



Innovation Types and the Distribution of Turnover Growth Rates

Morina, Drini
Lucas, Henning
Heiden, Stefanie

Gottfried Wilhelm Leibniz University Hannover
Institute for Innovation Research,
Technology Management and
Entrepreneurship



Agenda

1. Introduction – Innovation and Firm Growth
2. Data – The Community Innovation Survey (CIS)
3. Methods – Quantile Generalized Additive Models (QGAM)
4. Results
5. Discussion
6. Outlook



1. Introduction: Is R&D beneficial?

- Policy makers assume that innovation drives firm growth
- But R&D is uncertain and risky (Bloom & Bloom, 2007)
- Empirically inconclusive until 2008 (Coad & Rao, 2008)
- Maybe solved by “*utilizing sophisticated statistical methods*” (Starbuck, 1971: p. 126)

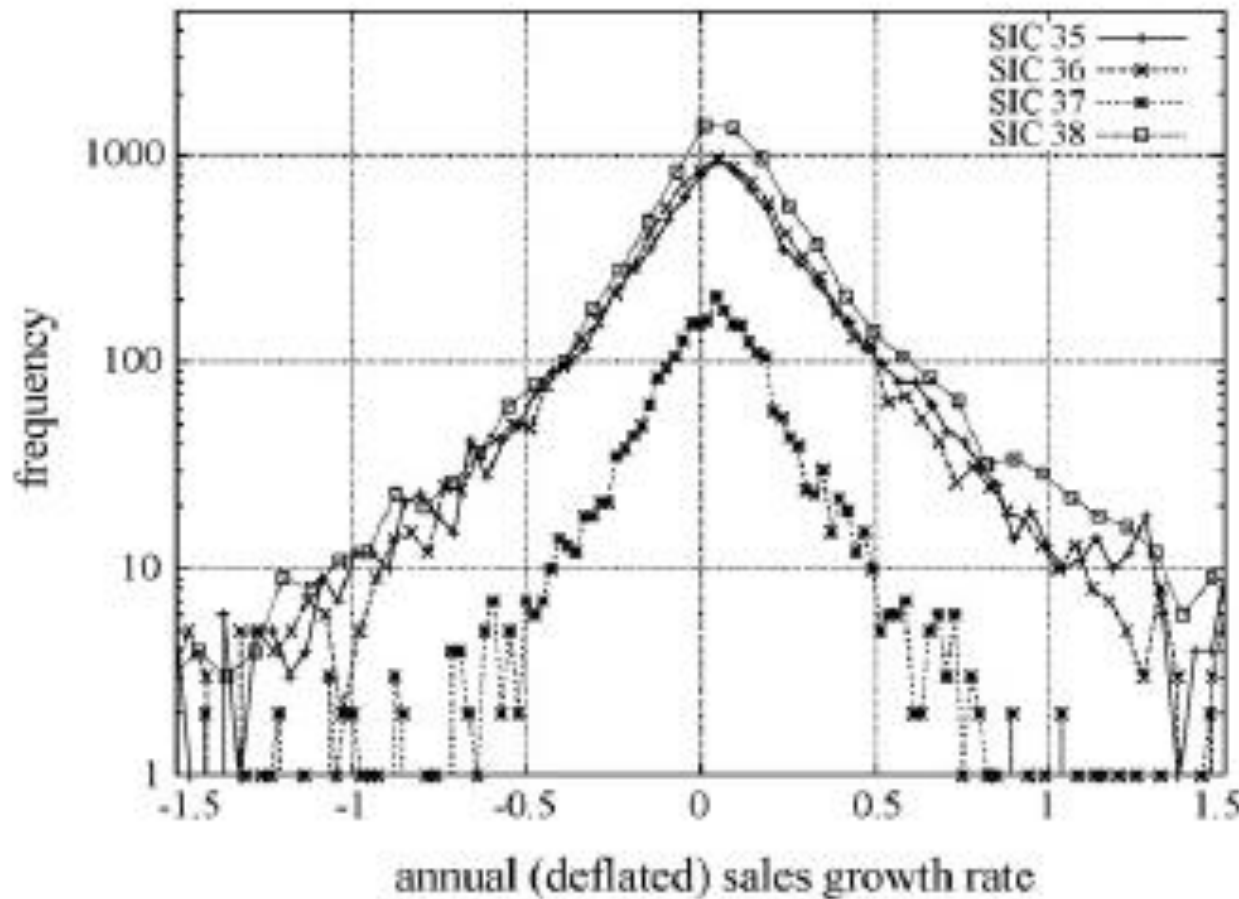


1. Introduction: Is R&D beneficial?

- Policy makers assume that innovation drives firm growth
- But R&D is uncertain and risky (Bloom & Bloom, 2007)
- Empirically inconclusive until 2008 (Coad & Rao, 2008)
- Maybe solved by “*utilizing sophisticated statistical methods*” (Starbuck, 1971: p. 126)
- Innovation affects the distribution of of firm growth (e.g. Coad & Rao, 2008; Kaiser, 2009; Hölzl, 2009; Falk, 2010; Mazucatto & Demirel, 2012; Segerra & Teruel, 2014; Bianchini et al., 2018; Moschella et al., 2019; Guarascio & Tamagni, 2019; Calvino, 2021 among others)



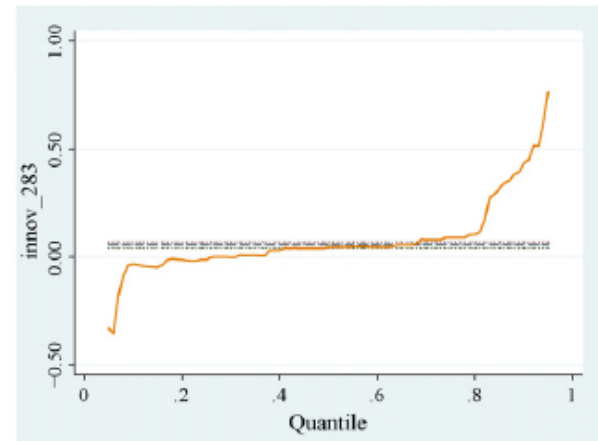
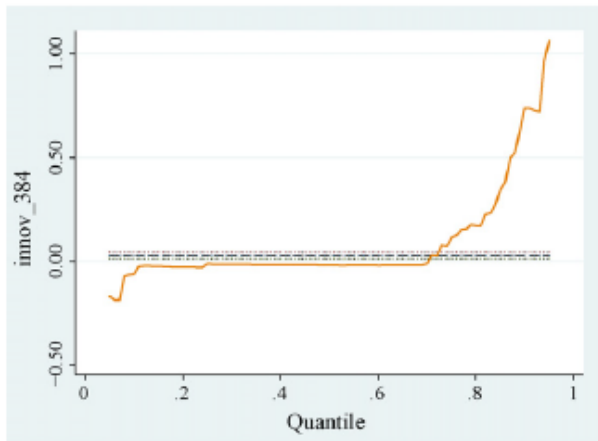
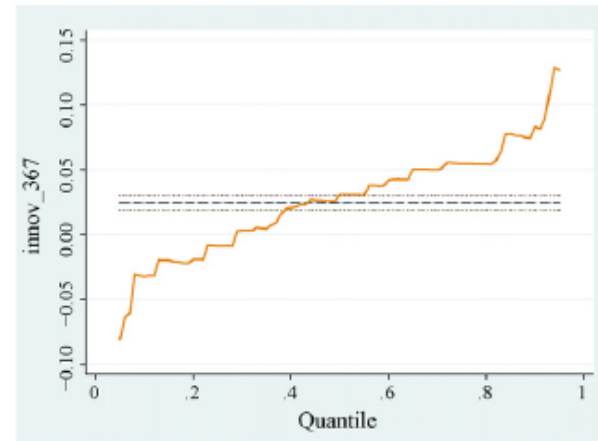
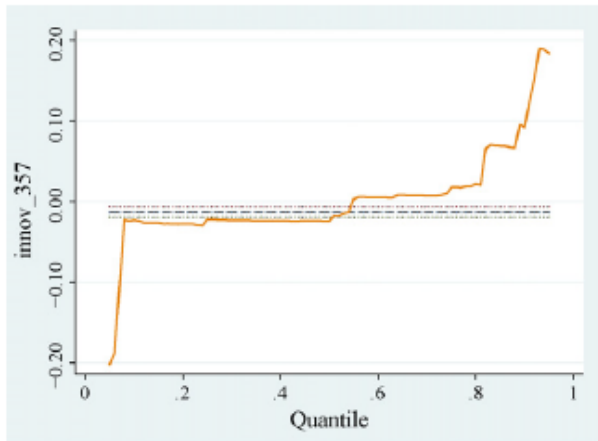
1. Introduction: Distribution of firm growth



Source:
Coad & Rao (2008)



1. Introduction: Quantile-specific effects



Source:
Coad & Rao (2008)



2. Data: Source

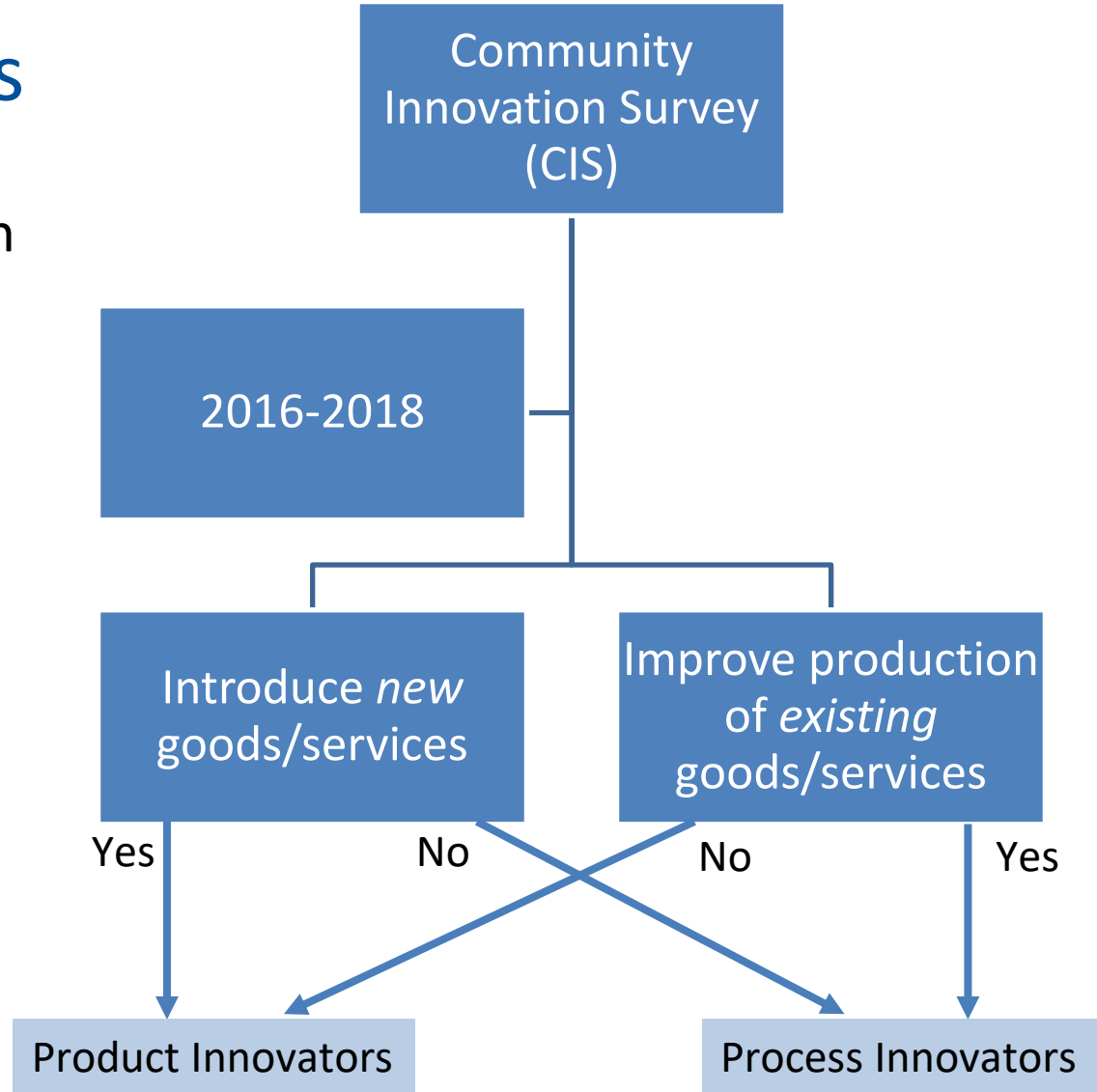
- The Community Innovation Survey (CIS):
 - Series of firm surveys (in two-year waves)
 - Ensured reliability and validity
 - Detailed data on innovation activity
 - Around 70.000 observations of 241 variables
- Our subsample:
 - Period 2016 – 2018 (short run)
 - Industries: NACE codes 13–96





2. Data: Variables

- Subsample construction by *Innovation Type*
- Variables:
 - Response variable:
 - Turnover growth denoted by $G_{i,t}$
 - Predictor variable:
 - R&D intensity denoted by $RDI_{i,t}$





2. Data: Description

- The Community Innovation Survey (CIS):
 - Product innovation subsample:

Variable	mean	median	standard deviation	<i>n</i> (before cleaning)	<i>n</i> (after cleaning)
Turnover growth	0.47	0.1	0.26	7910	3001
R&D intensity	0.15	0.03	0.49	7910	3001

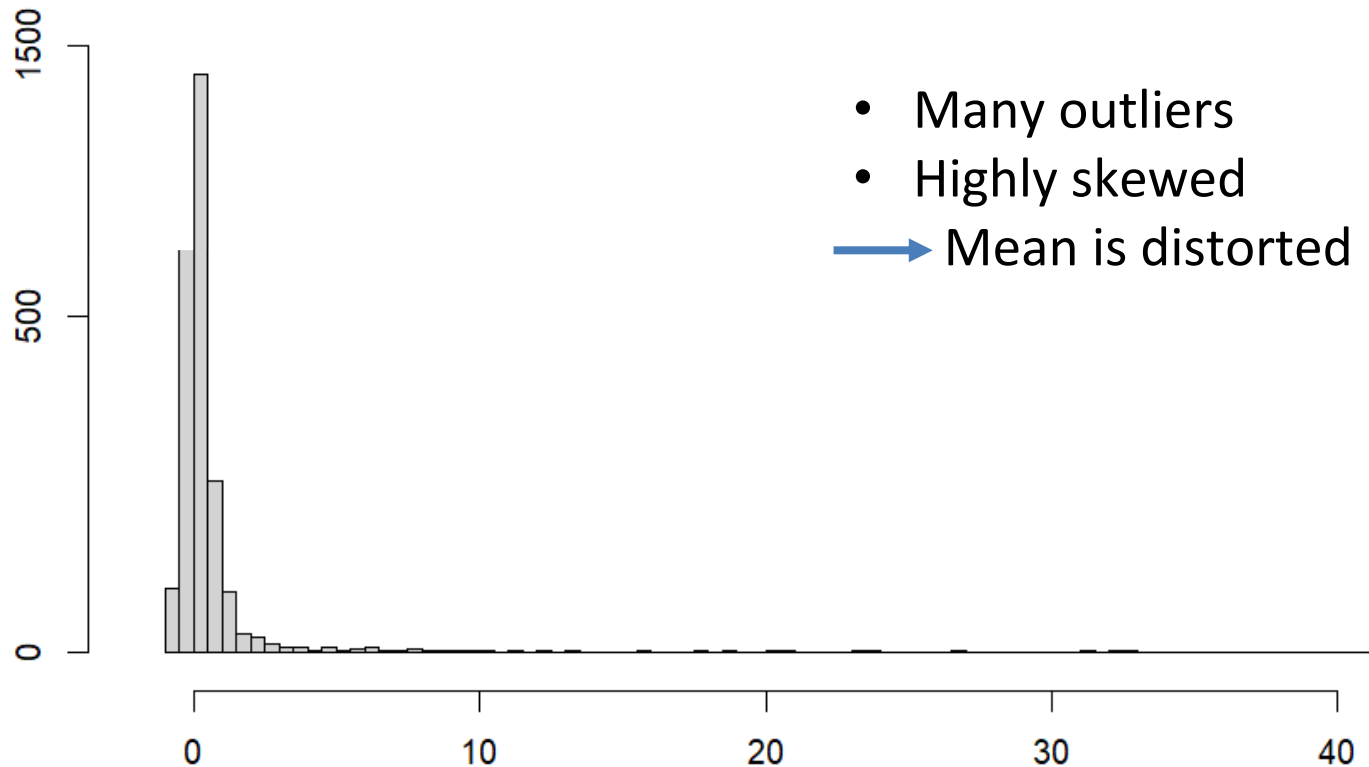
- Process innovation subsample:

Variable	mean	median	standard deviation	<i>n</i> (before cleaning)	<i>n</i> (after cleaning)
Turnover growth	0.50	0.1	3.3	9019	1847
R&D intensity	0.20	0.02	1.05	9019	1847



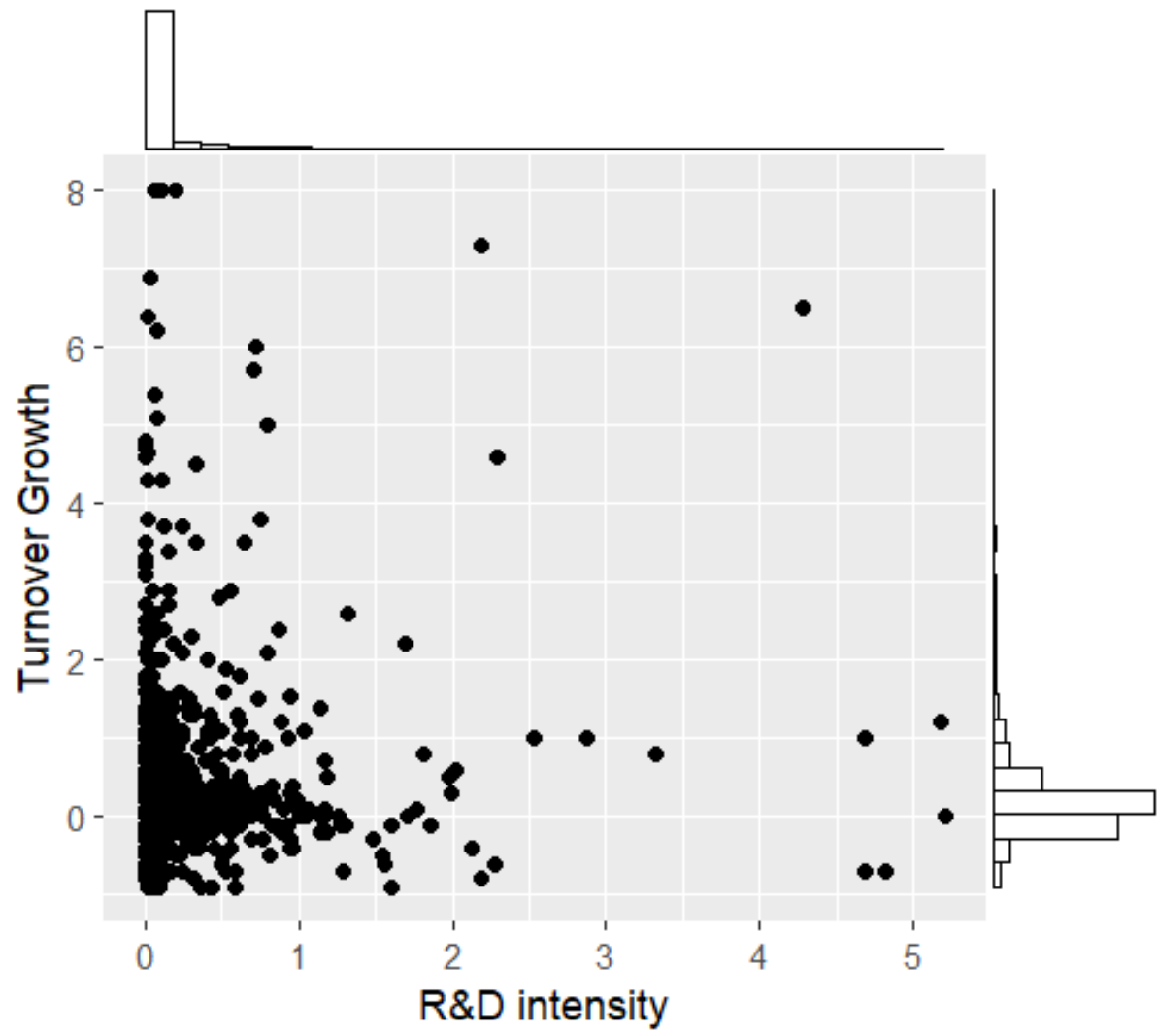
2. Data: Dependent Variabel

- Distribution of Turnover Growth



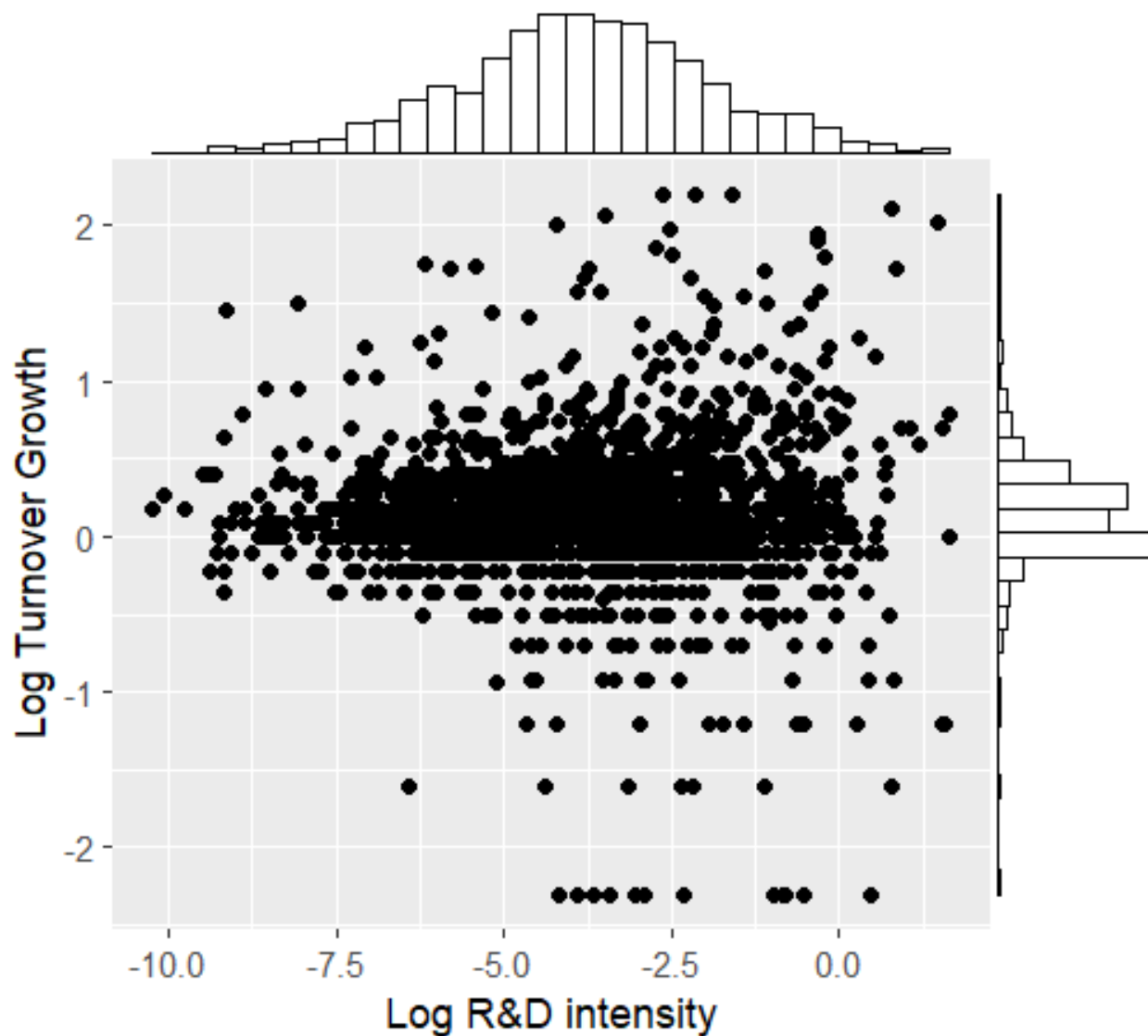


2. Data



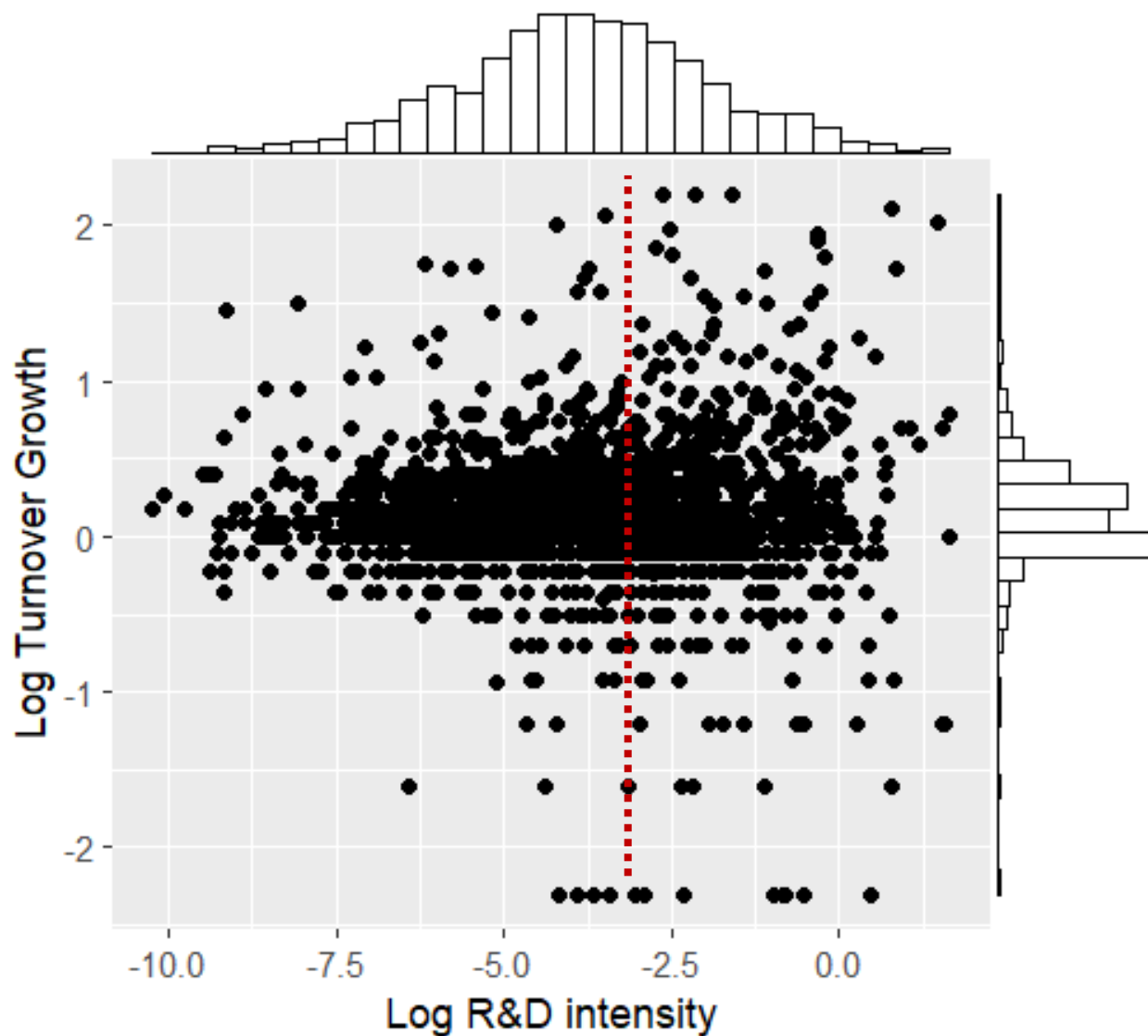


2. Data





2. Data



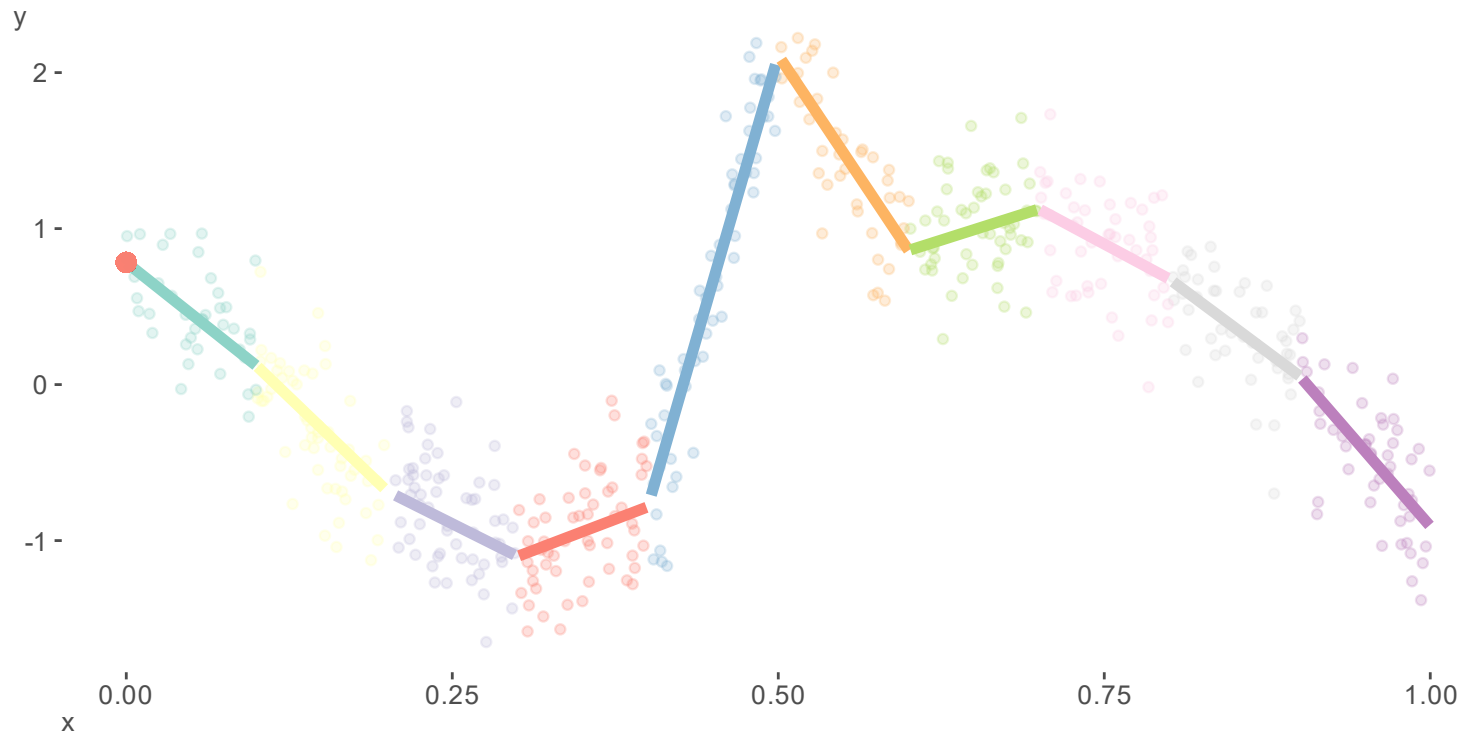


3. Methods: Non-linear Regression

- Fitting by
 - Polynomial regression
 - Try and error: $y = \beta_0 + \beta_1 x$
 $y = \beta_0 + \beta_1 x + \beta_2 x^2$
 $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$
 - Spline-based regression
 - Series of polynomial segments (basis function expansions)
 - Connected at so called *Knots*

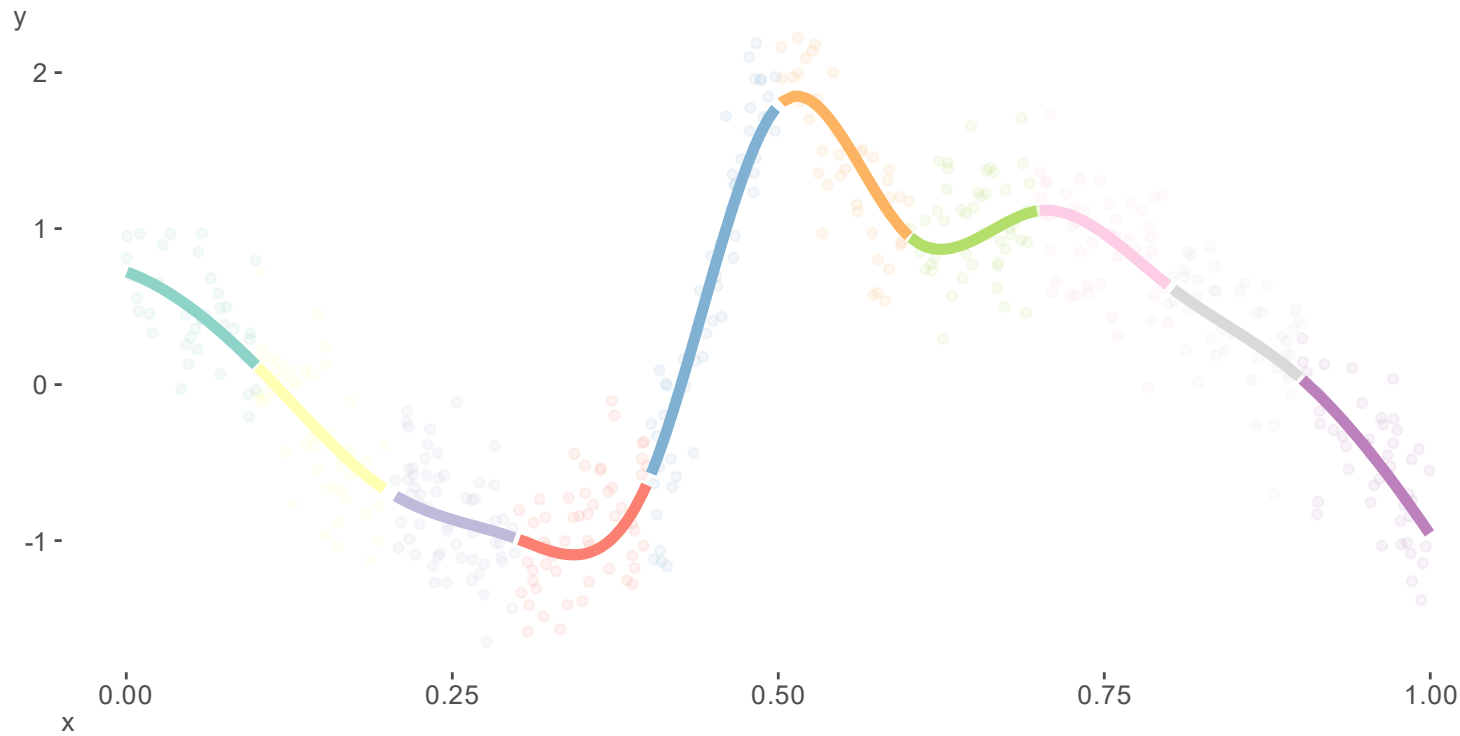


3. Methods: Spline-based regression





3. Methods: Spline-based regression





3. Methods: Non-linear regression

- Fitting by
 - Polynomial regression
 - Try and error: $y = \beta_0 + \beta_1 x$
 $y = \beta_0 + \beta_1 x + \beta_2 x^2$
 $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$
 - Spline-based regression
 - Series of polynomial segments (basis function expansions)
 - Connected at so called *Knots*
 - Generalized additive model (GAM, Hastie and Tibshirani, 1990)
 - Automatically determines smoothing parameters



3. Method

- Generalized Additive Models, GAMs (Hastie and Tibshirani, 1990)
 - Response variable Y depends on predictor variables x_i
 - The expected value of y_i is related to x_i via a structure such as

$$h(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_m(x_m)$$

- h is a link function
- f_i may be specified parametrically or non-parametrically



3. Method

- Generalized Additive Models, GAMs (Hastie and Tibshirani, 1990)
 - Response variable Y depends on predictor variables x_i
 - The expected value of y_i is related to x_i via a structure such as

$$h(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_m(x_m)$$

- h is a link function
 - f_i may be specified parametrically or non-parametrically
 - We use *thin plate regression splines*
 - Estimate the degrees of freedom
 - Estimate penalty to smoothness
- } Both automatically based on GCV or REML



3. Method

- Generalized Additive Models, GAMs (Hastie and Tibshirani, 1990)
 - Response variable Y depends on predictor variables x_i
 - The expected value of y_i is related to x_i via a structure such as

$$h(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_m(x_m)$$

- h is a link function
 - f_i may be specified parametrically or non-parametrically
 - We use thin plate regression splines
 - Estimate the degrees of freedom
 - Estimate penalty to smoothness
- } Both automatically based on GCV or REML
- see Woods et al. (2017)



3. Methods

- Quantile Regression (Kroenker, 2005)
 - Estimates conditional quantiles of the response variable (instead of the conditional mean):

$$Q_{Y|X}(q) = X\beta_q \text{ for the } q\text{th quantile}$$

- Quantile Generalized Additive Models (QGAMs):
 - Developed by Fasiolo et al. (2017)
 - Framework for additive quantile regression
 - Uses probability density similar to the pinball loss (The „Extended Log-F“ density)
 - Inference method is *Restricted Maximum Likelihood* (REML)
 - R software package „qgam“ (Fasiolo et al. 2017)



3. Methods

- Empirical model:

$$h(G_{i,t}) = \beta_0 + f(RDI_{i,t-1}) + \beta_1 G_{i,t-1} + \beta_2 X_{i,t-1} + \beta_3 IND_i + \beta_4 YEAR_t + u_{i,t}$$

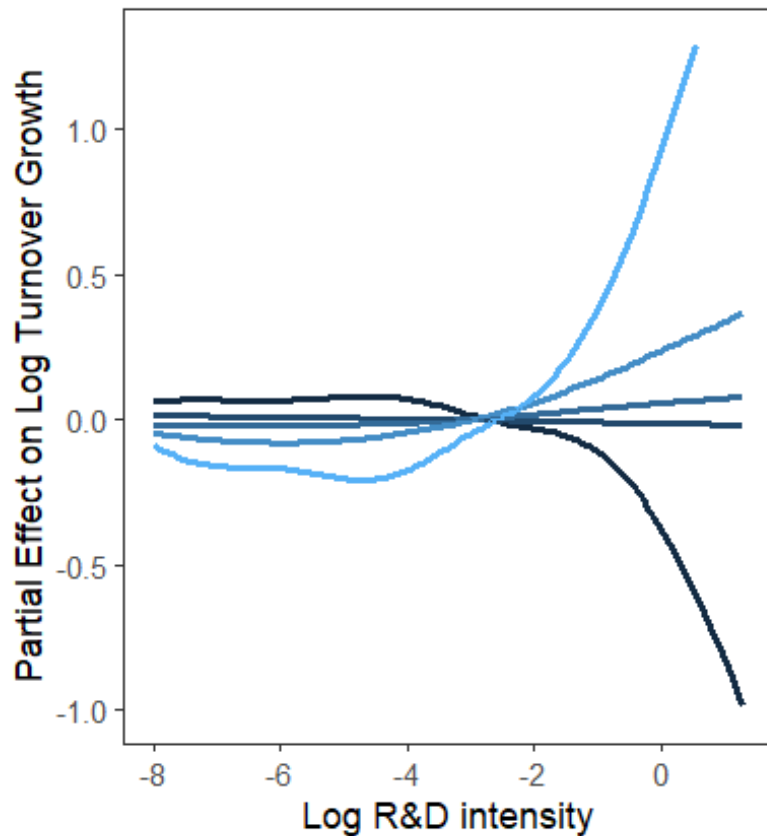
where

- $G_{i,t}$ is log-transformed turnover growth
- $RDI_{i,t-1}$ is log-transformed R&D intensity
- $X_{i,t-1}$ is a set of control variables (number of employees, age)
- IND_i and $YEAR_t$ are fixed effects
- $u_{i,t}$ is an error term

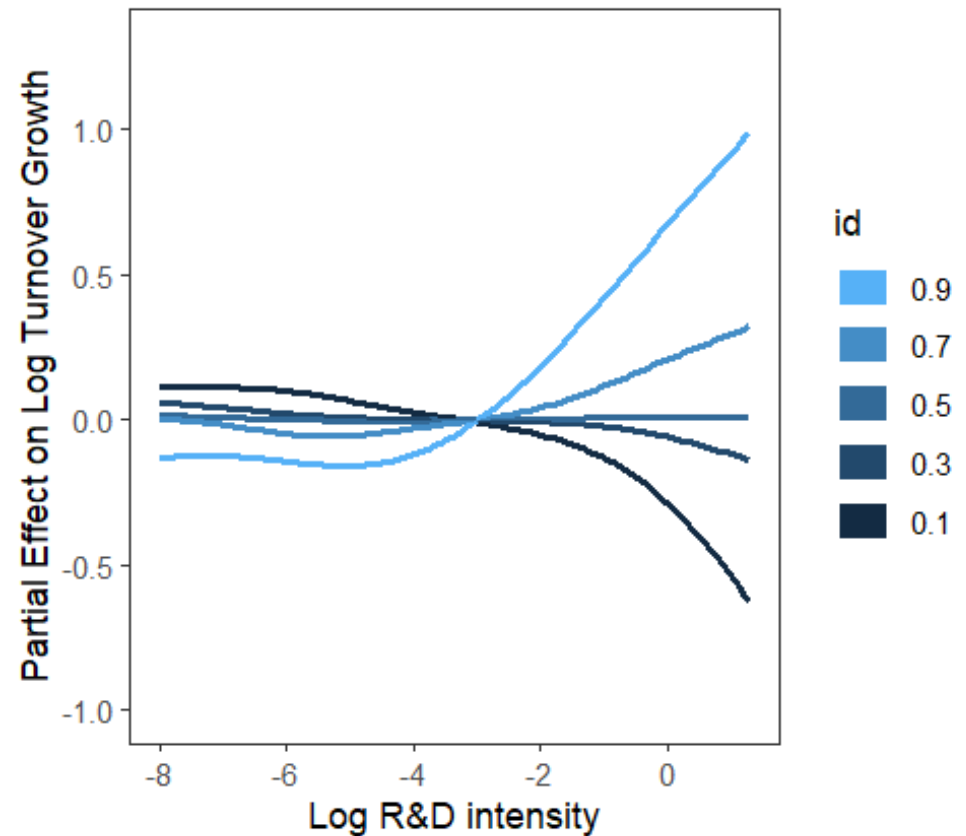


4. Results: Partial effect of R&D intensity

Product Innovation



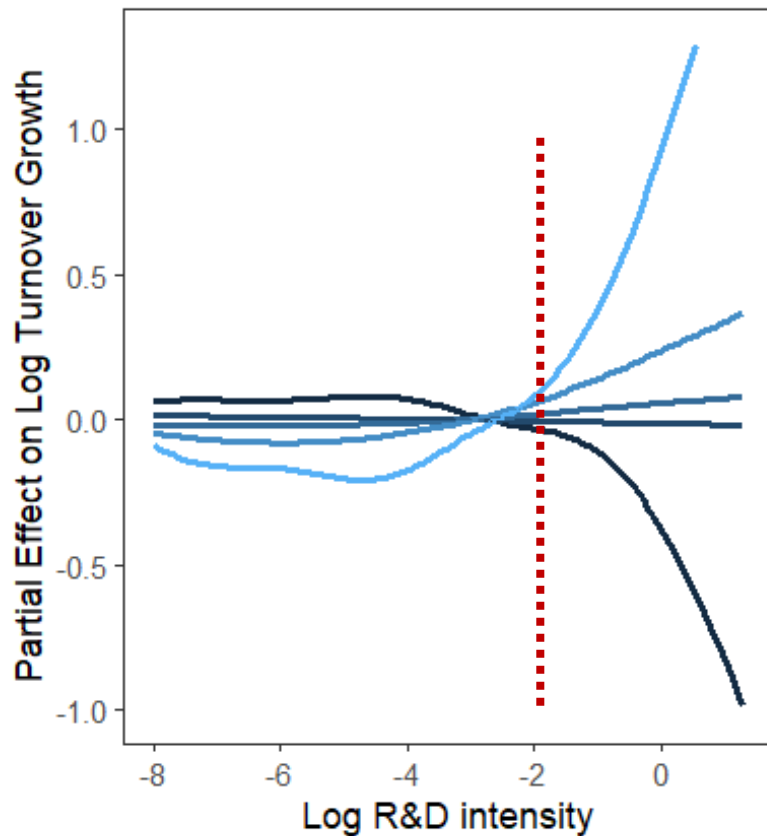
Process Innovation



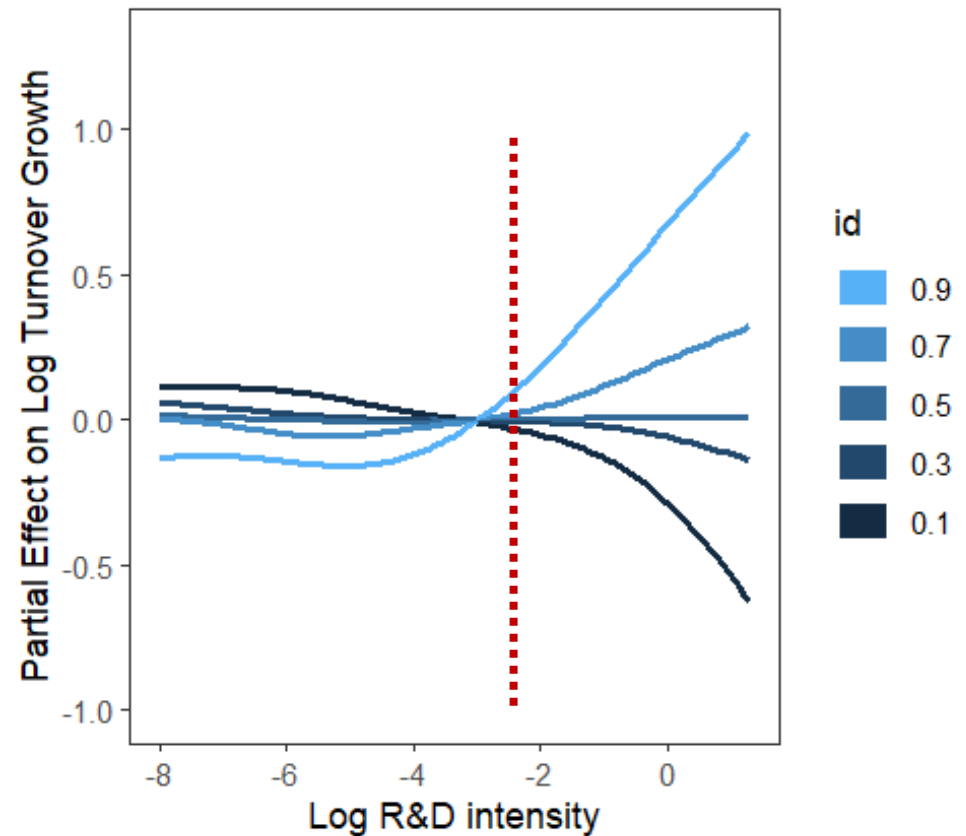


4. Results: Partial effect of R&D intensity

Product Innovation

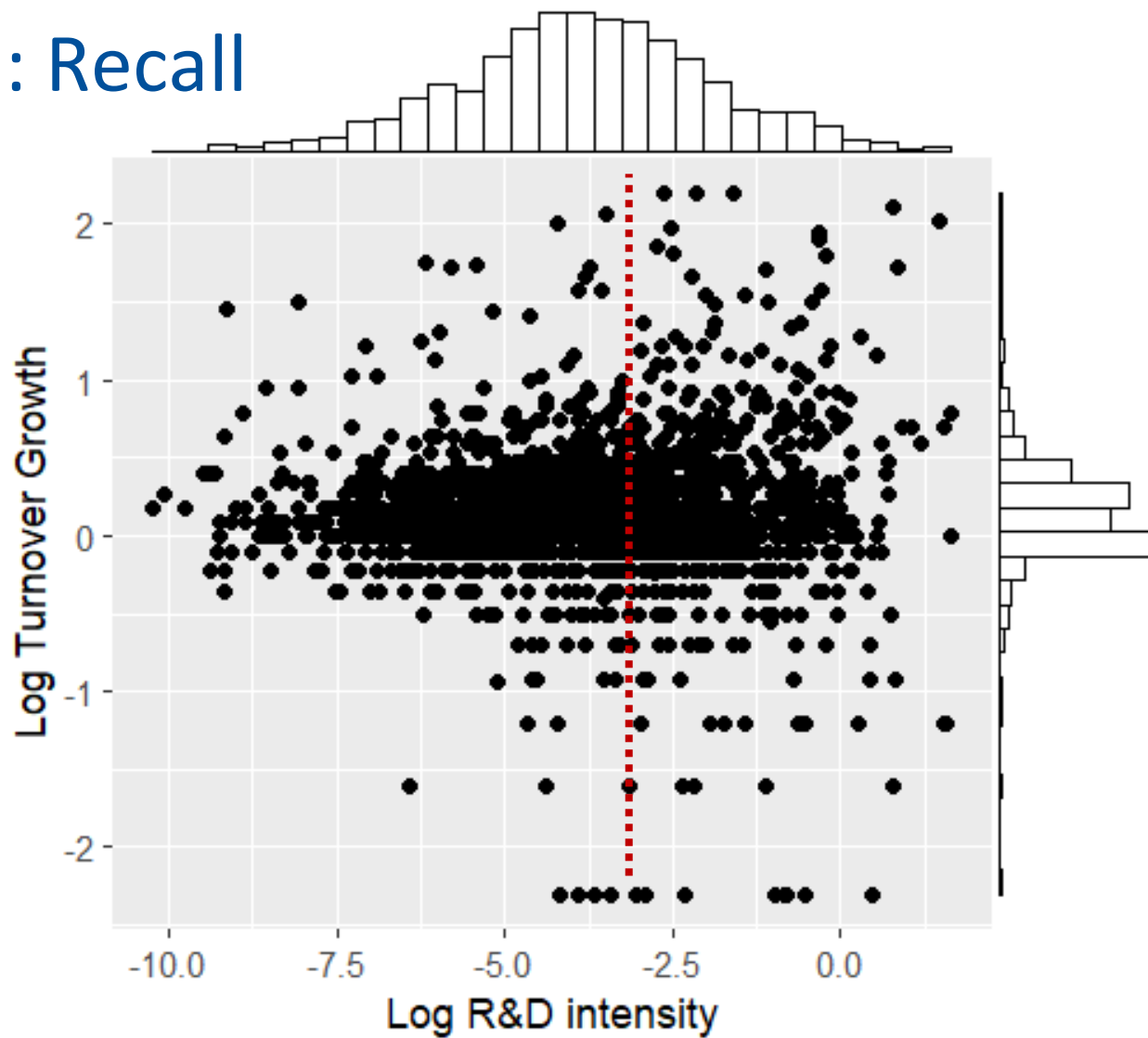


Process Innovation





2. Data: Recall





5. Discussion

- Main findings:
 - Link is indeed non-linear
 - Relationship depends on R&D intensity level
- Indicating that
 - linear quantile regression results are skewed
 - R&D can also be beneficial for firms in „bad conditions“



6. Outlook

- Our results describe the *very short run*
- Further analysis could
 - Employ different time lags and timing
 - Ensure causality (e.g. by instrumental variables)
 - Estimate Industry-specific models



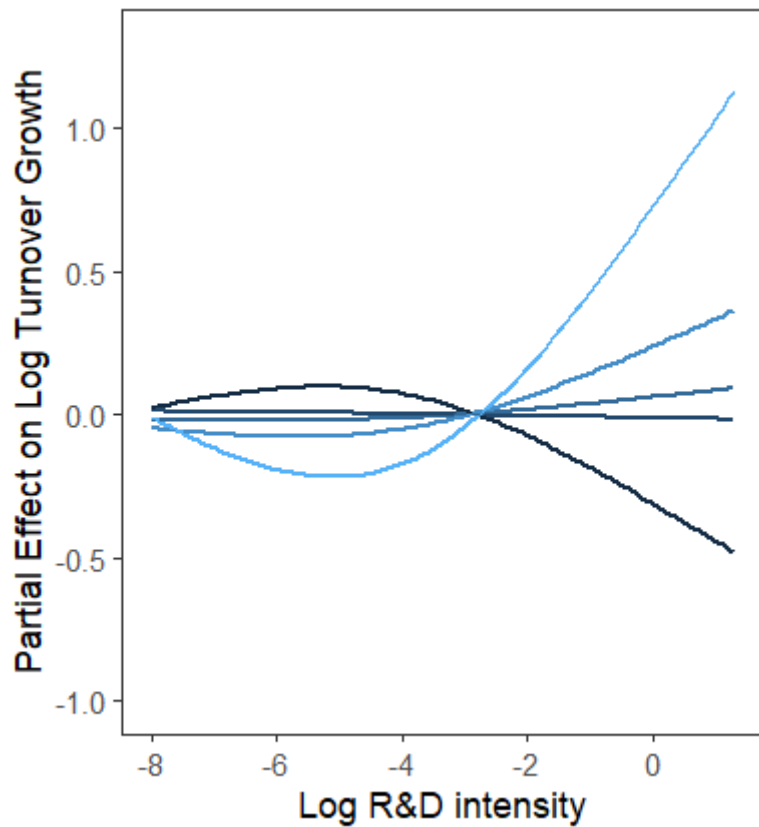
Thank you for your attention

Any questions?





Product Innovation



Process Innovation

