Journal of Health Economics 49 (2016) 120–135

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Cost savings of developmental screenings: Evidence from a nationwide program

Martin Halla [a,b,c], Gerald J. Pruckner [c,d,*], Thomas Schober [c,d]

[a] University of Innsbruck, Austria
[b] IZA, Institute for the Study of Labor, Bonn, Germany
[c] Christian Doppler Laboratory – Aging, Health and the Labor Market, Linz, Austria
[d] Johannes Kepler University of Linz, Austria

ARTICLE INFO

Article history:
Received 4 September 2015
Received in revised form 13 June 2016
Accepted 28 June 2016
Available online 4 July 2016

JEL classification:
I12
J13
I18
H51
H75

Keywords:
Early intervention
Child development
Developmental disorder
Developmental screening
Healthcare cost

ABSTRACT

Early intervention is considered the optimal response to developmental disorders in children. We evaluate a nationwide developmental screening program for preschoolers in Austria and the resulting interventions. Identification of treatment effects is determined by a birthday cutoff-based discontinuity in the eligibility for a financial incentive to participate in the screening. Assigned preschoolers are 14.5 percentage points more likely to participate in the program. For participants with high socio-economic status (SES), we find little evidence for interventions and consistently no effect on healthcare costs in the long run. For low SES preschoolers, we find evidence for substantial interventions, but only weak evidence for cost savings in the long run.

1. Introduction

A growing body of literature across different academic disciplines traces the origins of life-cycle well-being to the very early stages of life (Currie and Rossin-Slater, 2015). One important aspect is early-life health. A variety of policies, such as prenatal care, family leave, nurse home visiting, or early childhood center-based interventions, have the potential to improve health conditions at different stages of early childhood. In this study, we are concerned with medical care interventions for preschoolers with respect to developmental disorders. We are particularly interested in the identification of affected preschoolers, a step that predates any diagnosis or treatment.

An estimated 14 percent of all children in the US have some form of developmental disorder (Boyle et al., 2011). There is widespread agreement among medical specialists and policymakers that early identification of developmental disorders in children is essential for optimal intervention.[1] Developmental disorders, or delayed development, can be caused by specific medical conditions and may indicate an increased risk of other medical complications, as well as emotional and behavioral disorders. Early identification of developmental problems enables further evaluation, diagnosis, and treatment (Chakrabarti and Fombonne, 2001).

Successful intervention improves the well-being of families with affected children. If affected families have predominantly lower socioeconomic status (SES), such early intervention can be perceived as socially fair, since it helps to reduce (health) inequalities. The economic efficiency-based argument for early intervention rests on the simple comparison between the costs of intervention (today) and the costs of non-intervention (later). Proponents typically assume

[1] This view is in line with a growing body of literature pointing to the importance of early childhood in building the foundations for lifelong health. David J. Barker (see, e.g., Barker, 1995) developed the argument that the prenatal environment affects health conditions in adulthood, including heart disease and diabetes. Equivalent reasoning is documented in the literature on human capital, in which substantial benefit from early interventions arises because human capital formation is dynamic in nature (Cunha and Heckman, 2007; Almond and Currie, 2011).

* Corresponding author. Department of Economics, Johannes Kepler University of Linz, Altenberger Straße 69, A-4040 Linz, Austria.
E-mail address: gerald.pruckner@jku.at (G.J. Pruckner).

http://dx.doi.org/10.1016/j.jhealeco.2016.06.011
0167-6296/

that early intervention is more cost-effective than later remediation (Conti and Heckman, 2013).

While these theoretical arguments make a compelling case for early intervention, they do not provide guidance on how to implement intervention. In practice, a crucial point is the identification of developmental disorders that predate any diagnosis or treatment. Typically, developmental screening programs are used. For instance, the *American Academy of Pediatrics* officially recommends that a standardized developmental screening tests should be administered regularly at the ages of 9, 18, and 30 months.[2] Depending on age, these screening tests inspect the development of motor skills and coordination, visual and hearing abilities, communication and language skills, and cognitive abilities. Ideally, screening identifies all developmental disorders in these dimensions and initiates a comprehensive and purposeful response.

Thus far, the literature has not provided rigorous evaluation of physician-based developmental screenings for preschoolers. This is especially surprising given the extensive recommendations made by professional organizations and government agencies. Moyer and Butler (2004) conduct a systematic review of the literature for any pediatrician-based developmental screening and conclude that methodologically sound randomized controlled trials (RCTs) of developmental screenings do not exist. A more recent systematic review focusing on vision screening (Chou et al., 2011) concludes that there is no RCT that compares the effect of screening with non-screening.[3] Cadman et al. (1987) is the only exception we are aware of. Based on an RCT, the authors evaluate the effectiveness of a screening program for 4- to 5-year-old children, which includes general health interviews, and hearing and vision tests administered by public health nurses. They could not detect any effect of the intervention on developmental attainment or school performance 3 years after the screening.[4]

Even if it seems obvious that early intervention is desirable and most likely efficient, the literature has paid insufficient attention to the identification of developmental disorders and associated costs. In this study, we are interested in not only the intervention, but also the screening process that precedes any intervention. Depending on the context, the costs of identifying developmental disorders may vary strongly.

We evaluate a nationwide developmental screening program of preschoolers and subsequent medical interventions in Austria. Austria is a high-income country with a Bismarckian healthcare system offering a prenatal and early postnatal healthcare program that is free of charge and fully financially incentivized. In a subsequent developmental screening program, parents are offered examinations for their children, inter alia, at the ages of 24, 36, and 48 months. Parents may consult any contracted pediatrician or general practitioner (GP) who executes a predefined age-specific developmental screening procedure. This comprises physical examinations, assessment of a child's mental development, and identification of behavioral problems. In case of any abnormal results, the doctor will either schedule a follow-up appointment or refer the child to other professionals. The developmental screening itself and any follow-up appointment are fully covered by statutory health insurance.

In 2000, one provincial government (Upper Austria) introduced a financial incentive to promote developmental screening participation. Irrespective of their household income, families are offered €185 if their child participates in all three screenings, including some stipulated vaccinations. The only eligibility criterion is that the child was born on January 1, 2000 or later. We exploit this sharp birthday cutoff-based discontinuity in the eligibility to obtain exogenous variation in participation. We find that assigned preschoolers are – irrespective of their SES – 14.5 percentage points more likely to be screened.

To assess the cost savings potential of this screening program, we use high-quality administrative data. These provide information on the scope of intervention and long-term healthcare costs. The scope of intervention is quantified by short-run healthcare expenditure for follow-up treatments by the screening doctors and referrals to other specialists. If screening participation increases the likelihood of identifying a disorder, we expect an increase in short-run follow-up expenditure, compared to the counter-factual situation of non-participation. The assessment of the program's cost-saving potential depends on whether and to what extent the savings in the long run exceed the increase in expenditure due to early intervention. We observe the healthcare spending for preschoolers up to 11 years of age. Since we do not observe any direct measures of preschoolers' well-being, we cannot provide a comprehensive welfare-based cost–benefit analysis. The program may generate quality of life increases that are possible through early intervention only and not later spending.[5]

Based on a fuzzy *regression discontinuity design* (RDD), we find that the program is clearly *not* effective for preschoolers with higher SES, who comprise about 75 percent of the total preschooler population. For this group, we obtain a consistent picture with little evidence for interventions (the only exception is follow-up examinations by ophthalmologists), and consequently there is no effect on healthcare costs in the long run. For low SES preschoolers, who comprise about 25 percent of the total preschooler population, the interpretation is less clear. While we find clear evidence for interventions with follow-up examinations by several medical specialists, there is only weak evidence for cost savings in the long run. As an alternative interpretation, we consider the increase in healthcare expenditure in the short run not as an intervention addressing developmental disorders, but as supply-induced over-treatment resulting from profit-maximizing screening doctors. An additional estimation analysis focusing on "referred" follow-up examinations, without any financial benefit for the screening doctors, provides evidence that at least part of the increase in short-run healthcare expenditure is due to justified interventions.

These results have to be interpreted in the context of the Austrian healthcare system. There are financial incentives for health screenings up to the second year of life and participation rates are high (see Section 3.1.1). Moreover, parents can always consult medical specialists independently of participation in the program and free of charge. We consider our results representative of a European welfare state, for which we conclude that general physician-based

[2] See Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children With Special Needs Project Advisory Committee (2006) and reaffirmation for this policy in the American Academy of Pediatrics (2010).

[3] Williams et al. (2002) compare more intensive and less intensive screening. They focus on the detection and early treatment of amblyopia. The control group was assigned to a single intensive orthoptic screening at 37 months of age. The treatment group was screened five times (at 8, 12, 18, 25, and 37 months of age). The main result is that amblyopia was significantly less prevalent among the treatment group at the age of 7.5 years.

[4] In contrast to screening activities in the physician's office, there is substantial evidence on the effectiveness of home visit programs and more comprehensive center-based preschool interventions. Both types of interventions typically focus on children at risk because of low parental income or other adverse social circumstances and often comprise a developmental screening component. Meta-analyses confirm the effectiveness of home visit programs (Avellar and Supplee, 2013; Peacock et al., 2013; Sweet and Appelbaum, 2004) and center-based preschool interventions (Duncan and Magnuson, 2013) for improving the outcomes of participating families along several dimensions, including children's long-run health outcomes (Campbell et al., 2014; Carneiro and Ginja, 2014).

[5] For example, the prescription of glasses for children with visual impairment in due time generates health benefits, irrespective of short- and long-run out-of-pocket healthcare costs.

developmental screenings for preschoolers should be promoted only among preschoolers with lower SES.

The remainder of this paper is organized as follows. In Section 2, we briefly outline the theoretical determinants of the effectiveness of developmental screening programs. In Section 3, we present our research design. In Section 4, we present our estimation results along with several robustness checks. In Section 5, we provide an overall assessment of our estimation results. Finally, Section 6 formulates policy recommendations and concludes.

2. Theoretical considerations

From a theoretical point of view, a necessary condition for a cost-effective intervention is that the screening program identifies any developmental disorders that can be treated. Thus, if the subjects have no or only non-treatable disorders, the intervention is bound to fail (*Case 1*). Similarly, if treatable disorders exist but are not identified, then the intervention would certainly be ineffective (*Case 2*).

Assuming that the screening program identifies treatable disorders, the outcome in the counter-factual situation is decisive. It is possible that the disorder would have been diagnosed and treated without any screening (*Case 3*). For instance, proactive parents could have consulted a pediatrician anyway. In this case, the timing is relevant. If the screening caused an earlier treatment that would have improved the child's well-being and potentially reduced future healthcare costs, then the intervention could be cost-effective. If the screening caused no (or no beneficial) earlier treatment, we would consider it ineffective; resources would have been wasted on subjects without any disorders. The final case of a potentially effective screening is provided: if disorders would not have been diagnosed in the counter-factual situation without any screening (*Case 4*).

In *Cases 3* and *4*, in which screening leads to an (earlier) treatment of a disorder, cost-effectiveness would require that long-run cost decreases outweigh the increases in short-term treatment costs, including those for the screening itself. The screening costs should not only cover direct out-of-pocket expenditure, but also incorporate potential harm or considerable discomfort caused by the screening procedure itself, as well as any costs due to false positive outcomes that result in anxiety and/or over-treatment. In addition, costs should account for any follow-up medical examinations that are triggered by the screening. Screening-performing doctors may conduct further examinations or refer preschoolers to specialists for more detailed diagnostic services. The likelihood of further examinations increases with the doctor's level of risk aversion and decreases with his or her expertise and experience. Moreover, depending on the institutional setting, it cannot be ruled out that some further examinations are not justified medically. If screening-performing doctors benefit financially from further examinations, supply-induced demand may lead to unnecessary follow-up examinations.

Thus, it is ultimately an empirical question whether a certain developmental screening program is effective. Nonetheless, we conclude that a developmental screening program is more likely to be effective with the following conditions:

- if it focuses on subjects who are likely to have easily identifiable and treatable disorders,
- if untreated disorders cause substantial costs for the patient and society,
- if disorders most likely remain undetected for a long time in the counter-factual situation without screening, and
- if screening costs are low.
  - Screening costs are expected to be low if performing doctors are specialized and experienced, and have no financial interest in further examinations.

3. Research design

In this section, we first describe the institutional background, including the details of the developmental screening program and the financial incentives for participation in Upper Austria. Then, we present our data sources and estimation strategy, and discuss the identifying assumptions.

3.1. Institutional background

Austria has a Bismarckian welfare system with almost universal access to high-quality healthcare. Insurance is compulsory and, in general, is linked to employment. Our data cover private-sector employees (about 75 percent of the population) who are, depending on the type and location of the employer, insured with one of nine regional health insurance funds (in German, *Gebietskrankenkassen*).[6] Thus, workers have no choice regarding their healthcare provider or insurance package.

The outpatient healthcare system is funded predominantly by wage-related social security contributions of employers and employees. The inpatient sector is co-financed by social security contributions and general tax revenues from different federal levels. Health insurance contributions increase proportionally with income up to a ceiling, but are independent of the personal risk of the insured. The health insurance funds cover all healthcare expenditure in the inpatient and outpatient sector, including maternity and the institutionalized mother–child screening program.

The Organisation for Economic Co-operation and Development (OECD) concludes that the Austrian healthcare system delivers good quality and easily accessible services, albeit at very high costs (Gönenc et al., 2011).[7] The extensive provision of care with wide patient choice among in- and outpatient providers (i.e., there is no strict gate-keeping) combined with fee-for-service remuneration of doctors tends to produce high volumes of services. It is argued that the governance and funding structure is highly fragmented and overuses expensive inpatient healthcare services. In summary, the system predominantly operates on a supply-driven basis and does not have clear mechanisms to optimize spending on a cost–benefit or cost-effectiveness basis.

3.1.1. Mother–child-pass examination program

Public prenatal care has been established in Austria for decades. In 1974, the *Austrian Federal Ministry of Health* launched the first nationwide prenatal screening program.[8] This so-called *Mother–Child-Pass Examination Program* (MCPEP) consisted initially of four prenatal examinations. Over time, the aim and scope of the MCPEP have expanded substantially. Currently, it is a comprehensive screening program that monitors the health of expectant mothers and their children over a period of about 70 months. It starts with the first diagnosis of pregnancy (ideally before the 16th week) and lasts until the 5th year of the child's life. In total, it comprises five prenatal examinations, five postnatal examinations of infants (up to 14 months), and three developmental screenings of toddlers and

[6] Non-employed individuals are also covered by the regional health insurance funds. Farmers, other self-employed people, civil servants, and employees of the Austrian Railway Company and the mining industry have their own nationwide health insurance institutions. Moreover, there are six company-specific health insurance funds.

[7] Both the life expectancy and per capita of total health spending of Austria are above the OECD average. Infant mortality is below the OECD average; however, it is significantly higher compared to Scandinavian countries.

[8] At the time this program was launched, infant mortality was comparably high in Austria, amounting to 24 deaths of infants under the age of 1 year per 1000 live births. This was somewhat above the US figures. Since then, infant mortality rates in Austria have declined continuously, and are currently well below the US rates (own calculations based on data from the *World Bank*).

Table 1
Overview of the mother–child-pass examination program.

No.	Age	Examinations	Financial incentive	
			Nationwide	Upper Austria
Prenatal screening examinations				
1	Until 16th week	Obstetric examination; laboratory tests	Yes	Yes
	8th to 12th week	Ultrasound[a]		
2	17th to 20th week	Obstetric examination; internal examination	Yes	Yes
	18th to 22nd week	Ultrasound[a]		
3	25th to 28th week	Obstetric examination; laboratory tests	Yes	Yes
4	30th to 34th week	Obstetric examination	Yes	Yes
	30th to 34th week	Ultrasound[a]		
5	34th to 38th week	Obstetric examination	Yes	Yes
Postnatal screening examinations of infants (up to 14 months)				
1	1st week	Child examination; hip ultrasound[b]	Yes	Yes
2	4th to 7th week	Child examination; orthopedic examination	Yes	Yes
	6th to 8th week	Hip ultrasound[b]		
3	3rd to 5th month	Child examination	Yes	Yes
4	7th to 9th month	Child examination; ear, nose, and throat examination	Yes	Yes
5	10th to 14th month	Child examination; eye examination	Yes	Yes
Developmental screening examinations of toddlers and preschoolers (from 2 to 4 years)				
D1	22nd to 26th month	Anamnesis; physical examination; nutritional status; behavior; mental development;	No	Yes[c]
D2	34th to 38th month	speech and language; comprehensive eye and vision examination at D1	No	Yes[c]
D3	46th to 50th month		No	Yes[c]

[a] Ultrasound exams are recommended but not required for receipt of financial incentive.
[b] The hip ultrasound examination is recommended but not required for receipt of financial incentive.
[c] Only children born on January 1, 2000 or later are eligible for the financial incentive.

preschoolers (from 24 to 48 months). Table 1 summarizes the timeframe, type, and incentive structure for all examinations. Our focus is on the three developmental screenings (D1 to D3), which we discuss in more detail below. All stipulated examinations are free of charge, even for mothers without social health insurance coverage. Generally, the examinations are provided by outpatient care gynecologists, pediatricians, and GPs.[9]

As part of the program, expectant mothers receive the so-called "mother–child pass" (in German, *Mutter-Kind-Pass*). This official document issued by the *Austrian Federal Ministry of Health* is a booklet documenting all examinations and their results. In addition, an expectant mother receives an international certificate of vaccination for her child, and an additional information booklet containing advice on a variety of relevant topics. This documentation is important for mothers, not least because participation in the program is a prerequisite for receipt of financial payments.

Participation in the MCPEP traditionally has been financially incentivized at a federal level. However, the subsidy has never applied to the developmental screening part of the program (D1, D2, and D3). Mothers receive financial transfers if they have participated in the five prenatal and five postnatal examinations of infants up to 14 months (see column four of Table 1). The specific regulations have varied over time. For our empirical analysis, the period from 1997 through 2001 is mostly relevant, when each eligible mother who participated in the five prenatal and five postnatal examinations of infants up to 14 months received a one-time payment of €145.

3.1.2. Developmental screening program

Although the development screening part of the MCPEP (D1, D2, and D3) has never been incentivized at a federal level, the Upper Austrian provincial government introduced a so-called "mother–child subsidy" (MC subsidy) (in German *Mutter-Kind-Zuschuss*) for all children born on January 1, 2000 or later.[10] Eligible mothers received €185 if they participated in all three developmental screening examinations (D1, D2, and D3), including vaccinations. This regulation was enacted by the Upper Austrian government in November 2001. All mothers, irrespective of their household income, were eligible. The application had to be filed within 1 year of the 5th birthday of the child.

In our empirical analysis, we exploit this sharp discontinuity in eligibility by date of birth for the Upper Austrian MC subsidy. This creates a clear distinction between treated and control units. Mothers whose children were born before January 1, 2000 had no financial incentive to participate in D1, D2, and D3. By contrast, mothers whose children were born on January 1, 2000 or later (henceforth, "assigned mothers") received €185. As we show below, assigned mothers are 14.5 percentage points more likely to participate in D1 to D3. Under very weak identifying assumptions (see Section 3.3), this allows us to evaluate the effect of these preschool examinations on the subsequent healthcare expenditure up to preadolescence.

In support of the federal incentive, the Upper Austrian government offered another €185 to mothers whose children were born after January 1, 2000 if they participated in the five prenatal and five postnatal examinations of infants up to 14 months. Given that this regulation was not enacted before November 2001, it has no impact on the evaluation of the developmental screenings. Since the age of the youngest child in our estimation sample was 19 months, he or she was too old for the mother to react to this incentive. The only potential confounding factor is an income effect. Children born after January 1, 2000 who participated in all stipulated examinations (without their mothers knowing that financial incentives would be introduced in the future) received ex post €185. We consider this amount too low to affect child health or children's volume of healthcare expenditure.

In the examinations D1 to D3, the child's development is examined at the ages of about 24, 36, and 48 months. The examinations

[9] For risk pregnancies, hospitals provide selected services related to prenatal diagnostics. In general, the first postnatal check-up is conducted in the same hospital immediately after birth. Moreover, if any complication occurs, both the mother and child are always referred to the hospital for follow-up examinations that are beyond the scope of the MCPEP.

[10] Upper Austria is one of nine provinces in Austria. It comprises about one sixth of the Austrian population and workforce. It was the only federal state that offered a financial incentive.

have a preventative character that promotes early detection and timely treatment of health risks. The examinations include an anamnesis that covers potential behavioral disorders, previous diseases, and the speech and language development of the child. In a physical examination, the doctor tests the child's ears and eyes and examines his or her teeth and organs. Furthermore, the child's weight, height, and head circumference are recorded, and motor skills, nutritional status, and mental development are assessed. Finally, a comprehensive eye and vision examination is scheduled for D1, and measurements of blood pressure are scheduled for D2 and D3.

Parents can freely choose either a contracted pediatrician or a GP to conduct the screening. It turns out that parents, irrespective of their SES, predominantly choose pediatricians, who perform about two-thirds of all screenings. Ophthalmologists carry out the eye examination part of the screenings. For D1 to D3, physicians receive a fixed payment of €21.8 per examination during the entire study period. This payment is worthwhile for the screening doctor given the reimbursement for a general consultation in the outpatient sector, which in 2011, for example, was €20.6 on average.

3.2. Data

In our empirical analysis, we use administrative data from the *Upper Austrian Health Insurance Fund*. It covers the sub-population of all private sector employees and their dependents in the province of Upper Austria. These data include, among others, detailed information on healthcare service utilization in the outpatient sector (i.e., medical attendance and drug use) and some inpatient sector information (e.g., number of days of hospitalization). Thus, we observe participation in the examinations stipulated by the MCPEP, as well as any other single doctor visit and each drug prescription. Furthermore, we can trace the screening doctors' referral behavior. It must be noted, however, that the Austrian outpatient sector does not impose a strict gate-keeping system. Even if health insurance funds strongly recommend a formal referral of a GP for any consultation of a medical specialist, each patient is allowed to visit one specialist per quarter in each medical field without written referral by the family doctor. The consultation of radiologists and utilization of laboratory services are excluded from this regulation.

We focus on children born between September 1998 and April 1999 and between September 1999 and April 2000. The reason for choosing these time windows is to obtain cohorts of children who are in the same school grade.[11] We verify that the choice of the window width is not decisive. Estimations based on a larger window (e.g., 6 months before and after the cutoff) are very comparable to those obtained by the main specification.

We observe healthcare costs for these children up to 10.5 years of age. The latest year available in our dataset is 2011 and all expenditure is measured in 2011 Euros (€1 is equivalent to 1.39 US$). We complement these data with information from the *Austrian Birth Register*, including detailed information about the birth (e.g., gestational length and birth weight) and socioeconomic information about the mother. We use this information to generate covariates for our regression analysis and to explore heterogeneous treatment effects in different sub-samples of the population.

3.3. Estimation strategy

Our estimation strategy exploits variation in the probability of developmental screening participation of preschooler *i* resulting from the birthdate B_i cutoff-based discontinuity in the eligibility for a financial incentive for participation. In other words, children born shortly before the cutoff ($B_i < c$) are less likely to be screened compared to children born shortly after the cutoff ($B_i > c$). Since the probability of participation does not jump from zero to one at this birthday cutoff, this represents a fuzzy RDD. The design can be translated into a *two-stage least square* setup, in which the birthday cutoff serves as an instrumental variable for treatment status.

In the first-stage equation, the dependent variable is the treatment status S_i. In our main specification, we define a binary indicator S_i^3, which is equal to one if preschooler *i* participated in *all* three developmental screening examinations (D1, D2, and D3). The explanatory variable of primary interest is the instrumental variable T_i, which is equal to one if preschooler *i* is born after January 1, 2000 ($T_i = 1$ if $B_i > c$), and zero otherwise:

$$S_i^3 = \alpha_0 + \alpha_1 T_i + \alpha_2 (B_i - c) + \alpha_3 (B_i - c)^2 + \alpha_4 (B_i - c) \times T_i + \alpha_5 (B_i - c)^2 \times T_i + \Theta \mathbf{X}_i + \epsilon_i. \quad (1)$$

Furthermore, we allow for different quadratic time trends in participation (second order polynomials) before ($(B_i - c)$, $(B_i - c)^2$) and after ($(B_i - c) \times T_i$, $(B_i - c)^2 \times T_i$) the birthday cutoff and control for a set of covariates ($\mathbf{X}_i$). The latter includes information on the preschooler's parity, gestational length, birth weight, legitimacy of birth, and the mother's age, citizenship, educational attainment, and a binary indicator distinguishing between urban and rural place of residence. All covariates are measured at the time of birth. The parameter of primary interest is α_1, which gives us the change in the likelihood of participation in the whole developmental screening program, depending on the eligibility for the financial incentive.

In the second-stage equation, we regress our respective outcome variable O_i on the predicted screening participation from the first stage $\hat{S}_i^3$:

$$O_i = \beta_0 + \beta_1 \hat{S}_i^3 + \beta_2 (B_i - c) + \beta_3 (B_i - c)^2 + \beta_4 (B_i - c) \times T_i + \beta_5 (B_i - c)^2 \times T_i + \Psi \mathbf{X}_i + \mu_i. \quad (2)$$

To evaluate the cost saving potential of the developmental screening program, we examine two outcome dimensions. First, to assess the scope of intervention, we examine subsequent days of hospitalization and short-run healthcare expenditure for follow-up treatments by the screening doctors or other medical specialists, and for medication. We aggregate different categories of expenditure (e.g. for different types of resident medical specialists) over the age range from 18 to 54 months of children's lives. Second, we measure healthcare expenditure in the long run, defined as the age range from 6 to 10.5 years of children's lives.[12] The upper age limit of 10.5 years results from the fact that the latest year available in our data is 2011. The lower age limit of 6 years is chosen against the background that we conceptually attempt to capture healthcare costs that do not comprise expenditure for intervention triggered by the screening. Given that the last screening exam, D3, takes place at the age of 4 years, we are confident that our measurements for long-run healthcare costs can be interpreted meaningfully as outcomes of screening participation (i.e., we implicitly assume that medical interventions do not last longer than 2 years). To check the importance of this assumption, we consider alternative maximal intervention durations of three and four years below.

In our baseline specification, we use data covering 4 months before and 4 months after the birthday cutoff date. Thus, we include all preschoolers born between September 1, 1999 and April 30, 2000. This results in a sample size of $N = 4788$.

[11] According to Austrian Law, children enter primary school in a given year if they turn 6 years of age before September 1 of that year. Children who turn 6 years old thereafter have to delay enrollment by one year (Zweimüller, 2014).

[12] We exclude from the analysis children with extremely high expenditure for medication (above the 99.5 percentile in short- or long-term expenditure).

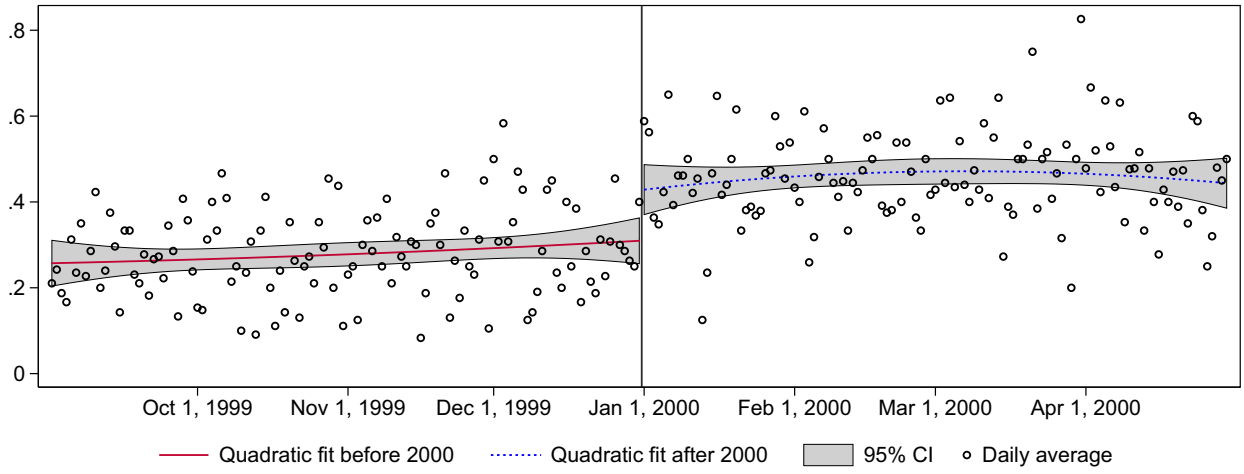

Fig. 1. Participation rate in developmental screenings per day of birth.

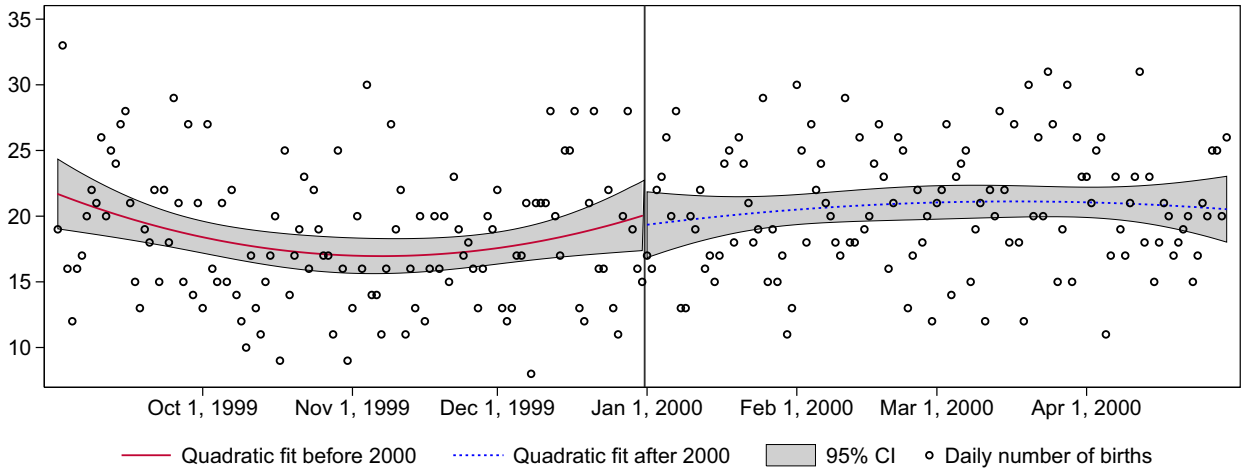

Fig. 2. Number of births per day.

3.3.1. Identifying assumptions

Three conditions need to hold for $\hat{\beta}_1$ to be informative about the effects of screening participation. First, the eligibility for the financial incentive T_i must predict participation in the screening. This condition is testable. Fig. 1 shows the first-stage relationship. Using birthdate, it plots the average share of preschoolers who have undergone all three screening examinations (referring to the binary treatment variable S_i^3 in our main specification). As expected, we observe a distinctive jump in the participation rate at the cutoff that can be attributed to the eligibility for the financial incentive. In other words, the probability of treatment is significantly higher for all eligible preschoolers. We show below that this condition also holds in a regression framework. Second, families do not precisely manipulate their children's dates of birth around the eligibility cutoff. This is the key identifying assumption behind any RDD. Since the Upper Austrian government enacted the financial incentive in retrospect on November 12, 2001, this assumption holds by definition. Accordingly, Fig. 2 shows that the average number of births per day does not vary around the cutoff date.[13]

Third, eligibility must not be correlated with any outcome-determining factor. We start by examining some observable characteristics. Fig. 3 shows that parity, legitimacy, mother's citizenship, and mother's educational attainment do not change discontinuously around the cutoff. The same holds for any other predetermined characteristics we observe. Thus, we again have no reason to expect a correlation between eligibility and any unobserved outcome-determining factor (included in μ_i).

We estimate two alternative specifications. First, we estimate equations (1) and (2) covering the 4 months before and 4 months after the cutoff. Second, we pool information from the previous turn of the year to account for any unobserved characteristics that follow a seasonal pattern.[14] Thus, we use information on all births between September 1998 and April 1999, and between September 1999 and April 2000, resulting in 9516 observations. Then, we include a series of binary indicators $M_{i,j}$ that capture the calendar month j of the child's birth to control for month fixed effects:

[13] More formally, the density-based test suggested by McCrary (2008) confirms this. We cannot reject the hypothesis that there is a shift in the discontinuity at the birthday cutoff: test statistic = 0.023, standard error = 0.102 (bin size = 1, default bandwidth calculation, bandwidth = 47.670).

[14] There is some evidence in the US that children born at different times of the year are born to mothers with significantly different characteristics (Buckles and Hungerman, 2013). In fact, seasonality in unobserved characteristics would constitute a threat to our identification only if unobserved outcome-determining factors varied discontinuously near the cutoff (December vs. January).

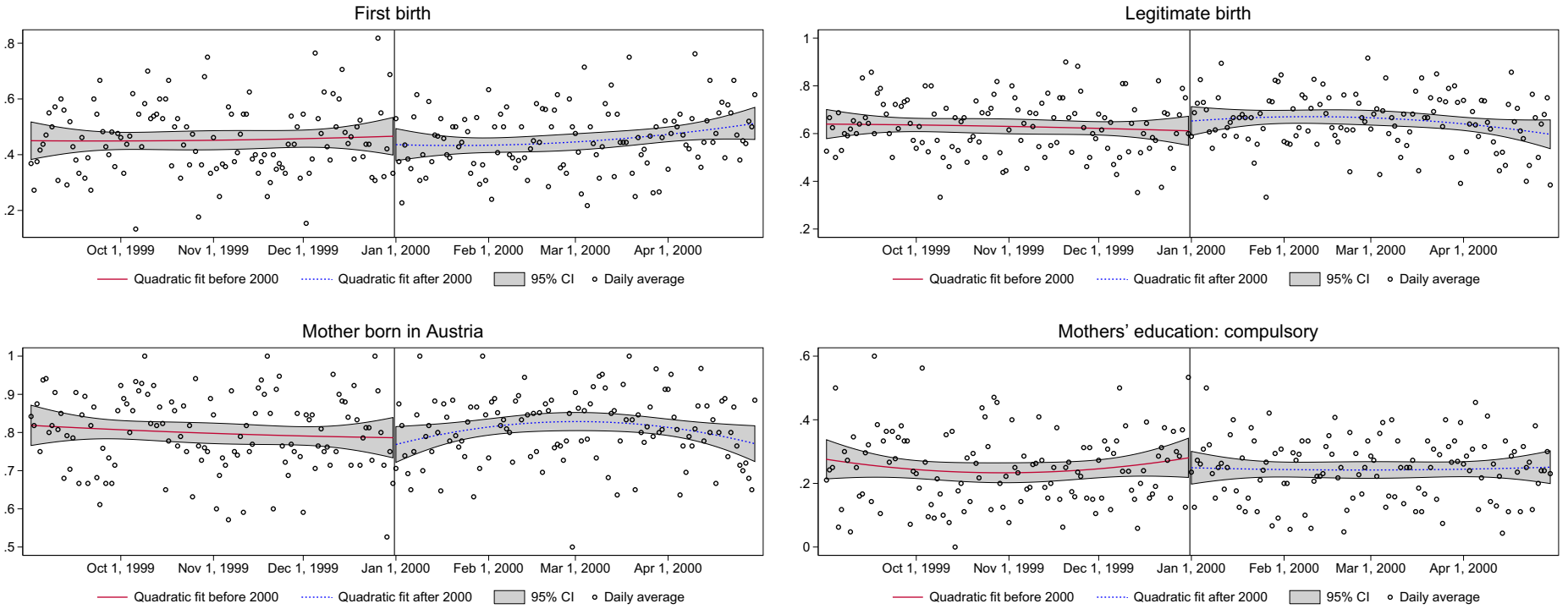

Fig. 3. Daily averages of selected covariates.

Table 2
Effect of the financial incentive on participation.

Dependent variable	Quadratic trends (see eq. (1))			Month fixed effects (see eq. (3))		
	S_i^3	$S_i^{2,3}$	S_i^{abs}	S_i^3	$S_i^{2,3}$	S_i^{abs}
	(1)	(2)	(3)	(4)	(5)	(6)
Treated, T_i	0.123***	0.142***	0.312***	0.145***	0.139***	0.358***
	(0.040)	(0.042)	(0.100)	(0.018)	(0.020)	(0.047)
Linear trend before, $B_i - c$	0.000	–0.000	0.002			
	(0.001)	(0.001)	(0.003)			
Linear trend after, $(B_i - c) \times T_i$	0.001	0.001	0.001			
	(0.002)	(0.002)	(0.004)			
Quad. trend before, $(B_i - c)^2$	0.000	0.000	0.000			
	(0.000)	(0.000)	(0.000)			
Quad. trend after, $(B_i - c)^2 \times T_i$	–0.000	–0.000	–0.000			
	(0.000)	(0.000)	(0.000)			
Period fixed effect, P_i				0.051***	0.051***	0.164***
				(0.013)	(0.014)	(0.033)
Month fixed effect, $M_{i,j}$	No	No	No	Yes	Yes	Yes
Further covariates						
First birth	0.160***	0.195***	0.518***	0.170***	0.211***	0.568***
	(0.016)	(0.016)	(0.037)	(0.011)	(0.011)	(0.027)
Preterm birth	0.027	0.078**	0.135	0.007	0.039	0.050
	(0.039)	(0.039)	(0.096)	(0.026)	(0.027)	(0.066)
Low birth weight	–0.002	0.013	0.021	0.012	0.005	0.006
	(0.037)	(0.036)	(0.089)	(0.025)	(0.026)	(0.064)
Mother's age	0.001	0.002	0.005	0.002**	0.003**	0.005*
	(0.001)	(0.002)	(0.004)	(0.001)	(0.001)	(0.003)
Legitimate birth	0.034**	0.035**	0.068*	0.030***	0.026**	0.053*
	(0.016)	(0.016)	(0.038)	(0.011)	(0.011)	(0.027)
Mother born in Austria	0.093***	0.068***	0.222***	0.071***	0.065***	0.195***
	(0.019)	(0.020)	(0.048)	(0.012)	(0.014)	(0.033)
Lives in city	0.031*	0.019	0.078*	0.039***	0.061***	0.146***
	(0.017)	(0.017)	(0.041)	(0.011)	(0.012)	(0.029)
Mother's educational attainment (base group: compulsory school)						
Vocational/lower sec.	0.050***	0.075***	0.180***	0.049***	0.056***	0.154***
	(0.017)	(0.018)	(0.044)	(0.011)	(0.013)	(0.030)
Upper sec./tertiary	0.051**	0.080***	0.183***	0.051***	0.068***	0.168***
	(0.022)	(0.023)	(0.055)	(0.015)	(0.016)	(0.039)
Constant	0.067	0.180***	0.772***	–0.000	0.140***	0.604***
	(0.053)	(0.056)	(0.134)	(0.033)	(0.037)	(0.087)
Preschoolers are born	09/1999–04/2000			09/98–04/99 and 09/99–04/00		
Number of observations	4788	4788	4788	9516	9516	9516
Mean of dependent variable	0.378	0.563	1.700	0.313	0.499	1.519
F statistic on T_i	9.2	11.7	9.8	62.9	49.9	57.9

Notes: This table summarizes estimation results on the effect of the eligibility for the financial incentive on developmental screening participation based on alternative specifications. The first three columns are based on the model described by eq. (1), which uses the sample of all children born between September 1999 and April 2000. The remaining three columns are based on the model described by eq. (3), which uses the sample of all children born between September 1998 and April 1999 and September 1999 and April 2000. In the first and fourth columns, the dependent variable is a binary indicator equal to one if child i has participated in all three developmental screening examinations, and zero otherwise. In the second and fifth columns, the dependent variable is a binary indicator equal to one if child i has participated in at least two developmental screening examinations, and zero otherwise. In the third and sixth columns, the dependent variable captures the absolute number of developmental screening examinations in which child i participated. The method of estimation is least squares and robust standard errors are reported in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

$$S_i^3 = \gamma_0 + \gamma_1 T_i + \sum_{j=2}^{8} \gamma_j M_{i,j} + \gamma_9 P_i + \Gamma \mathbf{X}_i + \eta_i. \quad (3)$$

$$O_i = \delta_0 + \delta_1 \hat{S}_i^3 + \sum_{j=2}^{8} \delta_j M_{i,j} + \delta_9 P_i + \Delta \mathbf{X}_i + \nu_i. \quad (4)$$

In addition, the binary variable P_i captures whether the preschooler i was born between September 1998 and April 1999 or between September 1999 and April 2000.

Moreover, we replicate all estimations with alternative treatment measurements. First, we define a binary indicator $S_i^{2,3}$ equal to one if the preschooler has participated in at least two of the screening examinations. Second, we define S_i^{abs} (where $S_i^{abs} \in \{0, 1, 2, 3\}$), capturing the absolute number of screening examinations in which the preschooler has participated.

4. Estimation results

We present our estimation results in three steps. First, we discuss the effects of the eligibility for financial incentives on screening participation. Second, we examine the effects of participation on the scope of intervention. Third, we consider the effect of participation on long-run healthcare costs. We conclude this section with an overall assessment of our results.

4.1. Effect of financial incentives on participation

The regression results summarized in Table 2 fully confirm the results suggested by Fig. 1. Across all specifications, we find that the probability of treatment is significantly higher for eligible preschoolers (i.e., all children born on January 1, 2000 or later). Columns (1) to (3) are based on equation (1), which allows for different quadratic trends before and after the cutoff date. The

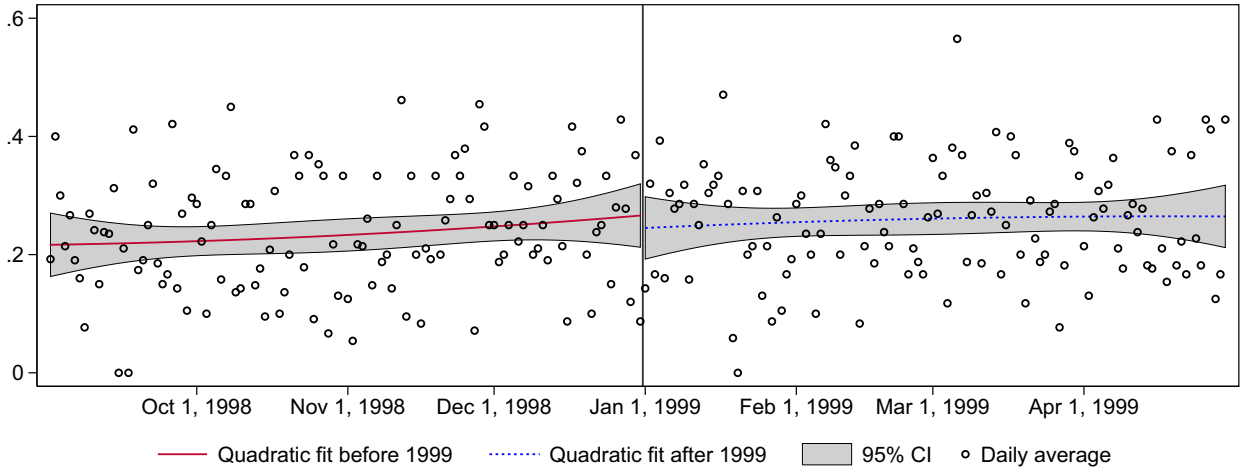

Fig. 4. Placebo check: Screening participation rate around January 1, 1999.

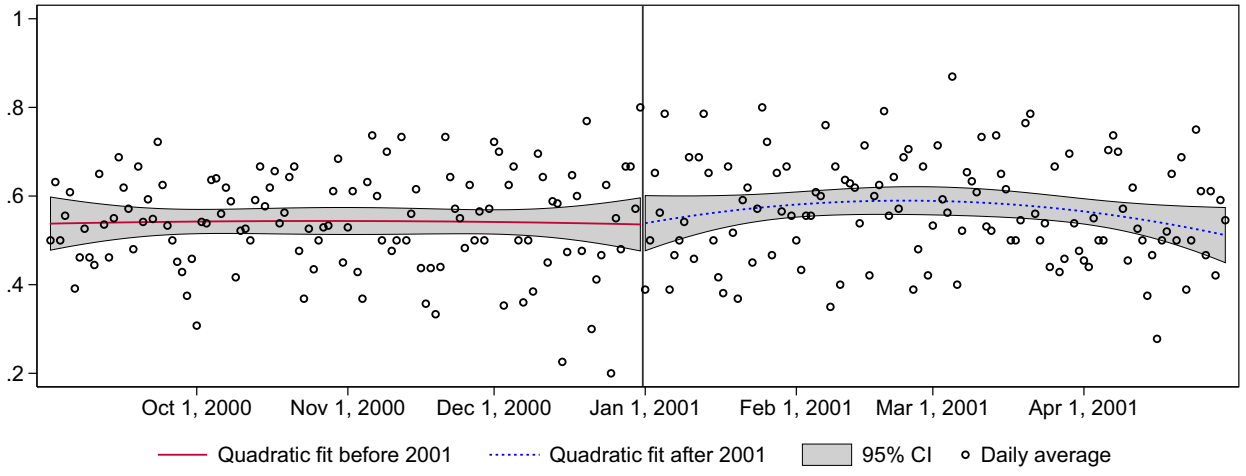

Fig. 5. Placebo check: Screening participation rate around January 1, 2001.

coefficients indicate that the eligibility for the financial incentive (T_i) increases the likelihood of participating in all three examinations by 12.3 percentage points, the probability of participation in two or more examinations by 14.2 percentage points, and the absolute number of performed exams by 0.31. Columns (4) to (6) are based on equation (3), which uses additional observations from children born 1 year earlier to control for birth month fixed effects. The estimated coefficients are very comparable and amount to 14.5 and 13.9 percentage points, respectively, and 0.36 additional screenings.[15] For our further analysis, we use the semi-parametric specification with month fixed effects. Since we obtain in this model large *F*-statistics for T_i (of at least 50), we can abstract from weak instrumental variable problems (Staiger and Stock, 1997).[16]

[15] First-stage relationships for participation in at least two screenings ($S_i^{2,3}$) and the number of screening examinations (S_i^{abs}) are illustrated in Figs. A.1 and A.2 in Appendix S1.

[16] Notably, all estimated coefficients remain essentially unchanged if we drop all further covariates. The estimated coefficients of these simple ordinary least square regressions are as follows: (1) 0.121, (2) 0.141, (3) 0.306, (4) 0.149, (5) 0.145, and (6) 0.373. This supports the presumption that our instrumental variable T_i is as good as randomly assigned.

4.1.1. Placebo tests

To underpin the validity of our identification strategy, we offer two different placebo tests. First, we show the participation rates in developmental screenings in Upper Austria in the year before and after the program cutoff date. As observed from Figs. 4 and 5, there is no discontinuity in participation rates in Upper Austria before and after January 1 in both the pre- and post-treatment year. Second, we provide evidence for potential participation effects based on data from the *Vorarlberg Health Insurance Fund*, which represents the equivalent institution to the *Upper Austrian Health Insurance Fund* for the federal state of Vorarlberg. The two health insurance funds operate under very similar conditions, for example, they offer the same MCPEP. However, the state of Vorarlberg has never offered a financial incentive for participation in the developmental screening part of the MCPEP (i.e., examinations D1, D2, and D3). Thus, we use Vorarlberg data to analyze any potential change in the screening participation rate around the birthdate cutoff, which determines eligibility for the financial incentive in Upper Austria. The only restriction we face with these data is that we have to focus on the participation in examination D3. Given that the available data from the *Vorarlberg Health Insurance Fund* cover only the period from 2003 to 2007, preschoolers are already 3 years and older, and we cannot observe their participation in the examinations D1 and D2. Consequently, we replicate our first-stage analysis

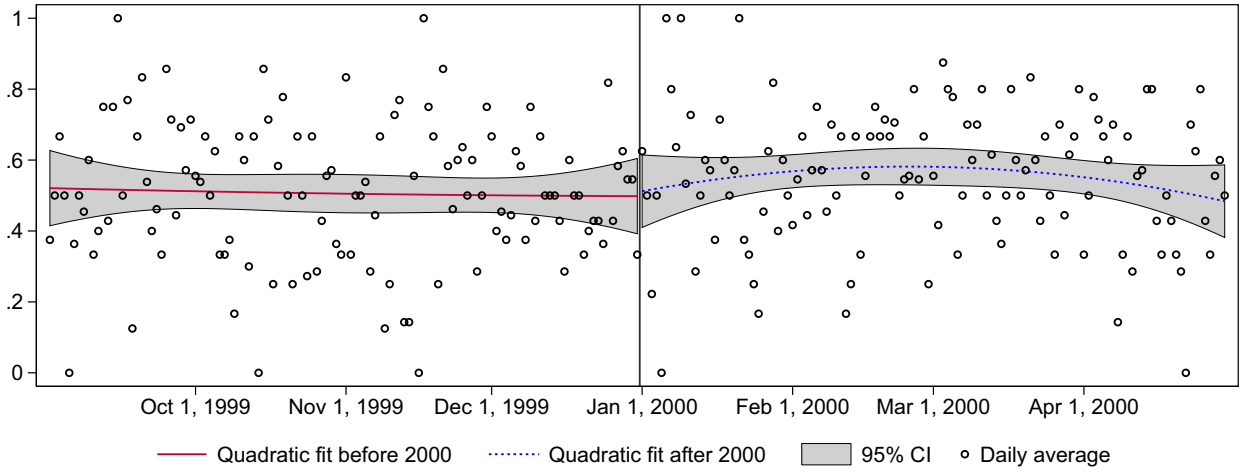

Fig. 6. Placebo check: Participation rate in developmental screening D3 per day of birth in Vorarlberg.

with Vorarlberg data for participation in D3. As expected, this placebo first stage shows no significant different screening likelihood for children born after January 1, 2000. This is depicted graphically in Fig. 6 and is observed within the regression framework summarized in column (1) of Table A.1 in Appendix S1. By contrast, the comparison with the equivalent analysis based on Upper Austrian data reveals a significant effect (see Fig. 7 and column (2) of Table A.1 in Appendix S1). The placebo test fully supports the assumption that no other national shock is driving our first-stage result based on the data from the *Upper Austrian Health Insurance Fund*.

4.1.2. Heterogeneity in participation

To study whether families with different characteristics react differently to the financial incentive, we stratify our sample by different important dimensions measured at the time of birth. Most importantly, we use information on mother's educational attainment to proxy for the family's SES. We distinguish between low SES preschoolers (the mother has compulsory schooling or less, 25 percent) and higher SES preschoolers (the mother has any degree higher than compulsory schooling, 75 percent). Further dimensions of sample splits are the mother's country of birth (Austria vs. foreign country), legitimacy of birth (mother is married vs. unmarried), parity (first birth vs. higher-order birth), child's sex, gestational length (full-term vs. pre-term), and birth weight (normal vs. low). Table 3 summarizes basic descriptive statistics and estimation results based on the month fixed-effects specification (3) using S_i^3 as a dependent variable. With two exceptions, we find very comparable effects of the financial incentive on screening participation between 12 and 16 percentage points.

Foreign-born mothers are one group that reacted to a lesser extent. Among them, their eligible preschoolers were only 7 percentage points more likely to participate. A lack of language proficiency and institutional knowledge are plausible explanations for this finding. Families with children who had problematic birth outcomes are substantially more likely to respond to the financial incentive. Eligible children with a low birth weight have an increased propensity to participate of 22 percentage points, while children with a birth weight of 2500 grams or more are only 14 percentage points more likely to be screened. Very comparable effects are observed for pre- versus full-term births. Interestingly, a similar pattern to the former two sample splits is obtained by splitting the sample at the 25th percentile of the birth-weight distribution. This sample split has the advantage that both sub-samples have a reasonably high number of observations, and consequently a higher *F*-statistic. With respect to the child's sex, we see only small differences. Eligible families with sons are marginally more likely to react to the financial incentive compared to their counterparts with daughters (16 vs. 13 percentage points).

Somewhat surprisingly, we find no difference across families with low and high SES. A potential explanation for this invariance could

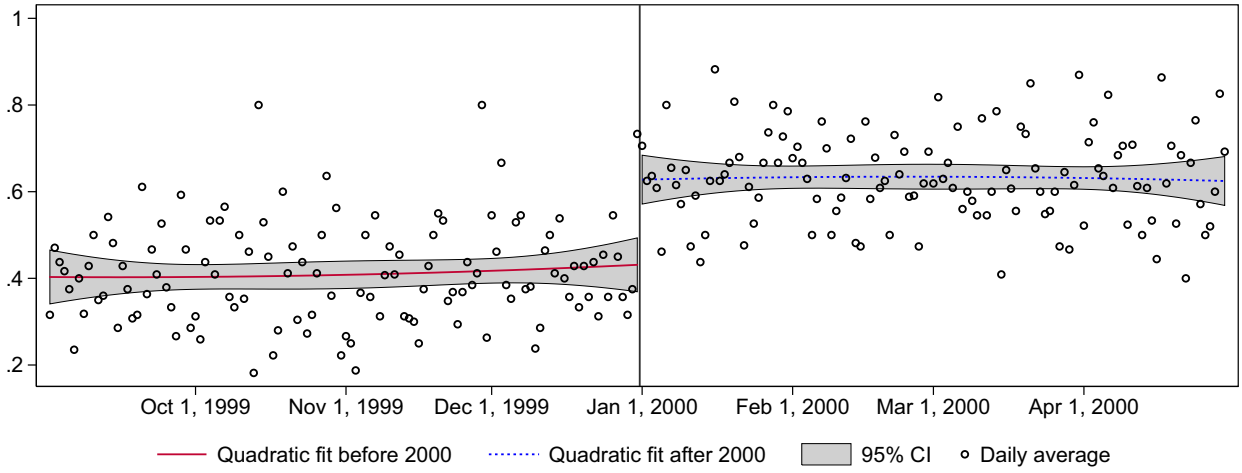

Fig. 7. Participation rate in developmental screening D3 per day of birth in Upper Austria.

Table 3
Heterogeneity in the effect of the financial incentive on participation.

	(1)	(2)	(3)	(4)	(5)
	N	Mean	Estimate	S.E.	*F*-stat.
Full sample	9516	0.313	0.145***	(0.018)	62.9
Socioeconomic status[a]					
Low SES	2346	0.242	0.145***	(0.034)	18.3
Higher SES	7170	0.336	0.143***	(0.022)	44.0
Mother's country of birth					
Foreign country	1810	0.238	0.070*	(0.039)	3.3
Austria	7706	0.330	0.162***	(0.021)	61.7
Legitimacy of birth					
Out of wedlock	3334	0.339	0.136***	(0.032)	18.3
Legitimate	6182	0.298	0.149***	(0.022)	44.7
Parity					
First birth	4252	0.404	0.132***	(0.029)	20.1
Higher order birth	5264	0.239	0.155***	(0.023)	46.4
Sex of child					
Boy	4847	0.312	0.159***	(0.025)	39.0
Girl	4669	0.314	0.126***	(0.026)	23.0
Gestation length[b]					
Pre-term birth	489	0.325	0.236**	(0.081)	8.5
Full-term birth	9027	0.312	0.138***	(0.019)	54.5
Birth weight[c]					
Low	520	0.346	0.222**	(0.077)	8.3
Normal	8996	0.311	0.140***	(0.019)	55.6
Below 25th percentile	2352	0.315	0.214***	(0.037)	33.4
Above 25th percentile	7164	0.312	0.123***	(0.021)	34.6

Notes: This table summarizes estimation results on the effect of the eligibility for the financial incentive on developmental screening participation for different sample splits. The estimations are equivalent to those presented in column (4) of Table 2 (i.e., the dependent variable is a binary indicator equal to one if child *i* has participated in all three developmental screening examinations, and zero otherwise; and the estimations follow the model described by eq. (3)). The method of estimation is least squares. Column (1) reports the number of observations, column (2) reports the mean of the dependent variable, column (3) reports the estimated coefficient, column (4) reports robust standard errors (* $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$), and column (5) reports the *F* statistic on T_i (test for a weak instrument).

[a] The preschooler's socioeconomic status (SES) is defined according to his or her mother's educational attainment at birth. If the mother's highest degree is compulsory schooling, then the preschooler is defined as low SES.

[b] A pre-term birth is defined by gestational length of 37 weeks or less.

[c] Low birth weight children weigh less than 2500 grams.

be that we observe the sum of countervailing effects. In order to respond to the financial incentive, parents have to (i) know and understand the institutional setting, (ii) assess the monetary value as sufficient in principle, and (iii) schedule and keep appointments for the screening procedures. It is possible that low SES families, while being on average more motivated by the monetary value, are comparably less informed and/or less good about keeping an appointment. Thus, low and high SES families may respond equally to the financial incentive, despite having a different *ex ante* evaluation of the monetary value.[17]

4.2. Effect of participation on scope of intervention

Our estimation results of the effect of the developmental screening on the scope of the intervention are summarized in Table 4.[18] The outcome variables are aggregated over children's age period from 18 to 54 months and capture the costs for follow-up treatment, excluding direct costs of the developmental screening. We present results for the overall sample, and for the two sub-samples of low and high SES preschoolers. It turns out that this is the only sample split that provides heterogeneous second-stage results. In the upper panel, we summarize the estimated effects on broad categories of healthcare expenditure: outpatient medical care, medication, and days in hospital. The lower panel provides results for expenditure for healthcare services at GPs and different types of outpatient medical specialists. For the overall sample, we list the reduced-form estimate (the *intention-to-treat effect*, ITT) and the second-stage estimate (the *local average treatment effect*, LATE). For the two sub-samples, we focus on the LATE. In each case, we list the sample mean of the respective outcome variable.

The first three columns of Table 4 show that for the overall sample, screening participation does not trigger a statistically significant increase in aggregate follow-up expenditure for medical attendance in the outpatient sector. However, the disaggregated analysis in the lower panel reveals significantly positive effects on follow-up treatments provided by pediatricians and ophthalmologists. Pediatricians and ophthalmologists charge €77.2 and €20.6, respectively, for additional medical treatment following the developmental screening procedure, which correspond to approximately 77 percent and more than 100 percent, respectively, of the average expenditure. The statistically insignificant coefficients of expenditure for other medical specialists indicate that screening participation does not trigger follow-up referrals to doctors in other medical fields. A first conclusion to be drawn from these short-term estimation results is that participation in developmental screening causes subsequent medical treatment of preschoolers by pediatricians and ophthalmologists.[19] The insignificant (negative) effect on GPs' subsequent medical services indicates that doctors in this group do not conduct their own additional treatment. If at all, follow-up costs borne by this group of doctors arise exclusively from subsequent referrals to pediatricians and ophthalmologists.

4.2.1. Heterogeneous results

The remainder of Table 4 summarizes the results for preschoolers with low and higher SES separately. This comparison reveals that the effects of screening on interventions are driven mainly by low SES preschoolers. Expenditure for outpatient medical care increases by as much as €394.7 (+149 percent) in this group. The disaggregated analysis reveals significant effects for pediatricians (€237.4) and ear, nose, and throat (ENT) specialists (€37.3). A weakly statistically significant effect is observed for radiologists. We conclude that screening doctors diagnose developmental disorders in low SES preschoolers, which leads to substantial further medical treatment. In addition, there is a weakly statistically significant effect on hospitalization; however, we find no significant effect on the consumption of medical drugs.

By comparison, the only significant increase in follow-up expenditure for high SES preschoolers is for ophthalmologists. Participation in developmental screening triggers additional expenditure for ophthalmologists in the amount of €24.9. Furthermore, participation in developmental screening is estimated to reduce expenditure for medical drugs by €72.1. A potential explanation is that families with higher SES may reduce or even stop unnecessary medication for their children as a consequence of doctor's consultation during the screening exam.[20]

[17] Figs. A.3 to A.8 in Appendix S1 illustrate participation in all three developmental screening examinations (S_i^3), participation in at least two screenings ($S_i^{2,3}$), and the number of screening examinations (S_i^{abs}) for high and low SES families.

[18] We focus on second-stage estimation results based on the month fixed-effects specification (see equations (3) and (4)). The results based on the quadratic trends specification (see equations (1) and (2)) are very comparable and detailed estimation output is available upon request.

[19] The ITT estimates are qualitatively identical to the LATE results. However, the coefficients are substantially lower due to imperfect compliance.

[20] In Appendix S1, we summarize estimation results for the models using the two alternative specifications of the treatment variable. See Table A.2 in Appendix S1 for the specification using participation in two or more screenings ($S_i^{2,3}$) and Table A.3 in Appendix S1 for the specification using the absolute number of screenings (S_i^{abs}). Both alternative specifications provide qualitatively identical and quantitatively very comparable results.

Table 4
Effect of participation on scope of intervention.

	Full sample of preschoolers			Low SES preschoolers		High SES preschoolers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	ITT	LATE	Mean	LATE	Mean	LATE
Aggregate health spending categories between 18 and 54 months of age							
Outpatient medical care (in euro)	291.1	11.5	79.8	265.3	394.7***	299.5	–32.5
		(10.0)	(66.1)		(129.7)		(83.5)
Medication (in euro)	64.8	–6.3	–43.8	63.5	24.5	65.3	–72.1**
		(4.0)	(29.3)		(55.8)		(35.8)
Hospitalization (in days)	1.9	0.3	2.1	2.6	10.0*	1.7	–0.6
		(0.2)	(1.6)		(5.6)		(1.3)
Spending on medical specialists between 18 and 54 months of age in euro							
Pediatrician	100.3	11.2**	77.2**	91.9	237.4***	103.0	24.0
		(5.3)	(33.8)		(71.4)		(41.7)
General practitioner	128.9	–4.5	–31.4	125.0	59.3	130.1	–65.1
		(5.4)	(38.1)		(72.3)		(46.3)
Ophthalmologist	18.9	3.0**	20.6**	15.7	4.4	20.0	24.9**
		(1.3)	(8.8)		(16.8)		(10.6)
ENT specialist	12.0	1.0	7.0	8.4	37.3**	13.2	–4.1
		(1.6)	(11.0)		(18.1)		(13.9)
Orthopedist	3.1	0.9*	6.2	2.3	10.0	3.4	4.9
		(0.5)	(3.8)		(6.7)		(4.6)
Laboratory	5.5	–0.1	–0.7	4.1	3.7	5.9	–2.7
		(0.9)	(5.9)		(9.5)		(7.4)
Radiologist	3.1	0.4	2.5	2.3	12.9*	3.3	–1.2
		(0.6)	(4.1)		(7.4)		(5.0)
Number of observations	9516			2346		7170	

Notes: This table summarizes estimation results of the effect of screening participation on the scope of intervention captured by different health spending categories aggregated over the children's age range between 18 and 54 months. Each entry reflects a separate regression, in which the dependent variable is indicated on the very left. The upper panel summarizes results for aggregate health spending categories, and the lower panel summarizes results for spending on medical specialists. The method of estimation is *two-stage least squares* based on the model described by eqs. (3) and (4), which includes covariates listed in Table 2. Robust standard errors are reported in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Columns (1) to (3) provide the mean of the dependent variable, the *intention-to-treat effect* (ITT, based on the reduced form) and the *local average treatment effect* (LATE, based on the second stage) for the full sample. Columns (4) and (5) provide the mean of the dependent variable and the LATE for the sample of low socioeconomic status (SES) preschoolers. Columns (6) and (7) provide the mean of the dependent variable and the LATE for the sample of high SES preschoolers. The preschooler's SES is defined according to his or her mother's educational attainment at birth. If the mother's highest degree is compulsory schooling, then the preschooler is defined as low SES.

In the robustness analysis, we split the high SES group into two groups. The first group consists of mothers with vocational education ($N = 3972$). The second group comprises mothers who at least attended any type of secondary school ($N = 3198$). The estimation results for the three resulting groups ((i) compulsory education, (ii) vocational education, and (iii) secondary and tertiary education) provide two additional insights compared to our baseline split. First, the increase in expenditure for ophthalmologists is driven by the group with vocational education and is not significant for families with secondary or tertiary education. Second, we observe a substitution of service utilization between GPs and pediatricians for the group with the highest education. For treated preschoolers, services provided by pediatricians increase, while GP expenditure decreases. Full estimation output for the alternative split by SES is provided in columns (1), (3) and (5) in Table A.4 in Appendix S1.

4.2.2. Mothers and siblings

It is possible that other family members change their healthcare utilization behavior due to the screening participation of the index child. Therefore, we analyze the healthcare expenditure of the preschoolers' mothers and siblings. We identify 7100 mothers and approximately 3400 older siblings of children in our main sample.[21] We measure their healthcare expenditure during the intervention period of the index child, who is between 18 and 54 months old during this period. The estimates reveal a positive effect of screening participation on low SES mothers' expenditure for GP and laboratory services over this period. Similarly, we find a positive impact of developmental screening on expenditure for outpatient medical care for siblings in low SES families. This increase is driven by expenditure for pediatricians. In high SES families, we find no statistically significant effects for siblings and mothers. The developmental screening participation of low SES preschoolers leads to further medical interventions for these children but also triggers additional expenditure for medical attendance in their mothers and siblings. The full estimation output for mothers and siblings is summarized in Tables A.5 and A.6 in Appendix S1.[22]

4.3. Effect of participation on long-run healthcare costs

Our estimation results of the effect of developmental screening on long-run healthcare costs are summarized in Table 5. The different expenditure categories are the sum over the period when children are between 6 and 10.5 years of age. Column (3) shows that on average, developmental screening participation has no impact on long-term healthcare costs. This holds for the inpatient and outpatient sectors. As a result, it can be concluded that the increase in follow-up expenditure for services provided by pediatricians and ophthalmologists cannot be compensated by a reduction in long-term diagnostic and/or therapeutic services. The separate analysis

[21] We restrict the analysis to the next older sibling. Younger siblings may be conceived post treatment, and as such are potentially endogenous. On average, the siblings are 3.26 years older than the children in the main analysis.

[22] Finally, we examine whether the labor market behavior of mothers of treated children is affected. Estimates of a linear probability model for the likelihood of the mother's employment 0,1,2, … years after the child's first developmental screening (D1) indicate no significant effect (see Table A.7 in Appendix S1 for the detailed regression output).

Table 5
Effect of participation on long-run healthcare costs.

	Full sample of preschoolers			Low SES preschoolers		High SES preschoolers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	ITT	LATE	Mean	LATE	Mean	LATE
Aggregate health spending categories between 6 and 10.5 years of age							
Outpatient medical care (in euro)	969.3	–23.5	–162.1	915.5	–588.7	986.9	–18.7
		(41.2)	(286.9)		(572.2)		(337.5)
Medication (in euro)	151.4	–13.5	–93.2	147.5	–131.2	152.6	–87.7
		(14.1)	(98.1)		(168.9)		(119.6)
Hospitalization (in days)	2.2	0.8	5.3	3.6	8.1	1.8	2.6
		(0.9)	(6.0)		(22.5)		(4.4)
Spending on medical specialists between 6 and 10.5 years of age in euro							
Pediatrician	69.4	1.9	13.4	74.2	37.7	67.8	3.8
		(5.2)	(35.6)		(71.0)		(41.8)
General practitioner	198.1	–10.8*	–74.4	208.3	–171.5*	194.7	–44.4
		(6.5)	(46.0)		(97.7)		(53.2)
Dentist	359.3	1.0	6.9	294.4	–6.1	380.5	25.1
		(22.9)	(158.2)		(259.8)		(195.1)
Ophthalmologist	43.7	0.6	4.2	38.9	–12.5	45.3	9.4
		(2.5)	(17.1)		(34.5)		(20.1)
ENT specialist	28.1	0.1	0.7	25.4	5.6	28.9	–0.8
		(2.8)	(19.4)		(39.0)		(22.5)
Orthopedist	5.3	0.1	0.7	4.5	–0.3	5.5	0.7
		(0.8)	(5.2)		(9.7)		(6.2)
Laboratory	9.7	1.5	10.1	9.2	–5.8	9.8	14.7
		(1.2)	(8.0)		(15.0)		(9.7)
Radiologist	9.0	–1.1	–7.7	9.1	–8.8	9.0	–7.5
		(1.1)	(7.9)		(16.7)		(9.0)
Number of observations	9516			2346		7170	

Notes: This table summarizes estimation results of the effect of screening participation on long-run healthcare costs captured by different health spending categories aggregated over the children's age range between 6 and 10.5 years. Each entry reflects a separate regression, in which the dependent variable is indicated on the very left. The upper panel summarizes results for aggregate health spending categories, and the lower panel summarizes results for spending on medical specialists. The method of estimation is *two-stage least squares* based on the model described by eqs. (3) and (4), which includes covariates listed in Table 2. Robust standard errors are reported in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Columns (1) to (3) provide the mean of the dependent variable, the *intention-to-treat effect* (ITT, based on the reduced form) and the *local average treatment effect* (LATE, based on the second stage) for the full sample. Columns (4) and (5) provide the mean of the dependent variable and the LATE for the sample of low socioeconomic status (SES) preschoolers. Columns (6) and (7) provide the mean of the dependent variable and the LATE for the sample of high SES preschoolers. The preschooler's SES is defined according to his or her mother's educational attainment at birth. If the mother's highest degree is compulsory schooling, then the preschooler is defined as low SES.

for preschoolers from low and higher SES preschoolers broadly supports this conclusion based on the full sample.

The case of high SES preschoolers is quite clear. Given that we find little evidence for subsequent medical interventions, with the exception of expenditure for ophthalmologists, we cannot expect substantial cost savings in the long run. This expectation is consistently supported by our regression results summarized in column (7). The effects on expenditure in the outpatient sector (in the aggregate and, to a lesser extent, for the different medical specialists) are estimated quite precisely and suggest a zero effect. While the effects on medication and hospitalization are estimated less precisely, they do not indicate cost savings in the long run either.

The case of low SES preschoolers is harder to assess. First, since we observe significant interventions, there is at least some potential for future cost savings. Indeed, with few exceptions, we observe negative point estimates in column (5). The only marginally significant effect is a reduction in the expenditure for GPs. Participation is estimated to reduce long-run expenditure for GP treatment by €171.5, or about 82 percent. Both alternative specifications of the treatment variable confirm these long-run results and suggest that for low SES preschoolers at the 10 percent level, there is a significantly negative effect of screening participation on GP expenditure, whereas the effects on all other outcomes remain statistically insignificant.[23] The standard errors of the estimates are, however, large. Even if we do not face a weak instrument problem (the *F*-statistic is 18.3), the large variation in some outcome variables and the smaller sample size for the group of low SES preschoolers have negative impacts on the precision of the instrumental variable estimates. Given these limitations, the results must be interpreted with caution, in particular for the low SES preschoolers.[24]

The significance levels of coefficients are consistent with those from the reduced-form specification (not shown). This specification provides smaller standard errors and confirms the insignificant and negative point estimates of screening participation on long-run healthcare costs for low SES preschoolers. While it is clear that the savings in the long run cannot compensate for the increase in expenditure due to the intervention in the short run, the results provide at least some evidence for cost-saving effects of the screening program in low SES preschoolers.

5. Discussion

For high SES preschoolers, we find a consistent picture with little evidence for interventions and consistently no effect on

[23] Detailed estimation output summarized in Appendix S1 shows that participation in two or more screenings (see Table A.8 in Appendix S1) reduces long-term GP expenditure by €125.1, and participation in one additional screening reduces the same expenditure category by €48.6 (see Table A.9 in Appendix S1).

[24] To check the sensitivity of our results with respect to the assumption of a maximum intervention duration of 2 years, we re-calculate our results with the alternative assumptions of 3 and 4 years. This corresponds with aggregating the long-run healthcare costs between 7 and 10.5 years of age, and between 8 and 10.5 years of age. Tables A.10 and A.11 in Appendix S1 summarize the corresponding results. The only notable difference observed is that under the assumption of a maximum duration of 4 years (vs. 2 or 3 years), the reduction in the expenditure for GPs among low SES preschoolers becomes more significant.

Table 6
Effect of participation on scope of intervention: Expenditure borne by referrals.

	Full sample of preschoolers			Low SES preschoolers		High SES preschoolers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	ITT	LATE	Mean	LATE	Mean	LATE
Spending on medical specialists due to referrals between 18 and 54 months of age in euro							
Pediatrician	4.8	–0.3	–2.1	7.0	7.3	4.0	–5.5
		(0.9)	(6.2)		(16.9)		(6.3)
General practitioner	4.7	–1.0	–6.6	4.0	0.5	5.0	–9.2*
		(0.6)	(4.3)		(7.0)		(5.3)
Ophthalmologist	5.4	0.9	6.4	6.4	–0.2	5.1	8.1
		(0.7)	(4.5)		(10.4)		(5.1)
ENT specialist	5.1	1.8**	12.3**	4.7	19.3*	5.3	9.7
		(0.8)	(5.8)		(11.1)		(6.9)
Orthopedist	1.3	0.1	0.4	1.1	4.0	1.4	–0.8
		(0.3)	(2.2)		(4.2)		(2.7)
Laboratory	5.1	0.1	0.5	3.8	5.7	5.5	–1.8
		(0.8)	(5.7)		(9.1)		(7.2)
Radiologist	2.9	0.2	1.4	2.3	12.5*	3.1	–2.5
		(0.6)	(4.0)		(7.3)		(4.9)
No. of observations	9516			2346		7170	

Notes: This table summarizes estimation results of the effect of screening participation on the scope of intervention captured by spending on medical specialists borne by referrals and aggregated over the children's age range between 18 and 54 months. Each entry reflects a separate regression, in which the dependent variable is indicated on the very left. The method of estimation is *two-stage least squares* based on the model described by eqs. (3) and (4) and robust standard errors are reported in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Columns (1) to (3) provide the mean of the dependent variable, the *intention-to-treat effect* (ITT, based on the reduced form) and the *local average treatment effect* (LATE, based on the second stage) for the full sample. Columns (4) and (5) provide the mean of the dependent variable and the LATE for the sample of low socioeconomic status (SES) preschoolers. Columns (6) and (7) provide the mean of the dependent variable and the LATE for the sample of high SES preschoolers. The preschooler's SES is defined according to his or her mother's educational attainment at birth. If the mother's highest degree is compulsory schooling, then the preschooler is defined as low SES.

healthcare costs in the long run. There are two equally plausible explanations for this finding, with the same broad conclusion. First, very few high SES preschoolers have developmental disorders that would require an intervention. Alternatively, developmental disorders among high SES preschoolers are identified in the counterfactual situation without screening participation. In this case, observant parents identify disorders themselves and consult medical specialists proactively. Furthermore, these parents might consult a pediatrician regularly, irrespective of any screening program, and the doctor might identify a developmental disorder during these consultations. In either case, we conclude that the developmental screenings are not cost-effective for this group.

The results for low SES preschoolers are more difficult to assess. We find clear evidence for interventions, but only weak evidence for cost savings in the long run. We identify two competing interpretations for this estimation result. First, screening doctors identify developmental disorders in a significant share of low SES preschoolers. However, the subsequent intervention does not dampen future healthcare costs. Nevertheless, the treated preschoolers might still benefit from the intervention. An alternative explanation relates to supply-induced or at least supply-determined follow-up healthcare services. In other words, screening doctors do *not* identify developmental disorders, but recommend additional diagnostic and therapeutic services to raise their income. The fact that the increase in follow-up treatment almost exclusively occurs in low SES preschoolers might indicate that when parents' level of health literacy is lower, physicians' efforts to sell additional unnecessary healthcare services are more successful. To distinguish between these two explanations, we provide further empirical analysis as outlined in the following paragraphs.

5.1. Supply inducement

Follow-up treatments are provided by either the screening doctor him- or herself or another outpatient care doctor. In the latter case, we can distinguish between cases in which parents consulted the non-screening doctor, first, owing to a referral by the screening (or any other) doctor, and second, due to their own initiative. Whereas there is a clear economic incentive for a doctor to schedule in-office follow-up treatments for a patient, the referral of a patient to any other doctor is not refunded by the health insurance fund. We assume there are no informal side payments for the referring doctor. Thus, "referred" interventions cannot represent any form of supply inducement and it seems safe to assume that the medical indication is the main reason for referring a preschooler. If we find an effect of screening participation on follow-up expenditure borne by referrals, we can unequivocally interpret this as true interventions. By contrast, from a zero effect on referrals, we cannot conclude that increases in follow-up expenditure are due to supply inducement.

Following this logic, we re-perform our regression analysis on the effects of participation on the scope of intervention (depicted in Table 4) based on a new dependent variable that captures only follow-up expenditure due to "referred" interventions. The new results are summarized in Table 6. A comparison of the means in the first column of these two tables shows which proportion of total spending on the respective medical specialists is triggered by referrals from other doctors.[25]

The estimation coefficients in Table 6 indicate that screening participation has a significant (*t*-statistics are about 1.7) and positive impact on "referred interventions" to ENT specialists and radiologists. As before, these effects are mainly driven by the group of low SES preschoolers. For this group, participation leads to a fourfold and fivefold increase in medically justified expenditure for ENT and radiological services, respectively. This supports our interpretation that screening doctors identify developmental disorders in low SES preschoolers and provide/initiate interventions.[26]

[25] In the case of the pediatrician, this proportion is very low (4.8/100.3 = 0.05), while for ophthalmologists (0.29) and ENT specialists (0.43) it is substantial, and for laboratories (0.93) and radiologists (0.94) it comprises the vast majority.

[26] The coefficients for pediatricians are economically and statistically insignificant. This suggests that GPs do not refer screened children to pediatricians for further interventions. The comparison with the highly significant effect for pediatricians in Table 4 (see column 5) indicates that screening pediatricians create their own revenues. In the absence of information on preschoolers' health status, we cannot evaluate whether these revenues are due to justified interventions or are supply induced.

6. Conclusions

We evaluated the cost-saving potential of a nationwide physician-based developmental screening program for preschoolers comprising examinations at the ages of 24, 36, and 48 months. Identification was based on a fuzzy RDD that exploited a sharp birthday cutoff-based discontinuity in the eligibility for a financial incentive for program participation. Families with low and higher SES responded equally to this incentive, and eligible children were 14.5 percentage points more likely to be screened.

For high SES preschoolers, we found a consistent picture with little evidence for interventions and consistently no effect on healthcare costs in the long run. Thus, high SES preschoolers have either little to no developmental disorders, or they also receive treatment in the counter-factual situation without screening participation. For low SES preschoolers, we found clear evidence for interventions, but only weak evidence for cost savings in the long run. Thus, low SES preschoolers have identifiable developmental disorders. However, long-run cost savings in the healthcare system cannot compensate for the expenditure on the interventions.

We conclude that this physician-based developmental screening program for preschoolers does not generate significant cost savings, as it is unable to compensate for short-run cost increases in the long run. This is particularly true for preschoolers with high SES. For this group, which comprises about 75 percent of the total preschooler population, we consider the program to be clearly ineffective, since it does not even lead to substantial interventions. The case of low SES preschoolers, who comprise about 25 percent of the total preschooler population, is less clear. We cannot conclude that the program is ineffective for this group. Since we find substantial interventions, the treatment may generate benefits for children without leaving any trace in long-term healthcare expenditure. Unlike screening programs for adults, which focus on early detection and treatment of life-threatening and expensive diseases, developmental screening examinations for children focus on identifying and reducing developmental deficits at an early stage of childhood. An early intervention with respect to these deficits could be expected to improve the conditions for their human capital development, and might not necessarily dampen future healthcare costs. For instance, our finding on the follow-up examinations by ophthalmologists may reflect the early detection of weak eye vision. Given that all children with weak vision will eventually require eyeglasses, early detection may not change future spending much. However, the returns for early detection could still be high if it promotes the acquisition of reading skills and mobility.

Moreover, we found some evidence that the effects of the program are not restricted to the treated children. The participation of low SES preschoolers leads to additional medical interventions in their mothers and siblings. More extensive network and peer effects, such as potential spillovers to classmates, friends, and teachers, cannot be analyzed due to data limitations. However, from an Austrian perspective, the program's reach should not be overrated. Cases of severe developmental disorders, with obvious negative societal externalities, will most likely be treated in the counter-factual situation as soon as they become manifest at school.

Our results suggest that a promising way to improve the accuracy and effectiveness of developmental screenings for preschoolers would be to focus – at least in comparable institutional settings – on subjects with a higher likelihood of otherwise untreated developmental disorders, such as low SES preschoolers. In the case of high SES preschoolers, the program could be reduced to an ophthalmological examination only. These programs should be accompanied by comprehensive evaluation, including direct health outcomes and other important developmental indicators focusing on human capital and social skills. For Austria, at the moment, these additional outcome dimensions cannot be analyzed owing to data restrictions. To continue implementing these programs, and to justify their substantial costs, evidence is needed that shows beneficial effects on other outcome dimensions beyond healthcare spending. When treated preschoolers are old enough, it would be interesting to study labor market outcomes and criminal records. Ideally, such evaluations would show that the program – while not reducing healthcare costs – might help to improve the starting conditions of underprivileged children as a prerequisite for healthy and professionally successful lives.

Finally, measures and guidelines that help to rule out supply-induced follow-up treatment are desirable. In particular, less educated parents with probable low health literacy may be more affected directly by supply-induced demand for pediatricians. Managed care techniques, including binding treatment pathways and reviewing processes for the medical necessity of specific services, could help to limit the amount of supply-determined healthcare services and to improve the quality of care.

Acknowledgments

For helpful comments, we would like to thank two anonymous referees and the Editor. The paper has also benefited from comments of and discussions with participants of the NRN research seminar "Labor Economics and the Welfare State" in Vienna, the DEEP Advanced Seminar in Economics at the University of Lausanne, the 2015 Annual Conference of the European Society for Population Economics in Izmir (Turkey), and the IWQW Research Seminar at the University of Erlangen-Nuremberg. The usual disclaimer applies. We gratefully acknowledge financial support from the Austrian Federal Ministry of Science, Research and Economic Affairs (bmwfw) and the National Foundation of Research, Technology and Development.

Appendix: Supplementary material

Supplementary data to this article can be found online at doi:10.1016/j.jhealeco.2016.06.011.

References

Almond, D., Currie, J., 2011. Human capital development before age five. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 4b. Elsevier, pp. 1315–1486.
American Academy of Pediatrics, 2010. Reaffirmation of policy statement: identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics 125 (2), e444–e445.
Avellar, S.A., Supplee, L.H., 2013. Effectiveness of home visiting in improving child health and reducing child maltreatment. Pediatrics 132 (Suppl. 2), S90–S99.
Barker, D.J., 1995. Fetal origins of coronary heart disease. British Medical Journal 311, 171–174.
Boyle, C.A., Boulet, S., Schieve, L.A., Cohen, R.A., Blumberg, S.J., Yeargin-Allsopp, M., Visser, S., Kogan, M.D., 2011. Trends in the prevalence of developmental disabilities in US children, 1997–2008. Pediatrics 127 (6), 1034–1042.
Buckles, K.S., Hungerman, D.M., 2013. Season of birth and later outcomes: old questions, new answers. Review of Economics and Statistics 95 (3), 711–724.
Cadman, D., Chambers, L.W., Walter, S.D., Ferguson, R., Johnston, N., McNamee, J., 1987. Evaluation of public health preschool child developmental screening: the process and outcomes of a community program. American Journal of Public Health 77 (1), 45–51.
Campbell, F., Conti, G., Heckman, J.J., Moon, S.H., Pinto, R., Pungello, E., Pan, Y., 2014. Early childhood investments substantially boost adult health. Science 343 (6178), 1478–1485.
Carneiro, P., Ginja, R., 2014. Long term impacts of compensatory preschool on health and behavior: evidence from head start. American Economic Journal Economic Policy 6 (4), 1–42.
Chakrabarti, S., Fombonne, E., 2001. Pervasive developmental disorders in preschool children. Journal of the American Medical Association 285 (24), 3093–3099.
Chou, R., Dana, T., Bougatsos, C., 2011. Screening for visual impairment in children ages 1–5 years: update for the USPSTF. Pediatrics 127 (2), e442–e479.
Conti, G., Heckman, J.J., 2013. The developmental approach to child and adult health. Pediatrics 131 (Suppl. 2), S133–S141.
Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children With Special Needs Project Advisory Committee, 2006. Identifying infants and

young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics 118 (1), 405–420.
Cunha, F., Heckman, J., 2007. The technology of skill formation. The American Economic Review 97 (2), 31–47.
Currie, J., Rossin-Slater, M., 2015. Early-life origins of lifecycle well-being: research and policy implications. Journal of Policy Analysis and Management 34 (1), 208–242.
Duncan, G.J., Magnuson, K., 2013. Investing in preschool programs. The Journal of Economic Perspectives: A Journal of the American Economic Association 27 (2), 109–132.
Gönenc, R., Hofmarcher, M.M., Wörgötter, A. 2011. Reforming Austria's highly regarded but costly health system, OECD Economics Department Working Papers 895, OECD, France.
McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: a density test. Journal of Econometrics 142 (2), 698–714.
Moyer, V.A., Butler, M., 2004. Gaps in the evidence for well-child care: a challenge to our profession. Pediatrics 114 (6), 1511–1521.
Peacock, S., Konrad, S., Watson, E., Nickel, D., Muhajarine, N., 2013. Effectiveness of home visiting programs on child outcomes: a systematic review. BMC Public Health 13 (1), 17–30.
Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica: Journal of the Econometric Society 65 (3), 557–586.
Sweet, M.A., Appelbaum, M.I., 2004. Is home visiting an effective strategy? A meta-analytic review of home visiting programs for families with young children. Child Development 75 (5), 1435–1456.
Williams, C., Northstone, K., Harrad, R.A., Sparrow, J.M., Harvey, I., 2002. Amblyopia treatment outcomes after screening before or at age 3 years: follow up from randomised trial. British Medical Journal 324 (7353), 1549–1551.
Zweimüller, M., 2014. The effects of school entry laws on educational attainment and starting wages in an early tracking system. Annals of Economics and Statistics 111–112, 141–169.