-2013
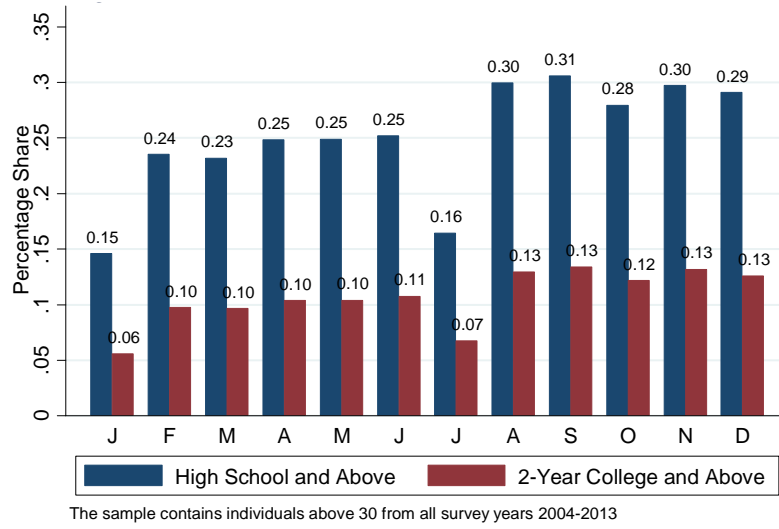
Figure 5: **Educational attainment across months of birth.**
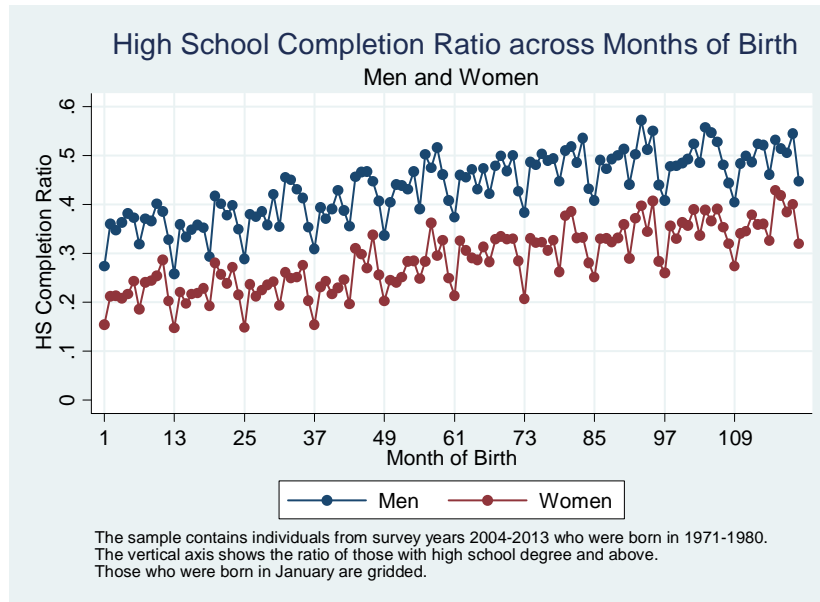
Figure 6: **High school completion rates across months of birth by gender categories.**
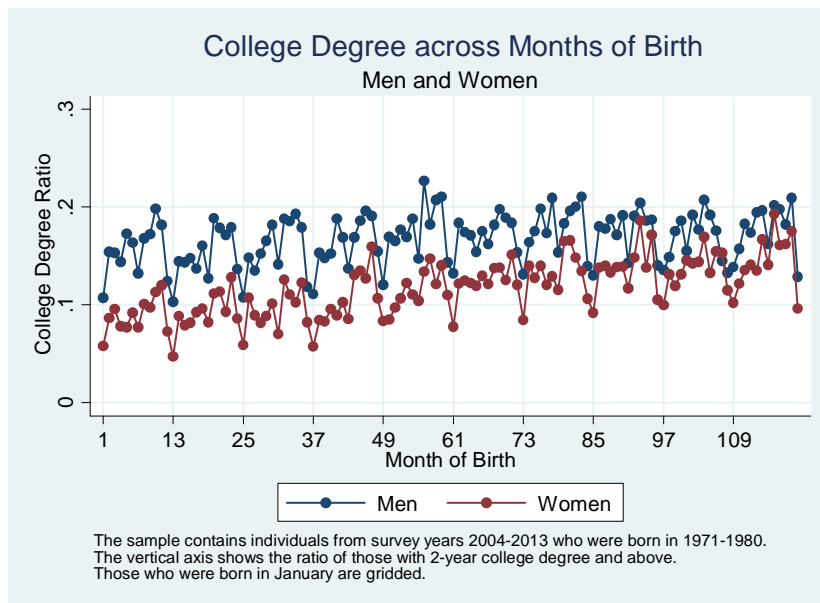
Figure 7: **College completion rates across months of birth by gender categories.**

The sample contains mothers of children below 15 from all survey years 2004-2013
The horizontal axis shows the months of birth of children.
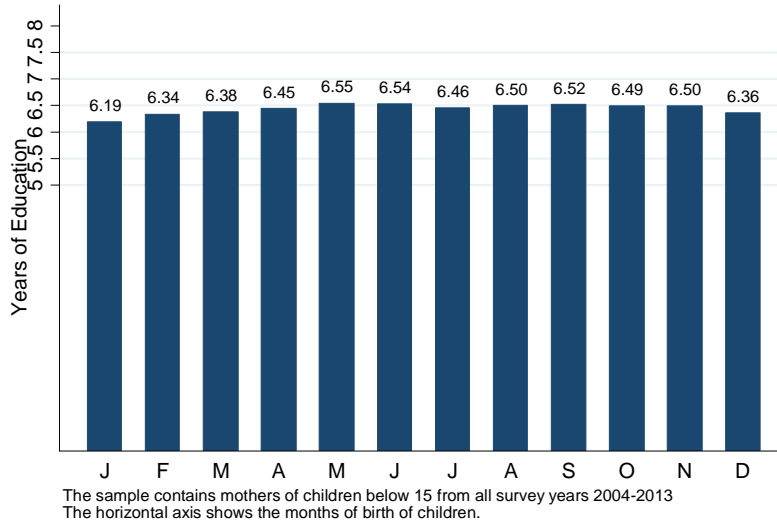
Figure 8: **Mother's years of education across months of birth.**

The sample contains fathers of children below 15 from all survey years 2004-2013.
The horizontal axis shows the months of birth of children.
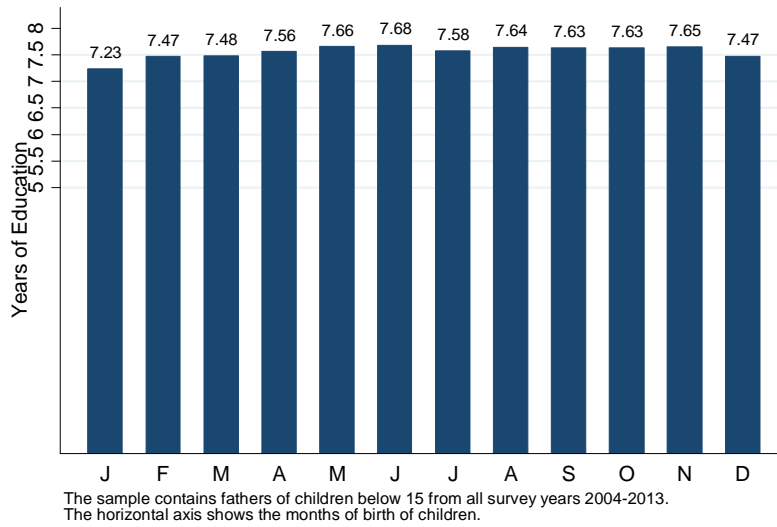
Figure 9: **Father's years of education across months of birth.**

Summary Statistics

M-o-b	Age				Education		Urban/Rural		Gender		Total
	0–19	20–39	40–59	Above 60	Below HS	HS and Above	Urban	Rural	Male	Female	
January	0.14	0.18	0.26	0.34	0.21	0.16	0.19	0.22	0.21	0.20	**0.20**
February	0.08	0.08	0.09	0.07	0.08	0.09	0.08	0.08	0.08	0.08	**0.08**
March	0.08	0.09	0.09	0.08	0.09	0.09	0.09	0.09	0.09	0.09	**0.09**
April	0.08	0.08	0.08	0.06	0.08	0.08	0.08	0.08	0.08	0.08	**0.08**
May	0.08	0.08	0.08	0.06	0.08	0.08	0.08	0.07	0.08	0.08	**0.08**
June	0.08	0.08	0.06	0.06	0.07	0.08	0.07	0.07	0.07	0.07	**0.07**
July	0.10	0.10	0.11	0.17	0.11	0.09	0.10	0.13	0.11	0.11	**0.11**
August	0.09	0.07	0.05	0.04	0.06	0.07	0.07	0.06	0.06	0.07	**0.07**
September	0.08	0.06	0.05	0.03	0.06	0.07	0.07	0.05	0.06	0.06	**0.06**
October	0.07	0.07	0.05	0.03	0.06	0.07	0.06	0.06	0.06	0.06	**0.06**
November	0.07	0.06	0.04	0.03	0.05	0.06	0.06	0.05	0.05	0.05	**0.05**
December	0.05	0.05	0.04	0.03	0.05	0.06	0.05	0.04	0.05	0.05	**0.05**

Table 1: **Summary Statistics.** The table reports the distribution of the population across the months of birth with respect to the specified variable. 2004–2013 waves of the Turkish Household Labor Force Survey are used. The total number of observations in our sample is 4,624,686.

Employment and Labor Force Participation

	Employment		Labor Force Part.	
	[1]	[2]	[3]	[4]
January	omitted	omitted	omitted	omitted
February	-0.0053***	-0.0114***	-0.0056***	-0.0118***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
March	-0.0064***	-0.0118***	-0.0063***	-0.0119***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
April	-0.0052***	-0.0120***	-0.0048***	-0.0117***
	(0.0010)	(0.0009)	(0.0009)	(0.0009)
May	-0.0074***	-0.0138***	-0.0078***	-0.0143***
	(0.0010)	(0.0009)	(0.0009)	(0.0009)
June	-0.0047***	-0.0127***	-0.0040***	-0.0120***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
July	0.0029***	0.0013	0.0034***	0.0017**
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
August	-0.0029***	-0.0136***	-0.0013	-0.0121***
	(0.0011)	(0.0011)	(0.0010)	(0.0010)
September	-0.0016	-0.0124***	-0.0017	-0.0126***
	(0.0011)	(0.0011)	(0.0011)	(0.0010)
October	-0.0038***	-0.0127***	-0.0042***	-0.0132***
	(0.0011)	(0.0011)	(0.0010)	(0.0010)
November	-0.0026**	-0.0132***	-0.0030***	-0.0136***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
December	-0.0043***	-0.0145***	-0.0045***	-0.0149***
	(0.0012)	(0.0012)	(0.0011)	(0.0011)
Education	No	Yes	No	Yes
Other Controls	Yes	Yes	Yes	Yes
Intercept	0.381***	0.368***	0.409***	0.394***
	(0.0019)	(0.0022)	(0.0018)	(0.0022)
R^2	0.302	0.318	0.352	0.368
# of Obs.	3,375,722	3,375,722	3,375,722	3,375,722

Table 2: **Employment and Labor Force Participation:** ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Robust standard errors are reported in parentheses. All individuals of age 15 and above are included into the sample. The sample size is 3,375,722. Turkish Household Labor Force Survey 2004–2013 waves are used in estimations. The coefficients should be read relative to January (the omitted month of birth dummy). The six education categories include no degree, primary school, middle school, high school, vocational high school, and college & above. Other controls include gender, age dummies, marital status, urban/rural status, region dummies, and survey-year dummies.

Real Wages and Formal/Informal Employment

	Real Wages		Formal Employment	
	[1]	[2]	[3]	[4]
January	omitted	omitted	omitted	omitted
February	31.50***	14.08***	0.0131***	0.0060***
	(2.197)	(2.085)	(0.0013)	(0.0013)
March	31.56***	15.70***	0.0156***	0.0090***
	(2.149)	(2.040)	(0.0013)	(0.0012)
April	34.70***	16.65***	0.0159***	0.0079***
	(2.232)	(2.118)	(0.0013)	(0.0013)
May	33.45***	15.67***	0.0142***	0.0066***
	(2.237)	(2.123)	(0.0013)	(0.0013)
June	41.38***	18.50***	0.0148***	0.0061***
	(2.356)	(2.237)	(0.0014)	(0.0014)
July	20.26***	12.02***	0.0032***	0.0024**
	(2.128)	(2.019)	(0.0012)	(0.0012)
August	54.98***	29.43***	0.0197***	0.0086***
	(2.431)	(2.308)	(0.0015)	(0.0015)
September	52.64***	26.38***	0.0212***	0.0100***
	(2.477)	(2.352)	(0.0015)	(0.0015)
October	47.06***	23.07***	0.0198***	0.0103***
	(2.456)	(2.332)	(0.0015)	(0.0015)
November	56.74***	30.38***	0.0227***	0.0119***
	(2.578)	(2.448)	(0.0016)	(0.0016)
December	47.40***	22.92***	0.0206***	0.0097***
	(2.672)	(2.537)	(0.0016)	(0.0016)
Education	No	Yes	No	Yes
Other Controls	Yes	Yes	Yes	Yes
Intercept	80.90***	90.66***	-0.0201***	-0.0719***
	(8.224)	(10.020)	(0.0043)	(0.0048)
R^2	0.210	0.289	0.442	0.455
# of Obs.	779,529	779,529	1,396,211	1,396,211

Table 3: **Real Wages and Formal/Informal Employment:** ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Robust standard errors are reported in parentheses. The sample size is 779,539 for the wage regressions and 1,396,211 for the informal employment regressions. Turkish Household Labor Force Survey 2004–2013 waves are used in estimations. The coefficients should be read relative to January (the omitted month of birth dummy). The six education categories include no degree, primary school, middle school, high school, vocational high school, and college & above. Other controls include gender, age dummies, marital status, urban/rural status, part/full-time job status, industry dummies, region dummies, and survey-year dummies.

Educational Outcomes

	Years of sch.	High school	College
January	omitted	omitted	omitted
February	0.730***	0.0863***	0.0377***
	(0.0104)	(0.0013)	(0.0010)
March	0.700***	0.0818***	0.0365***
	(0.0103)	(0.0013)	(0.0010)
April	0.784***	0.0914***	0.0411***
	(0.0107)	(0.0014)	(0.0010)
May	0.765***	0.0894***	0.0401***
	(0.0107)	(0.0014)	(0.0010)
June	0.843***	0.0990***	0.0463***
	(0.0112)	(0.0014)	(0.0010)
July	0.354***	0.0409***	0.0156***
	(0.0101)	(0.0013)	(0.0009)
August	1.073***	0.1260***	0.0592***
	(0.0119)	(0.0015)	(0.0011)
September	1.113***	0.1290***	0.0622***
	(0.0122)	(0.0016)	(0.0011)
October	0.920***	0.1060***	0.0518***
	(0.0119)	(0.0015)	(0.0011)
November	1.046***	0.1210***	0.0603***
	(0.0124)	(0.0016)	(0.0012)
December	1.036***	0.1190***	0.0564***
	(0.0125)	(0.0016)	(0.0012)
Controls	Yes	Yes	Yes
Intercept	8.414***	0.396***	0.182***
	(0.0167)	(0.0021)	(0.0015)
R^2	0.168	0.143	0.068
# of Obs.	2,460,398	2,460,398	2,460,398

Table 4: **Educational Outcomes:** ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Robust standard errors are reported in parentheses. All individuals of age 30 and above are included into the sample. The sample size is 2,460,398. Turkish Household Labor Force Survey 2004–2013 waves are used in estimations. The coefficients should be read relative to January (the omitted month of birth dummy). Controls include gender, age dummies, marital status, urban/rural status, region dummies, and survey-year dummies.

Regression Discontinuity Design Exercises

	Log Earnings		Years of Education
	[1]	**[2]**	**[3]**
Treatment (1965–66)	0.1684***	0.0300*	1.0027***
	(0.0232)	(0.0174)	(0.0713)
Treatment (1966–67)	0.0887***	-0.0030	0.9173***
	(0.0233)	(0.0175)	(0.0778)
Treatment (1967–68)	0.0962***	0.0157	0.7975***
	(0.0210)	(0.0170)	(0.0764)
Treatment (1968–69)	0.1594***	-0.0010	0.9737***
	(0.0219)	(0.0167)	(0.0735)
Treatment (1969–70)	0.1129***	0.0224	0.7426***
	(0.0203)	(0.0152)	(0.0664)
Treatment (1970–71)	0.1012***	0.0235*	0.7700***
	(0.0186)	(0.0142)	(0.0659)
Treatment (1971–72)	0.1367***	-0.0032	0.9320***
	(0.0199)	(0.0155)	(0.0699)
Treatment (1972–73)	0.0824***	-0.0093	0.8909***
	(0.0199)	(0.0153)	(0.0683)
Treatment (1973–74)	0.1134***	0.0312**	0.7775***
	(0.0190)	(0.0145)	(0.0678)
Treatment (1974–75)	0.1308***	0.0436***	0.8678***
	(0.0199)	(0.0016)	(0.0747)
Education	No	Yes	No
Other Controls	Yes	Yes	Yes

Table 5: **Regression Discontinuity Design:** ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Robust standard errors are reported in parentheses. Turkish Household Labor Force Survey 2004–2013 waves are used in estimations. Controls include gender, age dummies, marital status, urban/rural status, region dummies, and survey-year dummies.