I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18313

Difference–in–Kinks Design

Petri Böckerman
Sami Jysmä
Ohto Kanninen

DECEMBER 2025

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18313

Difference–in–Kinks Design

Petri Böckerman
University of Jyväskylä, Labour Institute for Economic Research and IZA

Sami Jysmä
Labour Institute for Economic Research and Tampere University

Ohto Kanninen
Labour Institute for Economic Research and University of Helsinki

DECEMBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Difference–in–Kinks Design*

This paper introduces the Difference-in-Kinks (DiK) design, an econometric framework that extends the standard regression kink design to settings in which the slope of a policy rule varies over time. By combining the key features of the regression kink and difference-in-differences approaches, the DiK design identifies causal effects from variation in kink intensity over time. We formalize both sharp and fuzzy versions of the estimator and derive the identification conditions under a parallel-trends assumption. Applying DiK to Finland's 2011 guarantee pension reform demonstrates that changes in marginal incentives significantly increased the probability of retirement, while the standard regression kink design would have obtained implausibly large estimates in the opposite direction. The DiK design thus offers a flexible framework for policy evaluation in dynamic, nonlinear environments.

JEL Classification: C21, C14, J26

Keywords: regression kink design, difference-in-differences, policy evaluation, causal inference, treatment effects

Corresponding author:
Petri Böckerman
Labour Institute for Economic Research
Pitkänsillanranta 3A
FI-00530 Helsinki
Finland
E-mail: petri.bockerman@labore.fi

* We would like to thank the Social Insurance Institution of Finland (KELA), OP Foundation and the Strategic Research Council of the Academy of Finland (project 345479) for financial support. The usual disclaimer applies.

1 Introduction

This paper introduces the Difference-in-Kinks (DiK) design. The method extends the standard regression kink design (RKD) by using policy changes over time that alter the slope of the policy rule at the kink. While the traditional RKD identifies causal effects using cross-sectional variation in the slope of a policy schedule at a known kink point, this new approach leverages variation over time in the kink's shape or intensity, thereby enabling causal inference when the level of the policy variable remains constant but its slope changes.

We formalize both sharp and fuzzy versions of the DiK estimator and derive the conditions under which the design identifies a policy relevant treatment-on-the-treated parameter. This extension is important because many real-world policies feature nonlinear benefit schedules whose slopes, though not necessarily their levels, change over time—settings in which conventional regression kink or difference-in-differences designs are inadequate.

We present identification assumptions—including a parallel trends assumption at the kink point—and develop estimators for both deterministic (sharp) and stochastic (fuzzy) policy functions. We illustrate the practical value of this new method through an empirical application examining the effects of Finland's guarantee pension reform on retirement behavior. The policy introduced a sharp kink in pension accrual, and the DiK design enables us to identify the causal impact of this change in marginal incentive to retire. Our empirical estimates show a sizable and economically meaningful elasticity of retirement behavior with respect to the guaranteed pension amount, highlighting the relevance of marginal incentives in retirement decisions. The Difference-in-Kinks design therefore offers a promising avenue for exploiting natural experiments in nonlinear policy environments that evolve over time.

2 Intuition behind the DiK

We study a setting in which the objective is to identify the effect of a policy variable B on an outcome of interest Y. In our empirical application, which examines the impact of Finland's guarantee pension reform on retirement behavior, B denotes the level of annual pension benefits, and Y denotes the probability of retirement.

In the sharp setup, the level of the policy variable is assumed to be completely determined by the running variable V, so that for $V = v$, the policy variable is $B = b(v)$. In our empirical example, V is the annual accrued pension without guaranteed pension. In the fuzzy case considered in Section 3.2, the policy variable is not fully determined by the running variable V.

Denote the average potential outcome when receiving treatment level d for those with running variable v and true policy level $b(v)$ by $\mathbb{E}[y(d)|v]$, where $y(d)$ is the potential potential outcome when receiving treatment level d. Note, that we only observe $Y(d = b(v), v) = \mathbb{E}[y(d = b(v))|v]$ at different levels of v. Let's now consider what happens to the observed average outcome when we move along the running variable from some point v_0 by some small amount $dv > 0$. We can decompose the change in the observed average outcome Y as v changes to two parts:

$$\begin{aligned} \frac{dY}{dv} &= \frac{\partial b}{\partial v}\frac{\partial Y}{\partial b} + \left.\frac{dY}{dv}\right|_b \\ &= \underbrace{\frac{\partial b}{\partial v}\frac{\partial Y}{\partial b}}_{\text{Change related to change in treatment}} + \underbrace{\text{Bias}}_{\text{Change related to everything else}} \end{aligned} \tag{1}$$

Here, the first term, $\frac{\partial b}{\partial v}\frac{\partial Y}{\partial b}$, gives us the change in Y related to the change in the treatment b as v changes. There are two parts in this term: first, $\partial b/\partial v$ measures how the policy variable changes with the running variable at v_0. This needs to be nonzero for there to be any impact through the policy. The second part is $\frac{\partial Y}{\partial b}$. This is a causal parameter that captures how the average potential outcome changes when the treatment changes, holding all else constant—that is, for the same group of individuals that have running variable v_0 and treatment $b(v_0)$. This average causal response on the treated ($ACRT$) parameter is the object of our analysis.

The second term, $\left.\frac{dY}{dv}\right|_b$, is a bias term. This needs to be dealt with one way or another for us to be able to identify $\frac{\partial Y}{\partial b}$. The bias term reveals us how the average outcome changes due to all factors other than the policy variable, which is held constant. This term captures both the direct impact of the running variable on Y and the impact of any compositional change (selection) as v increases slightly.

The decomposition in Equation 1 provides us with a useful framework to discuss how DiK differs from RKD. We will next consider how the standard RKD uses Equation 1, how it deals with the bias term, and what happens when it fails to do so. After that we will discuss how DiK deals with this bias.

2.1 RKD versus DiK

In a standard regression kink design (see e.g. Card et al. 2015), the bias term in equation 1 is dealt with by assuming that it is continuous at the kink point $v = 0$. Intuitively, this means that there is no kink in the bias term caused by either the direct impact of the running variable or by compositional differences. When this is the case, the difference between right and left limits of the bias term in equation 1 are equal at the kink point. Hence, if the bias-term is indeed continuous, we have

$$
\begin{aligned}
\frac{dY}{dv}^+ - \frac{dY}{dv}^- &= \underbrace{\frac{\partial b}{\partial v}^+ \frac{\partial Y}{\partial b}^+ - \frac{\partial b}{\partial v}^- \frac{\partial Y}{\partial b}^-}_{\text{Impact of the policy kink}} + \underbrace{[\text{Bias}^+ - \text{Bias}^-]}_{\text{RKD bias term}} \\
&= \frac{db}{dv}^+ \frac{\partial Y}{\partial b}^+ - \frac{db}{dv}^- \frac{\partial Y}{\partial b}^-
\end{aligned} \tag{2}
$$

The standard RKD estimand naturally arises from assuming that $\frac{\partial Y}{\partial b}$ is also continuous in v close to the kink point, so that $\frac{\partial Y}{\partial b}^- = \frac{\partial Y}{\partial b}^+ \equiv \frac{\partial Y(v=0)}{\partial b}$. Then we have that

$$
\frac{dY}{dv}^+ - \frac{dY}{dv}^- = \left[\frac{db}{dv}^+ - \frac{db}{dv}^-\right] \frac{\partial Y(v=0)}{\partial b}, \tag{3}
$$

which yields the sharp RKD estimand simply by dividing by the difference in slopes of the policy schedule. Naturally, this requires there to be a kink at $v = 0$, as otherwise the difference in slopes of the policy schedule would be zero.

Compared to RKD, DiK goes one step further and leverages variation in the slopes of the kink over time. In DiK, we have

$$
\begin{aligned}
&\frac{dY_1}{dv}^+ - \frac{dY_1}{dv}^- - \left[\frac{dY_0}{dv}^+ - \frac{dY_0}{dv}^-\right] \\
&= \underbrace{\frac{db_1}{dv}^+ \frac{\partial Y_1}{\partial b}^+ - \frac{db_1}{dv}^- \frac{\partial Y_1}{\partial b}^-}_{\text{Impact of the policy kink in } t=1} - \underbrace{\left[\frac{db_0}{dv}^+ \frac{\partial Y_0}{\partial b}^+ - \frac{db_0}{dv}^- \frac{\partial Y_0}{\partial b}^-\right]}_{\text{Impact of a pre-existing policy kink}} + \underbrace{[\text{Bias}_1^+ - \text{Bias}_1^- - (\text{Bias}_0^+ - \text{Bias}_0^-)]}_{\text{DiK bias term}}
\end{aligned} \tag{4}
$$

Here, the bias terms cancel out under the assumption that the post-pre difference of the bias term $(\text{Bias}_1 - \text{Bias}_0)$ is continuous in v. This is similar in spirit to the assumption made in RKD, but for the post-pre difference. A revealing way to think about this assumption is that it is really a parallel trends assumption at the kink point: there may be bias in the *cross-sectional* difference at the kink so that $\text{Bias}_t^+ - \text{Bias}_t^-$ may be non-zero, but this should stay the same between pre- and post-periods.

Under this parallel trends assumption, we have

$$
\begin{aligned}
&\frac{dY_1}{dv}^+ - \frac{dY_1}{dv}^- - \left[\frac{dY_0}{dv}^+ - \frac{dY_0}{dv}^-\right] \\
&= \frac{db_1}{dv}^+ \frac{\partial Y_1}{\partial b}^+ - \frac{db_1}{dv}^- \frac{\partial Y_1}{\partial b}^- - \left[\frac{db_0}{dv}^+ \frac{\partial Y_0}{\partial b}^+ - \frac{db_0}{dv}^- \frac{\partial Y_0}{\partial b}^-\right]
\end{aligned} \tag{5}
$$

The sharp DiK estimand then arises by assuming that the average causal response parameter $\frac{\partial Y_t}{\partial b}^{\pm}$ is both continuous at $v = 0$ and stable over time so that $\frac{\partial Y_1}{\partial b}^- = \frac{\partial Y_1}{\partial b}^+ = \frac{\partial Y_0}{\partial b}^- =$

$\frac{\partial Y_0}{\partial b}^+ \equiv \frac{\partial Y(v=0)}{\partial b}$. When this assumption is in place, we have

$$\begin{aligned}&\frac{dY_1}{dv}^+ - \frac{dY_1}{dv}^- - \left[\frac{dY_0}{dv}^+ - \frac{dY_0}{dv}^-\right] \\ &= \left[\frac{db_1}{dv}^+ - \frac{db_1}{dv}^- - \left[\frac{db_0}{dv}^+ - \frac{db_0}{dv}^-\right]\right] \frac{\partial Y(v=0)}{\partial b},\end{aligned} \tag{6}$$

which yields the sharp DiK estimand simply by dividing by the difference-in-differences of the slope of the policy schedule. This leads to the natural assumption that there must be a change in the kink at $v = 0$. What is noteworthy here is that this does *not* require one side of the kink to stay the same – both derivatives may change. In other words, we do not need a "clean" control group for identification.

Dividing by the difference-in-differences of the slope of the policy schedule at the kink point, the sharp DiK estimand is

$$\frac{\frac{dY_1}{dv}^+ - \frac{dY_1}{dv}^- - \left[\frac{dY_0}{dv}^+ - \frac{dY_0}{dv}^-\right]}{\frac{db_1}{dv}^+ - \frac{db_1}{dv}^- - \left[\frac{db_0}{dv}^+ - \frac{db_0}{dv}^-\right]} = \frac{\partial Y(v=0)}{\partial b}. \tag{7}$$

In sum, as long as all the derivatives and limits used in this section exist and are finite, the key identifying assumptions of sharp RKD can be summarized as 1) continuity of the bias term at the kink, 2) continuity of the average causal response parameter at the kink, and 3) the existence of a known policy kink. A similar list for the sharp DiK would instead be 1) parallel trends in the bias term (implying parallel trends without the change in the kink), 2) continuity and time-stability of the average causal response parameter at the kink, and 3) the existence of a known change in the policy kink. Hence, from a strict identification point of view, sharp DiK identifies the same average casual response parameter as sharp RKD, but allows time-invariant discontinuities at the kink-point. However, sharp DiK does require the additional assumption that the treatment effect is stable over time at the kink-point.

3 Formal identification

3.1 Sharp DiK

Potential outcomes. The outcome of an individual of type u is given by a function $y(v, b, u)$, where v is the running variable and b is the level of the policy (i.e. treatment). We denote the potential outcome of an individual receiving treatment d with running variable v had they received treatment b by $y(d|v, b, u)$. Moreover, we denote the average potential outcome of individuals at v in time-period t as $\mathbb{E}[Y_t(d)|v] = \int_u y(d|v,b,u)dF_t(u|v)$, where $F_t(u|v)$ is the conditional CDF of u at v in period t. Hence, we observe $\mathbb{E}[Y_t(b_t(v))|v]$.

We will begin by introducing the identifying assumptions of Sharp DiK. We consider an arbitrarily small interval around $v = 0$.

Assumption SDiK1 (**Sharp difference-in-kinks policy**).

i) There is a known policy schedule $b_t(v)$ in periods $t \in \{0,1\}$ that is continuously differentiable in a neighborhood of $v = 0$.

ii) At $v = 0$, there is a change in the policy schedule kink from $t = 0$ to $t = 1$, so that

$$\lim_{v\to0^-} b_1'(v) - \lim_{v\to0^+} b_1'(v) \neq \lim_{v\to0^-} b_0'(v) - \lim_{v\to0^+} b_0'(v).$$

Assumption SDiK2 (**Continuously differentiable $\mathbb{E}[Y_t(d)|v]$**). *The function $\mathbb{E}[Y_t(d)|v]$ is continuously differentiable with respect to d and v in some neighborhood of $v = 0$, but not necessarily at $v = 0$ for both $t \in \{0,1\}$.*

Assumption SDiK3 (**Existence of left and right limits**). *For each $t \in \{0,1\}$, the limits of the partial derivatives $\lim_{v\to0^\pm} \partial\mathbb{E}[Y_t(d = b_t(v)|v]/\partial d$, $\lim_{v\to0^\pm} \partial\mathbb{E}[Y_t(d = b_t(v))|v]/\partial v|_d$, and $\lim_{v\to0^\pm} b_t'(v)$ exist and are finite.*

Assumption SDiK4 (**Parallel trends at the kink**). *The impacts of the other parameters than b evolve similarly from $t = 0$ to $t = 1$ for left and right limits at the kink-point $v = 0$:*

$$\lim_{v\to0^-} \frac{\partial\mathbb{E}[Y_1(d = b_1(v))|v]}{\partial v}\bigg|_d - \lim_{v\to0^+} \frac{\partial\mathbb{E}[Y_1(d = b_1(v))|v]}{\partial v}\bigg|_d$$
$$= \lim_{v\to0^-} \frac{\partial\mathbb{E}[Y_0(d = b_0(v))|v]}{\partial v}\bigg|_d - \lim_{v\to0^+} \frac{\partial\mathbb{E}[Y_0(d = b_0(v))|v]}{\partial v}\bigg|_d$$

Assumption SDiK5 (**Stability and continuity of treatment effect at kink**). $\partial\mathbb{E}[Y_t(d)|v]/\partial d$ *is continuous and time-invariant at $v = 0$:*

$$\lim_{v\to0^+} \frac{\partial\mathbb{E}[Y_1(d = b_1(v))|v]}{\partial d} = \lim_{v\to0^-} \frac{\partial\mathbb{E}[Y_1(d = b_1(v))|v]}{\partial d}$$
$$= \lim_{v\to0^+} \frac{\partial\mathbb{E}[Y_0(d = b_0(v))|v]}{\partial d} = \lim_{v\to0^-} \frac{\partial\mathbb{E}[Y_0(d = b_0(v))|v]}{\partial d} \equiv \frac{\partial\mathbb{E}[Y(d)|v = 0]}{\partial d}$$

Assumption SDiK1 states that there is a known policy schedule that is a function of the running variable v. Moreover, at a point in that schedule, normalized to $v = 0$, there is a discontinuous change in the slope from one period to another. This is the change in the kink at that point. Note that Assumption $SDiK1$ allows both sides of the policy schedule to change, and there does not need to be a kink in both periods.

Assumption SDiK2 states that the average outcome is continuously differentiable near the kink, but there may be discontinuity at the kink point. This contrasts with the standard RKD setting, where such a discontinuity at the kink is ruled out.

Assumption SDiK3 is a technical assumption ensuring that the left and right limits of partial derivatives exist and are finite. This allows us to use these limits in our identification result.

Assumption SDiK4 states that other variables than b affect the limits in the same way on both sides. Assumption SDiK4 implies a more standard parallel trends setup, where in the absence of the policy change, the left and right limits of the observable total derivatives would have evolved in the same way. Assumption SDiK4 together with Assumption SDiK2 imply that while there may be discontinuities at the kink, these discontinuities do not change from period 0 to period 1. Similarly as with standard DiD, we can pseudo-test this with pre-trends if we have multiple pre-periods.

Assumption SDiK5 states that the treatment effect $\partial \mathbb{E}[Y_t(d = b_t(v))|v]/\partial d$ is both stable over time and and continuous at the kink. In other words the left and right limits at $v = 0$ are equal and they do not change from $t = 0$ to $t = 1$. This assumption fails if e.g. $b_0(0) \neq b_1(0)$ and the treatment effect is not constant at different levels of b. Another important situation where Assumption SDiK5 fails is when there are compositional changes near the kink resulting in different treatment effects between time-periods.

Denoting $\lim_{v \to 0^\pm} \mathbb{E}[Y_t|v] \equiv Y_t(0^\pm)$ and $\lim_{v \to 0^\pm} b_t'(v) \equiv b_t'(0^\pm)$, our identification result for sharp Difference-in-Kinks is:

Proposition 1 (Identification for Sharp Difference-in-Kinks)**.** *Under Assumptions SDiK1-SDiK5,*

$$\frac{\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right]}{b_1'(0^-) - b_1'(0^+) - [b_0'(0^-) - b_+'0(0^+)]} = \frac{\partial \mathbb{E}[Y(b)|v = 0]}{\partial b}$$

Proof: See Appendix A. The proof follows the intuition outlined in Section 2.

3.2 Fuzzy DiK

Instead of assuming a known deterministic policy function $b(v)$, we now allow the relationship between v and b to be stochastic, with an unobserved term ε. That is, we consider two policy functions, $b_0(v, \varepsilon)$ and $b_1(v, \varepsilon)$.

Again, for a group at the running variable level v, we denote the average potential outcome had they received policy level b as $\mathbb{E}[Y_t(d)|v] = \int_{\varepsilon,u} y(d|v, b, u)dG(\varepsilon, u|v)$, where

$G_t(\varepsilon, u|v)$ is the joint conditional distribution of ε and u at v in period t. We also denote the average potential outcome for a given ε by $\mathbb{E}[Y_t(d)|v,\varepsilon] = \int_u y(d|v,b,u)dF_t(u|v,\varepsilon)$, where $F_t(u|v,\varepsilon)$ is the conditional CDF of u for given v and ε.

In this Appendix, we analyze identification in a Fuzzy DiK setting. To that end, we make the following assumptions.

Assumption FDiK1 (**Fuzzy difference-in-kinks policy**).

i) For each ε, there is a policy schedule $b_t(v,\varepsilon)$ in periods $t \in \{0,1\}$ that is continuously differentiable in a neighborhood of $v = 0$.

ii) At $v = 0$, there is a change in the policy schedule from $t = 0$ to $t = 1$ so that

$$\lim_{v\to 0^-} \frac{\partial \mathbb{E}[b_1(v,\varepsilon)|v]}{\partial v} - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[b_1(v,\varepsilon)|v]}{\partial v} \neq \lim_{v\to 0^-} \frac{\partial \mathbb{E}[b_0(v,\varepsilon)|v]}{\partial v} - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[b_0(v,\varepsilon)|v]}{\partial v},$$

and the change in the kink is either non-negative or non-positive for all ε.

Assumption FDiK2 (**Continuously differentiable** $\mathbb{E}[Y_t(d)|v,\varepsilon]$). *The function $\mathbb{E}[Y_t(d)|v,\varepsilon]$ is continuously differentiable with respect to d and v in some neighborhood of $v = 0$, but not necessarily at $v = 0$ for all ε and both $t \in \{0,1\}$.*

Assumption FDiK3 (**Existence of left and right limits**). *For each $t \in \{0,1\}$ and all ε, the limits of the partial derivatives $\lim_{v\to 0^\pm} \partial \mathbb{E}[Y_t(d = b_t(v)|v,\varepsilon]/\partial d$, $\lim_{v\to 0^\pm} \partial \mathbb{E}[Y_t(d = b_t(v))|v,\varepsilon]/\partial v|_d$, and $\lim_{v\to 0^\pm} \frac{\partial \mathbb{E}[b_t(v,\varepsilon)|v,\varepsilon]}{\partial v}$ exist and are finite.*

Assumption FDiK4 (**Parallel trends at the kink**). *For all ε, the impacts of the other parameters than b evolve similarly from $t = 0$ to $t = 1$ for left and right limits at the kink-point $v = 0$:*

$$\lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d$$
$$= \lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d$$

Assumption FDiK5 (**Stability and continuity of treatment effect at kink**). $\partial \mathbb{E}[Y_t(d)|v,\varepsilon]/\partial d$ *is continuous and time-invariant at $v = 0$ for all ε:*

$$\lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial d} = \lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial d}$$
$$= \lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial d} = \lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial d} \equiv \frac{\partial \mathbb{E}[Y(d)|v = 0,\varepsilon]}{\partial d}$$

Assumption FDiK6 (**Aggregation of ε-level limits**). *For both left and right limits at $v = 0$, and for all ε, we have:*

$$\mathbb{E}\left[\lim_{v\to 0^{\pm}} \frac{d\mathbb{E}[Y_t|v,\varepsilon]}{dv}\middle| v\right] = \lim_{v\to 0^{\pm}} \frac{d\mathbb{E}[Y_t|v]}{dv},$$

and

$$\mathbb{E}\left[\lim_{v\to 0^{\pm}} \frac{d\mathbb{E}[b_t|v,\varepsilon]}{dv}\middle| v\right] = \lim_{v\to 0^{\pm}} \frac{d\mathbb{E}[b_t|v]}{dv}.$$

Denoting $\lim_{v\to 0^{\pm}} \mathbb{E}[Y_t|v] \equiv Y_t(0^{\pm})$ and $\lim_{v\to 0^{\pm}} \mathbb{E}[b_t|v] \equiv b_t(0^{\pm})$, our identification result for fuzzy Difference-in-Kinks is:

Proposition 2 (Identification for Fuzzy Difference-in-Kinks). *Under Assumptions FDiK1-FDiK6,*

$$\frac{\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right]}{\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]} = \tau_{FDiK},$$

where $\tau_{FDiK} = \mathbb{E}\left[\kappa(\varepsilon)\frac{\partial \mathbb{E}[Y(b)|v=0,\varepsilon]}{\partial b}\middle| v = 0\right]$, *and*
$\kappa(\varepsilon) = \left(\frac{db_1(0^-,\varepsilon)}{dv} - \frac{db_1(0^+,\varepsilon)}{dv} - \left[\frac{db_0(0^-,\varepsilon)}{dv} - \frac{db_0(0^+,\varepsilon)}{dv}\right]\right) / \left(\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]\right)$
is a positive weight with $\int_{\varepsilon} \kappa(\varepsilon) = 1$.

Proof: see Appendix B. The intuition behind the proof is that for each ε, one can obtain a similar result as for the sharp DiK. Aggregating over ε then gives then links these to observables under Assumption FDiK6.

Proposition 2 states that under a fuzzy policy schedule, the DiK estimand identifies a weighted average of causal responses, where the weights $\kappa(\varepsilon)$ are proportional to the intensity of the kink change for different ε. In other words, groups with larger changes in the kink at $v = 0$ receive more weight and groups with smaller changes receive less weight.

4 Estimation

In a sharp DiK design, the denominator in Proposition 1 is known, and we only need to estimate the numerator. We do this with local polynomial regressions:

$$
\begin{aligned}
&\min_{\beta_j^-} \sum_{i=1}^{n_t^-} \left[Y_{i,t}^- - \sum_{j=0}^{p} \beta_{j,t}^- (v_{i,t}^-)^j \right]^2 K\left(\frac{v_{i,t}^-}{h}\right) \\
&\min_{\beta_j^+} \sum_{i=1}^{n_t^-} \left[Y_{i,t}^+ - \sum_{j=0}^{p} \beta_{j,t}^+ (v_{i,t}^+)^j \right]^2 K\left(\frac{v_{i,t}^+}{h}\right)
\end{aligned}
\tag{8}
$$

for $t = 0$ and $t = 1$, where the $+$ and $-$ superscripts denote observations below and above the kink at $v = 0$, p is the order of the polynomial being fitted, K is the kernel, and h the bandwidth.

The sharp DiK estimator is then

$$
\hat{\tau}^{SDiK} = \frac{[\hat{\beta}_{1,1}^+ - \hat{\beta}_{1,0}^+] - [\hat{\beta}_{1,1}^- - \hat{\beta}_{1,0}^-]}{\left[\frac{db_1^+}{dv} - \frac{db_0^+}{dv}\right] - \left[\frac{db_1^-}{dv} - \frac{db_0^-}{dv}\right]}, \tag{9}
$$

where $\hat{\beta}_{1,t}^{\pm}$ is the regression coefficient estimate of the first order polynomial term (i.e., linear term) in Equation 8 for either below or above the kink point ($\pm$) in period t, and the denominator is a known scalar.

In a fuzzy DiK setting, the estimator would then be

$$
\hat{\tau}^{FDiK} = \frac{[\hat{\beta}_{1,1}^+ - \hat{\beta}_{1,0}^+] - [\hat{\beta}_{1,1}^- - \hat{\beta}_{1,0}^-]}{\left[\hat{\kappa}_{1,1}^+ - \hat{\kappa}_{1,0}^+\right] - \left[\hat{\kappa}_{1,1}^- - \hat{\kappa}_{1,0}^-\right]}, \tag{10}
$$

where $\hat{\kappa}_{1,t}^{\pm}$ is the local polynomial regression coefficient estimate of the first order polynomial term (similarly as for $\hat{\beta}_{1,t}^{\pm}$ in Equation 8).

5 Empirical illustration: the effect of guaranteed pension on retirement decisions in Finland

We apply the Difference-in-Kinks (DiK) design to analyze the causal impact of Finland's 2011 guarantee pension reform on retirement behavior. This reform introduced a kink in the pension accrual schedule, altering marginal incentives for individuals with low accrued pensions. The DiK framework exploits the change in the slope of the policy rule at the guarantee pension threshold to identify the treatment effect.

5.1 Institutional background

In March 2011, Finland transitioned from a two-tier to a three-tier pension system with the introduction of the guarantee pension. Tier 1 of the system is the earnings-related pension. The accrual rate for tier 1 in our study period (2008–2015) was 1.5% of annual earnings until age 52, 1.9% between ages 53 and 62, and 4.5% between ages 63 and 68, excluding the social insurance payment of between 4.3% and 4.5% in the study period. Full retirement age was 63, with the possiblity of claiming pension at 62 with a penalty of 0.4 pp. of full pension per month. It was also possible to delay claiming beyond age 68 for an increase of 0.4 pp. of full pension per month. Tier 2 is the national pension, which was between approximately 560 and 640 euros per month for singles and around 500 and 560 euros for couples during our study period. National pension is tapered by 50 cents for each euro of earnings-related pensions (see Figure A1 Panel A). Eligibility for national pension started at 62, with the full retirement age at 65 years. A reduction or increase of 0.4 pp. was applied for claiming before or after the age of 65 years.

The reform introduced a third tier to the pension system in the form of a guarantee pension. The guarantee pension provides a minimum flat-rate benefit to all individuals whose combined tier 1 and tier 2 pensions fall below a specified threshold. When introduced, it was around 690 euros per month and increased to around 750 euros per month by 2015 (see Figure A1 Panel A). There was a early claiming penalty of 0.4 pp. for claiming before the full retirement age of 65 years. For individuals above age 65 but below the eligibility threshold, there was a strong financial incentive to claim. However, a claimant at age 65 would forgo the late-claiming increases in the national pension (tier 2).

In our research design we define tiers 1 and 2 of the pensions system as the running variable. The reform changed the policy rule for total pension as a function of the running variable. Before the reform, the slope between the running variable and total pension was 1 at all points. After the reform, the slope in the treatment group, those eligible for guarantee pension, the slope changed from 1 to 0, creating our differences-in-kinks study design (see Panels B and C in Figure A1.

5.2 Data and sample

We use the FOLK dataset from Statistics Finland, a nationwide register-based dataset covering all individuals living in Finland. The data contain high-quality individual-level data on employment, income, education, and demographics. We complement this data from the

Center for Pensions, the Social Insurance Institution of Finland and the Finnish Tax Authority for complementing individual-level details on accrual and eligibility of the three tiers (earnings-related, national and guarantee pension) of the Finnish pension system.

In our empirical study we focus on those aged 65 and more, since they had unequivocal increase in incentives to claim earlier due to the reform. To do this, we restrict our sample to individuals aged 64–69 at the start of the year. We include 64-year-olds, as they reach age 65 during the one-year follow-up period and provide additional statistical power for inference. We restrict our sample to years 2008–2015. This only gives us 3 pre-treatment years, but we are bound by a previous reform implemented in 2005, which had major effects on claiming patterns for multiple years (see Gruber et al. (2022)). We also require that the individual has not claimed any type of pension to be included in the sample.

There is no established bandwidth estimator for the differences-in-kinks design. We choose an ad hoc main bandwidth and confirm the robustness of our estimates for bandwidth choice.[1]

Table A1 summarizes the characteristics of both the full sample and the estimation sample (individuals within the main bandwidth). In the main sample, the mean claiming rate is 72%, with an average accrued pension of €8,145 and average earnings of €4,104. These descriptive patterns highlight that the sample is concentrated around the eligibility threshold, a prerequisite for credible kink-based identification. Panel B shows that the main sample represents around one tenth of all non-retired 64 to 69 year-olds. Figure A1, Panel D, shows a histogram of our sample (N = 11,241) by treatment status and the running variable.

5.3 Baseline estimates and robustness

Panel A of Figure A2 plots the pre- and post-reform retirement rates across bins of accrued pension. The slope is considerably different in the treatment group from the control group in the pre-period. Our focus is on the change of the retirement probability before and after the treatment, which we use to estimate causal effect of the reform (Panel B). In Panel B, we find visual evidence of a gradual increase in retirement rates in the treatment group as the distance from the eligibility threshold grows, consistent with a behavioral effect of the reform. In Panel C, we present the slope estimates from equation **??**, when separating the regressions by year and treatment status. We observe a change in the treatment group following the reform. Panel D visualizes event-study type yearly differences-in-kinks estimates showing no signs of a pre-trend in the three preceding year of the reform and an apparent effect following the reform.

[1]See Böckerman et al. (2018) for an application of this approach in the context of an RKD.

Table A2 reports the estimated treatment effect. The DiK estimate of -0.219 (SE = 0.041) is statistically significant at the 1% level, indicating that the slope of the retirement rate with respect to accrued pension decreased following the introduction of guarantee pension. This corresponds to an effect of 0.219 percentage points per €1,000 reduction in marginal pension accrual. The implied elasticity of retirement with respect to pension income is 2.18, suggesting a high sensitivity of retirement behavior to marginal financial incentives near the eligibility cutoff.

Figure A3 confirms the robustness of our DiK estimates. Panel A shows that the point estimates are stable across different bandwidth choices, with varying statistical significance. Panel B evaluates sensitivity to donut hole sizes, Panel C presents placebo tests with artificial kinks, and Panel D examines different follow-up periods. In all cases, our results tend to remain consistent and robust to specification choices. Figure A4 documents additional validity tests. Panel A reveals no sign of non-smoothness of the difference in pre- and post-treatment density functions (see Table A3). Panels B to D analyze whether covariates (age, earnings and female dummy) change due to the reform. We find a significant estimate for age, and an insignificant estimate for earnings and female dummy. Controlling for age in the main regression lowers the main estimate by 32% from -0.219 to -0.150 and the corresponding t-value by 22%, yet the estimate remains significant at the 1% level.

6 Conclusion

This paper introduces the Difference-in-Kinks (DiK) design, a novel econometric framework that extends the traditional regression kink design to settings where the slope of a policy rule changes over time. By combining the strengths of both regression kink and difference-in-differences approaches, the DiK design enables causal inference from policy reforms that alter marginal incentives without affecting benefit levels. We formalize both sharp and fuzzy versions of the estimator and provide identification conditions under a parallel trends assumption at the kink point.

We employ the DiK design to analyze the effects of Finland's 2011 guarantee pension reform, which introduced a discontinuity in the slope of pension accrual. Our empirical results show that changes in marginal incentives significantly affect retirement behavior. Notably, we find a sizable elasticity of retirement with respect to pension income, highlighting that individuals close to the eligibility threshold are highly responsive to changes in benefit slopes. This effect is consistent with quasi-experimental evidence that local changes in pension accrual slopes materially affect retirement behaviour (Ye 2022; Kolsrud et al. 2024).

The DiK design provides a flexible and policy-relevant tool for empirical researchers studying dynamic non-linear policy environments. The method opens up new opportunities for evaluating marginal treatment effects in institutional settings where traditional RKD or DiD approaches fall short. Future research can build on this framework to examine a broader set of policy changes that involve shifts in slopes rather than levels, including taxation, subsidies, and social insurance programs.

References

Böckerman, P., O. Kanninen, and I. Suoniemi (2018): "A kink that makes you sick: The effect of sick pay on absence," *Journal of Applied Econometrics*, 33, 568–579.

Card, D., D. S. Lee, Z. Pei, and A. Weber (2015): "Inference on Causal Effect in a Generalized Regression Kink Design," *Econometrica*, 83, 2453–2483.

Gruber, J., O. Kanninen, and T. Ravaska (2022): "Relabeling, retirement and regret," *Journal of Public Economics*, 211, 104677.

Kolsrud, J., C. Landais, D. Reck, and J. Spinnewijn (2024): "Retirement consumption and pension design," *American Economic Review*, 114, 89–133.

Ye, H. (2022): "The effect of pension subsidies on the retirement timing of older women," *Journal of the European Economic Association*, 20, 1048–1094.

A Appendix – Sharp DiK

Potential outcomes. The outcome of an individual of type u is given by a function $y(v, b, u)$, where v is the running variable and b is the level of the policy (i.e. treatment). We denote the potential outcome of an individual receiving treatment d with running variable v had they received treatment b by $y(d|v, b, u)$. Moreover, we denote the average potential outcome of individuals at v in time-period t as $\mathbb{E}[Y_t(d)|v] = \int_u y(d|v, b, u)dF_t(u|v)$, where $F_t(u|v)$ is the conditional CDF of u at v in period t. Hence, we observe $\mathbb{E}[Y_t(b_t(v))|v]$.

In this Appendix, we analyze identification in a Sharp DiK setting. To that end, we make the following assumptions:

Assumption SDiK1 (**Sharp difference-in-kinks policy**).

i) There is a known policy schedule $b_t(v)$ in periods $t \in \{0, 1\}$ that is continuously differentiable in a neighborhood of $v = 0$.

ii) At $v = 0$, there is a change in the policy schedule kink from $t = 0$ to $t = 1$, so that

$$\lim_{v\to 0^-} b_1'(v) - \lim_{v\to 0^+} b_1'(v) \neq \lim_{v\to 0^-} b_0'(v) - \lim_{v\to 0^+} b_0'(v).$$

Assumption SDiK2 (**Continuously differentiable** $\mathbb{E}[Y_t(d)|v]$). *The function $\mathbb{E}[Y_t(d)|v]$ is continuously differentiable with respect to d and v in some neighborhood of $v = 0$, but not necessarily at $v = 0$ for both $t \in \{0, 1\}$.*

Assumption SDiK3 (**Existence of left and right limits**). *For each $t \in \{0, 1\}$, the limits of the partial derivatives $\lim_{v\to 0^\pm} \partial\mathbb{E}[Y_t(d = b_t(v)|v]/\partial d$, $\lim_{v\to 0^\pm} \partial\mathbb{E}[Y_t(d = b_t(v))|v]/\partial v|_d$, and $\lim_{v\to 0^\pm} b_t'(v)$ exist and are finite.*

Assumption SDiK4 (**Parallel trends at the kink**). *The impacts of the other parameters than b evolve similarly from $t = 0$ to $t = 1$ for left and right limits at the kink-point $v = 0$:*

$$\begin{aligned} &\lim_{v\to 0^-} \frac{\partial\mathbb{E}[Y_1(d = b_1(v))|v]}{\partial v}\bigg|_d - \lim_{v\to 0^+} \frac{\partial\mathbb{E}[Y_1(d = b_1(v))|v]}{\partial v}\bigg|_d \\ &= \lim_{v\to 0^-} \frac{\partial\mathbb{E}[Y_0(d = b_0(v))|v]}{\partial v}\bigg|_d - \lim_{v\to 0^+} \frac{\partial\mathbb{E}[Y_0(d = b_0(v))|v]}{\partial v}\bigg|_d \end{aligned}$$

Assumption SDiK5 (**Stability and continuity of treatment effect at kink**). $\partial\mathbb{E}[Y_t(d)|v]/\partial d$

is continuous and time-invariant at $v = 0$:

$$\lim_{v \to 0^+} \frac{\partial \mathbb{E}[Y_1(d = b_1(v))|v]}{\partial d} = \lim_{v \to 0^-} \frac{\partial \mathbb{E}[Y_1(d = b_1(v))|v]}{\partial d}$$
$$= \lim_{v \to 0^+} \frac{\partial \mathbb{E}[Y_0(d = b_0(v))|v]}{\partial d} = \lim_{v \to 0^-} \frac{\partial \mathbb{E}[Y_0(d = b_0(v))|v]}{\partial d} \equiv \frac{\partial \mathbb{E}[Y(d)|v = 0]}{\partial d}$$

Assumption SDiK1 states that there is a known policy schedule that is a function of the running variable v. Moreover, at a point in that schedule, normalized to $v = 0$, there is a discontinuous change in the slope from one period to another. This is the change in the kink at that point. Note that Assumption $SDiK1$ allows both sides of the policy schedule to change, and there does not need to be a kink in both periods.

Assumption SDiK2 states that the average outcome is continuously differentiable near the kink, but there may be discontinuity at the kink point. This contrasts with the standard RKD setting, where such a discontinuity at the kink is ruled out.

Assumption SDiK3 is a technical assumption ensuring that the left and right limits of partial derivatives exist and are finite. This allows us to use these limits in our identification result.

Assumption SDiK4 states that other variables than b affect the limits in the same way on both sides. Assumption SDiK4 implies a more standard parallel trends setup, where in the absence of the policy change, the left and right limits of the observable total derivatives would have evolved in the same way. Assumption SDiK4 together with Assumption SDiK2 imply that while there may be discontinuities at the kink, these discontinuities do not change from period 0 to period 1. Similarly as with standard DiD, we can pseudo-test this with pre-trends if we have multiple pre-periods.

Assumption SDiK5 states that the treatment effect $\partial \mathbb{E}[Y_t(d = b_t(v))|v]/\partial d$ is both stable over time and and continuous at the kink. In other words the left and right limits at $v = 0$ are equal and they do not change from $t = 0$ to $t = 1$. This assumption fails if e.g. $b_0(0) \neq b_1(0)$ and the treatment effect is not constant at different levels of b. Another important situation where Assumption SDiK5 fails is when there are compositional changes near the kink resulting in different treatment effects between time-periods.

Denoting $\lim_{v \to 0^\pm} \mathbb{E}[Y_t|v] \equiv Y_t(0^\pm)$ and $\lim_{v \to 0^\pm} b_t'(v) \equiv b_t'(0^\pm)$, our identification result for sharp Difference-in-Kinks is:

Proposition 1 (Identification for Sharp Difference-in-Kinks)**.** *Under Assumptions SDiK1-SDiK5,*

$$\frac{\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right]}{b_1'(0^-) - b_1'(0^+) - [b_0'(0^-) - b_+'0(0^+)]} = \frac{\partial \mathbb{E}[Y(b)|v = 0]}{\partial b}$$

Proof. Let's first note that using the potential outcomes notation, we can denote $\frac{dY_t(0^\pm)}{dv} = \lim_{v\to 0^\pm} d\mathbb{E}[Y(d = b_t|v]/dv$, where the dependence of b_t on v is left implicit to lighten notation. Using Assumption SDiK2 we then have that at a neighborhood around $v = 0$,

$$\frac{dY_t(0^\pm)}{dv} = \lim_{v\to 0^\pm} \frac{\partial\mathbb{E}[Y(d = b_t)|v]}{\partial d} b_t'(0^\pm) + \lim_{v\to 0^\pm} \left.\frac{\partial\mathbb{E}[Y(d = b_t)|v]}{\partial v}\right|_d .$$

This means that

$$\begin{aligned}
&\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right] \\
&= \lim_{v\to 0^-} \frac{\partial\mathbb{E}[Y(d = b_1)|v]}{\partial d} b_1'(0^-) - \lim_{v\to 0^+} \frac{\partial\mathbb{E}[Y(d = b_1)|v]}{\partial d} b_1'(0^+) \\
&- \left[\lim_{v\to 0^-} \frac{\partial\mathbb{E}[Y(d = b_0)|v]}{\partial b} b_0'(0^-) - \lim_{v\to 0^+} \frac{\partial\mathbb{E}[Y(d = b_0)|v]}{\partial d} b_0'(0^+)\right] \\
&+ \lim_{v\to 0^-} \left.\frac{\partial\mathbb{E}[Y(d = b_1)|v]}{\partial v}\right|_d - \lim_{v\to 0^+} \left.\frac{\partial\mathbb{E}[Y(d = b_1)|v]}{\partial v}\right|_d \\
&- \left[\lim_{v\to 0^-} \left.\frac{\partial\mathbb{E}[Y(d = b_0)|v]}{\partial v}\right|_d - \lim_{v\to 0^+} \left.\frac{\partial\mathbb{E}[Y(d = b_0)|v]}{\partial v}\right|_d\right]
\end{aligned}$$

By Assumption SDiK4, the four final terms equal 0, and by Assumption SDiK5 $\lim_{v\to 0^\pm} \frac{\partial\mathbb{E}[Y(d=b_t)|v]}{\partial d} = \frac{\partial\mathbb{E}[Y(d)|v=0]}{\partial d}$ for $t \in \{0, 1\}$. Hence, we have that

$$\begin{aligned}
&\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right] \\
&= \left[b_1'(0^-) - b_1'(0^+) - \left[b_0'(0^-) - b_0'(0^+)\right]\right] \frac{\partial\mathbb{E}[Y(d)|v = 0]}{\partial d}.
\end{aligned}$$

The result then follows by dividing by $b_1'(0^-) - b_1'(0^+) - \left[b_0'(0^-) - b_+'0(0^+)\right]$, which is non-zero by Assumption SDiK1. We then have that

$$\frac{\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right]}{b_1'(0^-) - b_1'(0^+) - [b_0'(0^-) - b_+'0(0^+)]} = \frac{\partial\mathbb{E}[Y(d)|v = 0]}{\partial d},$$

proving Proposition 1.

B Appendix – Fuzzy DiK

Instead of assuming a known deterministic policy function $b(v)$, we now allow the relationship between v and b to be stochastic, with an unobserved term ε. That is, we consider two policy functions, $b_0(v,\varepsilon)$ and $b_1(v,\varepsilon)$.

Again, for a group at the running variable level v, we denote the average potential outcome had they received policy level b as $\mathbb{E}[Y_t(d)|v] = \int_{\varepsilon,u} y(d|v,b,u)dG(\varepsilon,u|v)$, where $G_t(\varepsilon,u|v)$ is the joint conditional distribution of ε and u at v in period t. We also denote the average potential outcome for a given ε by $\mathbb{E}[Y_t(d)|v,\varepsilon] = \int_u y(d|v,b,u)dF_t(u|v,\varepsilon)$, where $F_t(u|v,\varepsilon)$ is the conditional CDF of u for given v and ε.

In this Appendix, we analyze identification in a Fuzzy DiK setting. To that end, we make the following assumptions.

Assumption FDiK1 (Fuzzy difference-in-kinks policy).

i) For each ε, there is a policy schedule $b_t(v,\varepsilon)$ in periods $t \in \{0,1\}$ that is continuously differentiable in a neighborhood of $v = 0$.

ii) At $v = 0$, there is a change in the policy schedule from $t = 0$ to $t = 1$ so that

$$\lim_{v\to 0^-} \frac{\partial \mathbb{E}[b_1(v,\varepsilon)|v]}{\partial v} - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[b_1(v,\varepsilon)|v]}{\partial v} \neq \lim_{v\to 0^-} \frac{\partial \mathbb{E}[b_0(v,\varepsilon)|v]}{\partial v} - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[b_0(v,\varepsilon)|v]}{\partial v},$$

and the change in the kink is either non-negative or non-positive for all ε.

Assumption FDiK2 (Continuously differentiable $\mathbb{E}[Y_t(d)|v,\varepsilon]$). *The function $\mathbb{E}[Y_t(d)|v,\varepsilon]$ is continuously differentiable with respect to d and v in some neighborhood of $v = 0$, but not necessarily at $v = 0$ for all ε and both $t \in \{0,1\}$.*

Assumption FDiK3 (Existence of left and right limits). *For each $t \in \{0,1\}$ and all ε, the limits of the partial derivatives $\lim_{v\to 0^\pm} \partial \mathbb{E}[Y_t(d = b_t(v)|v,\varepsilon]/\partial d$, $\lim_{v\to 0^\pm} \partial \mathbb{E}[Y_t(d = b_t(v))|v,\varepsilon]/\partial v|_d$, and $\lim_{v\to 0^\pm} \frac{\partial \mathbb{E}[b_t(v,\varepsilon)|v,\varepsilon]}{\partial v}$ exist and are finite.*

Assumption FDiK4 (Parallel trends at the kink). *For all ε, the impacts of the other parameters than b evolve similarly from $t = 0$ to $t = 1$ for left and right limits at the kink-point $v = 0$:*

$$\lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d$$
$$= \lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d - \lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial v}\bigg|_d$$

Assumption FDiK5 (Stability and continuity of treatment effect at kink). $\partial \mathbb{E}[Y_t(d)|v,\varepsilon]/\partial d$ *is continuous and time-invariant at* $v = 0$ *for all* ε*:*

$$\lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial d} = \lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_1(d = b_1(v,\varepsilon))|v,\varepsilon]}{\partial d}$$
$$= \lim_{v\to 0^+} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial d} = \lim_{v\to 0^-} \frac{\partial \mathbb{E}[Y_0(d = b_0(v,\varepsilon))|v,\varepsilon]}{\partial d} \equiv \frac{\partial \mathbb{E}[Y(d)|v=0,\varepsilon]}{\partial d}$$

Assumption FDiK6 (Aggregation of ε-level limits). *For both left and right limits at* $v = 0$*, and for all* ε*, we have:*

$$\mathbb{E}\left[\lim_{v\to 0^\pm} \frac{d\mathbb{E}[Y_t|v,\varepsilon]}{dv}\middle| v\right] = \lim_{v\to 0^\pm} \frac{d\mathbb{E}[Y_t|v]}{dv},$$

and

$$\mathbb{E}\left[\lim_{v\to 0^\pm} \frac{d\mathbb{E}[b_t|v,\varepsilon]}{dv}\middle| v\right] = \lim_{v\to 0^\pm} \frac{d\mathbb{E}[b_t|v]}{dv}.$$

Assumption FDiK1 states that there is variation in the policy schedule based on unobservable ε. One can think of ε as indicating different policy schedules. In fuzzy DiK, we also have a kink in the average policy schedule at $v = 0$. The kink is assumed to have the same sign for all ε. This rules out situations where some have a positive kink at $v = 0$ and some have a negative kink at $v = 0$. Although one should note that there may be individuals without any kink at $v = 0$.

Assumptions FDiK2-FDiK4 are similar to the sharp DiK assumptions SDiK2-SDiK4, but at the level of each ε.

Assumption FDiK6 is a regularity assumption we need so that we are able to switch the expectation sign with limits and derivation.

Denoting $\lim_{v\to 0^\pm} d\mathbb{E}[Y_t|v]/dv \equiv dY_t(0^\pm)/dv$ and $\lim_{v\to 0^\pm} d\mathbb{E}[b_t|v]/dv \equiv db_t(0^\pm)/dv$, our identification result for fuzzy Difference-in-Kinks is:

Proposition 2 (Identification for Fuzzy Difference-in-Kinks)**.** *Under Assumptions FDiK1-FDiK6,*

$$\frac{\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right]}{\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]} = \tau_{FDiK},$$

where $\tau_{FDiK} = \mathbb{E}\left[\kappa(\varepsilon)\frac{\partial \mathbb{E}[Y(b)|v=0,\varepsilon]}{\partial b}\middle| v = 0\right] = \mathbb{E}\left[\kappa(\varepsilon)\tau_{SDiK}(\varepsilon)|\, v = 0\right]$*, and*
$\kappa(\varepsilon) = \left(\frac{db_1(0^-,\varepsilon)}{dv} - \frac{db_1(0^+,\varepsilon)}{dv} - \left[\frac{db_0(0^-,\varepsilon)}{dv} - \frac{db_0(0^+,\varepsilon)}{dv}\right]\right) / \left(\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]\right)$
is a positive weight with $\int_\varepsilon \kappa(\varepsilon) = 1$*.*

Proof. Note that for each ε, the policy schedules $b_0(v, \varepsilon)$ and $b_1(v, \varepsilon)$ are fixed. Hence, following the proof for Sharp DiK, we have that under Assumptions FDiK1-FDiK5,

$$\frac{dY_1(0^-, \varepsilon)}{dv} - \frac{dY_1(0^+, \varepsilon)}{dv} - \left[\frac{dY_0(0^-, \varepsilon)}{dv} - \frac{dY_0(0^+, \varepsilon)}{dv}\right]$$
$$= \left[\frac{db_1(0^-, \varepsilon)}{dv} - \frac{db_1(0^+, \varepsilon)}{dv} - \left[\frac{db_0(0^-, \varepsilon)}{dv} - \frac{db_0(0^+, \varepsilon)}{dv}\right]\right] \frac{\partial \mathbb{E}[Y(d, \varepsilon)|v = 0]}{\partial d}.$$

Moreover, taking expectations over ε on both sides, and using Assumption FDiK6 this becomes

$$\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right]$$
$$= \left[\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]\right] \mathbb{E}\left[\kappa(\varepsilon) \frac{\partial \mathbb{E}[Y(d, \varepsilon)|v = 0]}{\partial d}\middle| v = 0\right],$$

where $\kappa(\varepsilon) = \left(\frac{db_1(0^-,\varepsilon)}{dv} - \frac{db_1(0^+,\varepsilon)}{dv} - \left[\frac{db_0(0^-,\varepsilon)}{dv} - \frac{db_0(0^+,\varepsilon)}{dv}\right]\right) / \left(\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]\right)$ is a positive weight with $\int_\varepsilon \kappa(\varepsilon) = 1$ (the nominator of the weights have the same sign for all ε due to due to Assumption FDiK1).

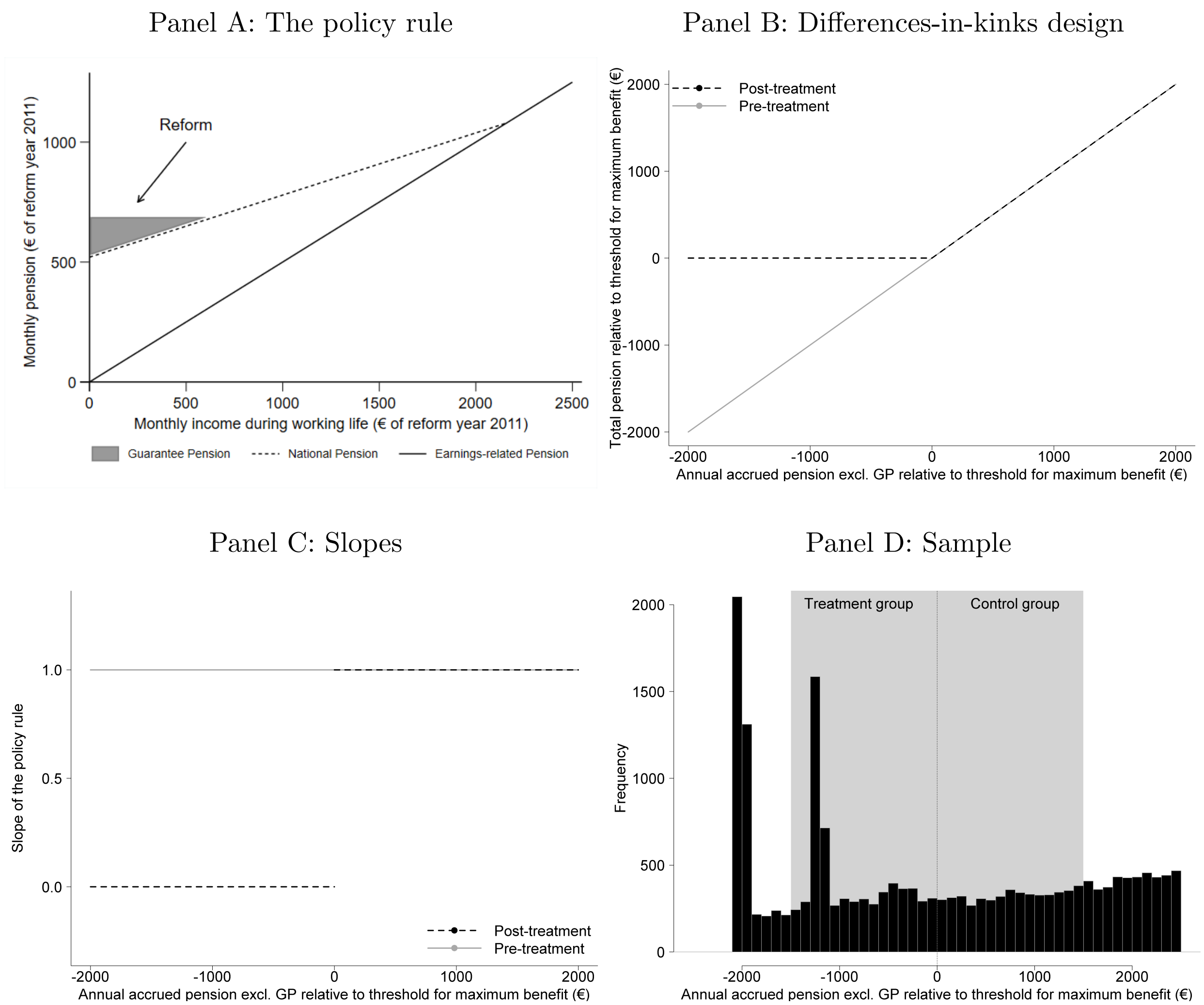
Under Assumption FDiK1, $\left[\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]\right] \neq 0$, so that

$$\frac{\frac{dY_1(0^-)}{dv} - \frac{dY_1(0^+)}{dv} - \left[\frac{dY_0(0^-)}{dv} - \frac{dY_0(0^+)}{dv}\right]}{\frac{db_1(0^-)}{dv} - \frac{db_1(0^+)}{dv} - \left[\frac{db_0(0^-)}{dv} - \frac{db_0(0^+)}{dv}\right]} = \mathbb{E}\left[\kappa(\varepsilon) \frac{\partial \mathbb{E}[Y(d, \varepsilon)|v = 0]}{\partial d}\middle| v = 0\right],$$

proving Proposition 2.

C Appendix - Figures and Tables

Figure A1: Research Design



Notes: **Panel A** depicts the three tiers of the Finnish pension system in the year of the introduction of guarantee pension (2011). Annual national pension was 7037.52 (single) or 6242.28 (with spouse) euros at zero earnings-related pension. Annual guarantee pension was set at 8252.88 euros. **Panels B–C** show the policy rule in terms of differences-in-kinks design. The running variable (x-axis) is the sum of earnings-related pension and national pension. With the introduction of the guarantee pension, the slope of the policy rule for the treatment group shifted from 1 to 0 (i.e., $\tau = -1$). **Panel D** shows the histogram of the sample and the main specification bandwidth of 1,500 euros. Sample size is 11,241.

Table A1: Descriptive statistics

Panel A: Means and SDs

	All observations		Main sample	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
Retirement rate	0.60	0.49	0.72	0.45
Accrued pension	23,430	16,364	8,145	913
Running variable	15,122	16,345	-142	914
Age	65.08	1.00	65.04	1.13
Earnings	24,842	30,678	4,104	10,539

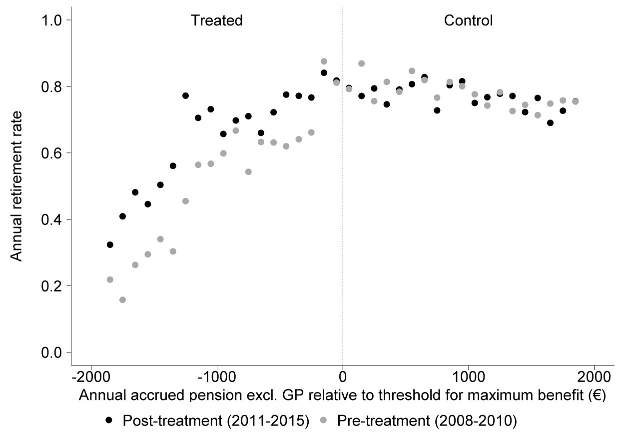
Panel B: Sample size

	2008	2009	2010	2011	2012	2013	2014	2015
All observations	9,895	11,906	14,332	16,045	17,243	18,762	19,047	20,938
Main sample	1,188	1,236	1,335	1,444	1,405	1,478	1,482	1,673

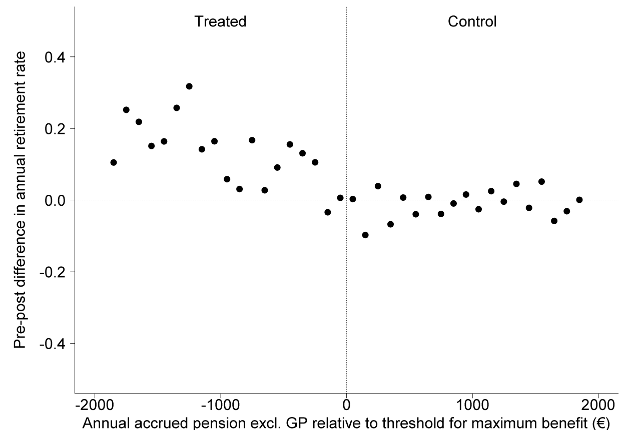
Notes: All values refer to individuals who are not retired and are aged between 64 and 68 years at the beginning of the year. The main sample includes only those within the CCT bandwidth of the guarantee pension eligibility threshold. Claiming rate is the annual proportion of sample claiming old-age pension. Accrued pension is the tier 1 (earnings-related pension), 2 (national pension) and 3 (guarantee pension) accrual at the start of the year. Running variable is accrued annual pension for tiers 1 and 2 normalized at the guarantee pension eligibility threshold. Total sample size of the main sample is 11,241.

Figure A2: Results

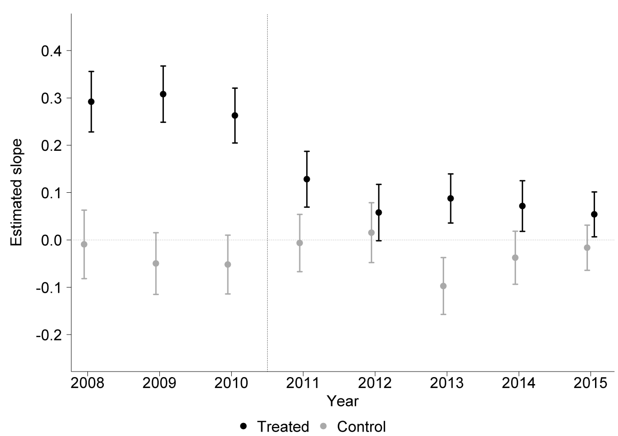
Panel A: Pre- and post-treatment means



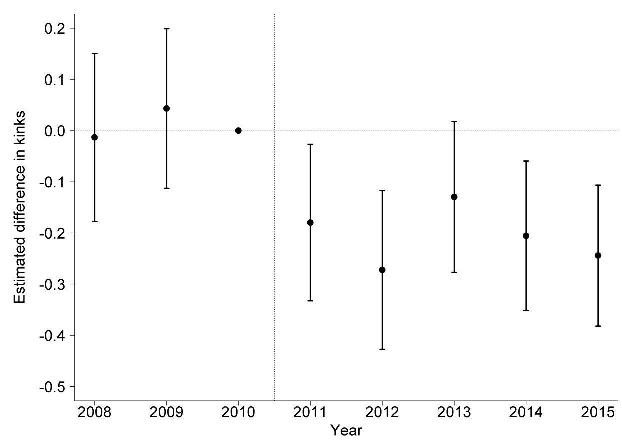
Panel B: Pre- and post-treatment differences



Panel C: Estimated slopes by treatment status



Panel D: Event study design



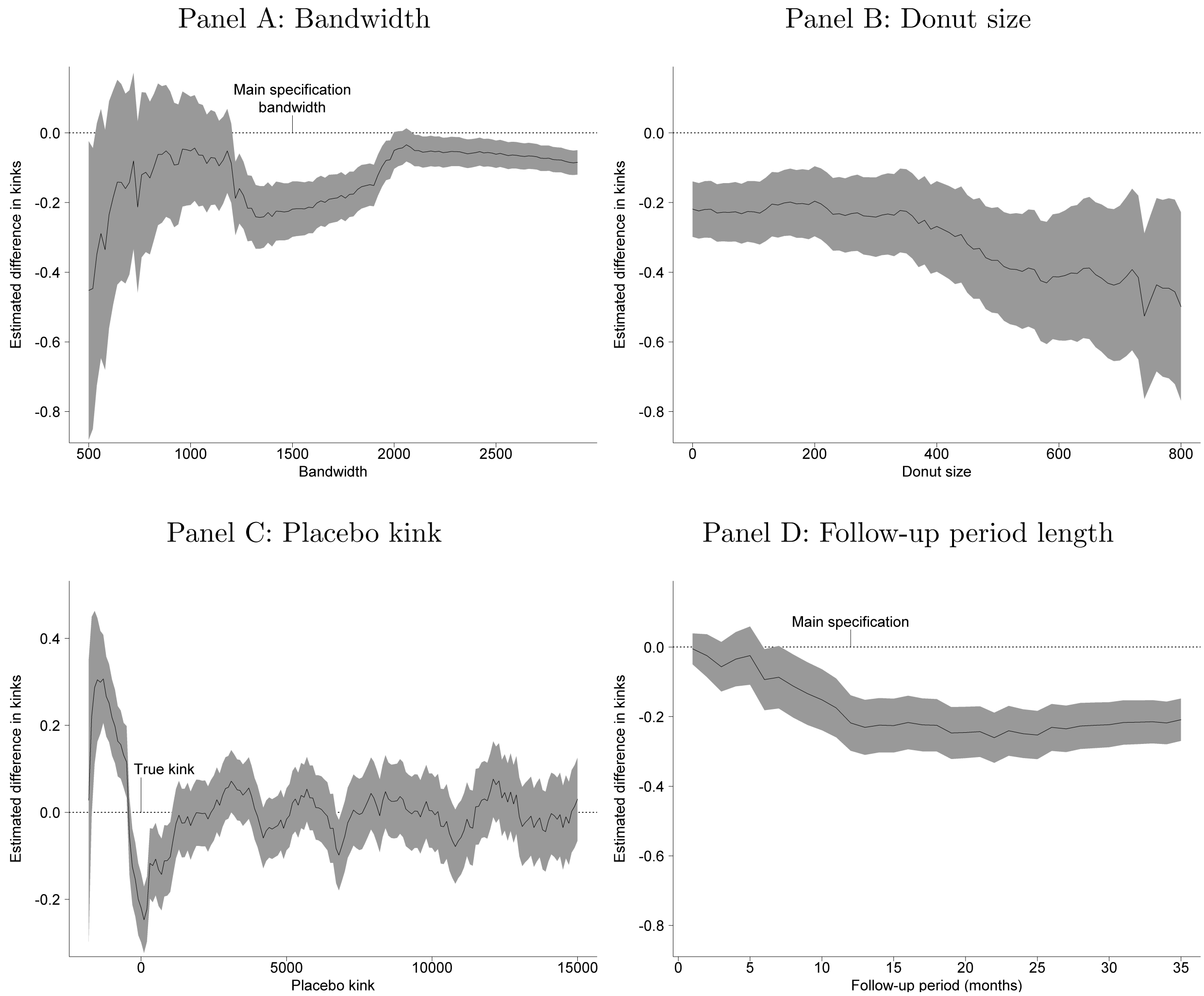
Notes: **Panel A** shows the annual pre- and post-treatment retirement rate means in bins of 100 euros of accrued annual tier 1 and 2 pensions. Tier 2 (national pension) accrual is a decreasing function of tier 1 (earnings-related pension) accrual. **Panel B** shows the differences between the pre- and post-treatment means depicted in Panel A. **Panel C** shows the estimated slopes by treatment status from a single regression (equation XXX) with a year interaction term. **Panel D** presents the event-study design, reporting yearly difference-in-kinks estimates relative to the baseline year 2010.

Table A2: Results

	(1)
Panel A: The estimate	
Difference-in-kinks estimate (β_1)	-0.219 (0.041)***
Baseline slope	-0.039* (0.020)
Baseline slope for post period	0.011 (0.012)
Baseline slope for treatment group	0.327 (0.034)***
Yearly fixed effects	✓
N	11,241
Panel B: The economic interpretation	
Change in policy rule (γ)	-1
The effect per 1,000 euros of GP ($\tau = \frac{\beta_1}{\gamma}$)	0.219
Panel C: Retirement elasticity w.r.t. pension	
Baseline retirement rate (R)	0.83
Baseline pension (Y)	8,290
Elasticity ($\frac{\Delta R / R}{\Delta Y / Y}$)	2.18

Notes: **Panel A** reports the key estimates from the difference-in-kinks (DiK) specification. The main coefficient of interest, β_1, measures the change in the slope of the outcome variable at the kink point for the treatment group relative to the control group. Standard errors are reported in parentheses. Triple asterisks (***) indicate statistical significance at the 1% level. All models include year fixed effects. **Panel B** provides the economic interpretation of the DiK estimate. The change in the slope of the policy rule due to the guarantee pension reform is $\gamma = -1$, implying that the DiK estimate can be directly interpreted as the marginal effect of a 1,000 euro increase in guarantee pension on the outcome of interest. **Panel C** derives the implied elasticity of retirement rate with respect to acccrued pension income. The elasticity is computed as the ratio of the relative change in retirement to the relative change in pension at pre-treatment baseline (calculated at within one tenth of main specification bandwidth from the threshold) retirement rate (R) and pension (Y) in the sample.

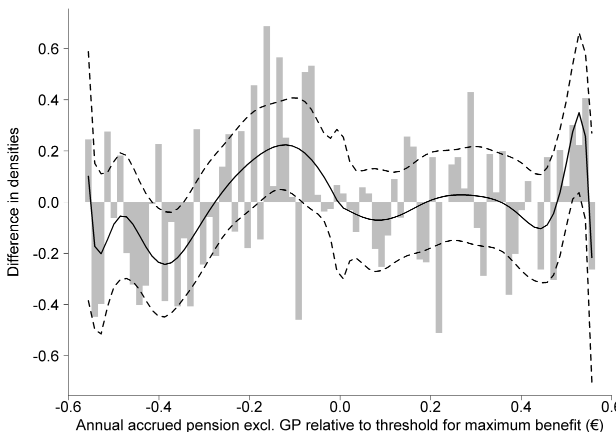
Figure A3: Robustness checks



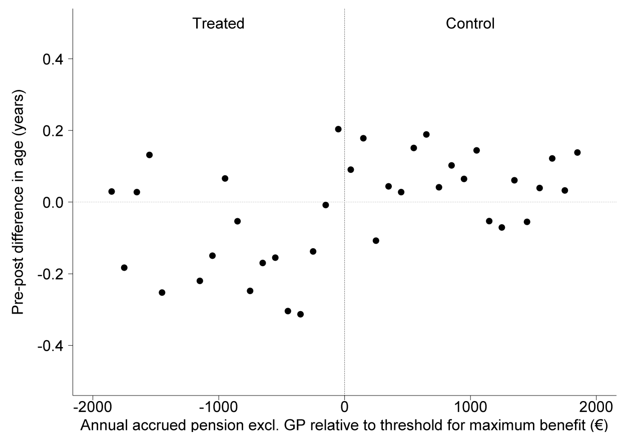
Notes: Each panel depicts the estimates from a series of regressions varying one parameter of the main specification. The gray area is the 95% confidence interval. **Panel A** shows the estimates from 121 regressions varying the bandwidth around the optimal RKD CCT bandwidth by increments of 20 euros. **Panel B** shows the estimates from 81 regressions increasing the the omitted data around the threshold for maximum benefit (donut) by increments of 10 euros. No data are omitted in the main specification. **Panel C** shows the estimates from 169 regressions in which a placebo kink is created by moving the kink point by increments of 100 euros. Of the 163 regressions, which are not around the true kink point, 12 (7.4%) are significant at the 5 % significance level. Most regressions to the left of the true kink lack sufficient mass to cover the whole support of the optimal bandwidth. **Panel D** shows the estimates from 35 regressions varying the follow-up period length. For follow-up periods shorter than 12 months, part of the sample does not reach the next round age, at which much of the claiming occurs. For follow-up periods longer than 14 months, the 2010 sample is affected by the guarantee pension reform, which took effect in March 2011, and for periods longer than 26 months, the 2009 sample is also affected by the reform.

Figure A4: Validity tests

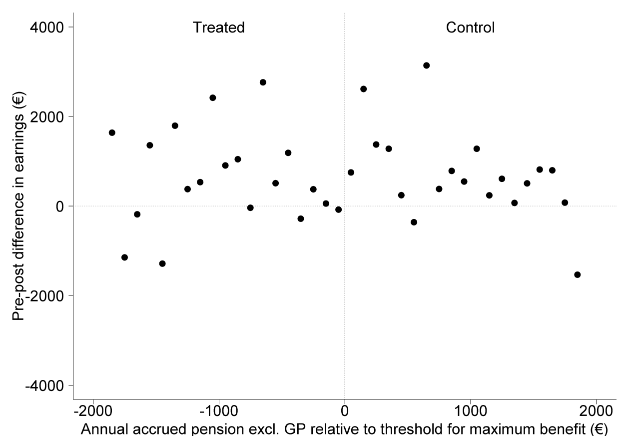
Panel A: Local polynomial density test



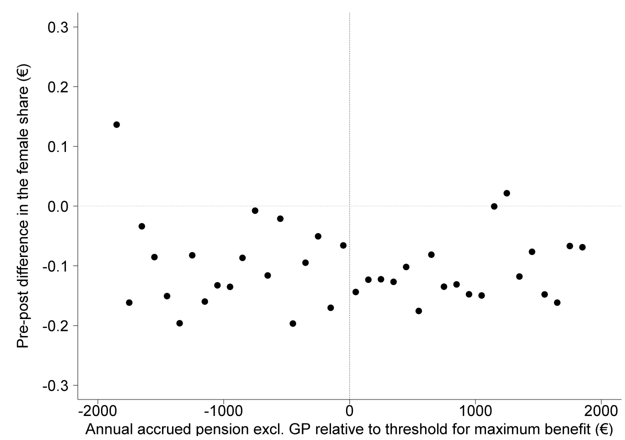
Panel B: Covariate linearity (age)



Panel C: Covariate linearity (earnings)



Panel D: Covariate linearity (female)



Notes: **Panel A** shows the estimated 13th order local polynomial over the pre-post difference of the density function and the corresponding histogram with 80 bins. The bandwidth and bin size are extracted from running the "rddensity" package for a local polynomial density test for regression discontinuity design in R. **Panels B, C and D** show the differences between pre- and post-treatment mean in age, earnings and female dummy in bins of 100 euros of the running variable.

Table A3: Validity tests

	Local polynomial	Covariate linearity		
		Age	Earnings	Female
	(1)	(2)	(3)	(4)
Estimate	-2.20	0.27**	540	0.001
	(10.07)	(0.09)	(927)	(0.039)
Year fixed effects	✓	✓	✓	✓
N	80	13,469	13,469	13,469

Notes: The first column reports the estimated change in the first-order term of the 13th-order local polynomial over the pre–post difference of the density function. Columns 2 to 4 report the covariate linearity regression for our main specification with age, earnings and female dummy as the dependent variable.