UNIVERSITEIT GENT

FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

Starting an R&D project under uncertainty

Sabien Dobbelaere
Roland Iwan Luttens
Bettina Peters

May 2009

2009/585

D/2009/7012/37

Starting an R&D project under uncertainty*

Sabien Dobbelaere† Roland Iwan Luttens‡ Bettina Peters§

May 2009

Abstract

We study a two-stage R&D project with an abandonment option. Two types of uncertainty influence the decision to start R&D. Demand uncertainty is modelled as a lottery between a proportional increase and decrease in demand. Technical uncertainty is modelled as a lottery between a decrease and increase in the cost to continue R&D. We relate differences in uncertainty to differences in risk premia. We deduct testable hypotheses on the basis of which we empirically analyze the impact of uncertainty on the decision to start an R&D project. Using data for about 4000 German firms in manufacturing and services (CIS IV), our model predictions are strongly confirmed.

JEL classification : D21, D81, L12, O31.
Keywords : Investment under uncertainty, R&D, demand uncertainty, technical uncertainty, entry threat.

1 Introduction

The decision to start a Research and Development (R&D) project is one of the most challenging firm decision problems. R&D projects usually take time to complete, their investments are irreversible and therefore represent sunk costs and above all, they are highly uncertain. Models of investment decisions in an uncertain environment have permeated different parts of the investment literature, ranging from the neoclassical theory of investment (e.g. Jorgenson, 1963; Eiser and Nadiri, 1968), over the real options approach (e.g. Dixit and Pindyck, 1994) to oligopolistic settings explicitly accounting for strategic interaction (e.g. Grenadier, 2000). While most of the investment literature considers uncertainty in input and output prices, for R&D projects other sources of uncertainty are possible. Grenadier and Weiss (1997) and Farzin et al. (1998) focus on uncertainty in technological progress. Besides input cost uncertainty, Pindyck (1993) also considers a second type of cost uncertainty, namely technical uncertainty. It implies that, although the input prices are known, the firm does not know at the beginning the amount of

*This research initiated when the first author was visiting the Centre for European Economic Research (ZEW) whose hospitality is greatly acknowledged. We are grateful to Paul Belleflamme, Boris Lokshin, Stephane Robin, LiWei Shi and participants at the International Industrial Organization Conference (Washington, 2008), the ZEW Conference on the Economics of Innovation and Patenting (Mannheim, 2008), the Danish Research Unit for Industrial Dynamics (DRUID) Conference (Copenhagen, 2008), the International Schumpeter Society Conference (Rio de Janeiro, 2008) and the European Association for Research in Industrial Economics (EARIE) Conference (Toulouse, 2008) for helpful comments and suggestions. All remaining errors are ours.

†Ghent University, Free University Amsterdam, Tinbergen Institute, IZA Bonn. Postdoctoral Fellow of the Research Foundation - Flanders (FWO). Corresponding author: sdobbelaere@feweb.vu.nl

‡SHERPPA, Ghent University and CORE, Université Catholique de Louvain. Postdoctoral Fellow of the Research Foundation - Flanders (FWO).

§Centre for European Economic Research (ZEW).

time, effort and materials ultimately needed to complete the project. Importantly, this type of cost uncertainty can only be solved by starting the R&D project. Market uncertainty is related to the future value of the innovation which is strongly determined by market demand (Tyagi, 2006). For example, if firms have successfully developed the new product or production technology, uncertainty still exists about market acceptance and hence innovation rents.

Despite considerable empirical evidence on the impact of uncertainty on firm-level investment (e.g. Dorfman and Heien, 1989; Leahy and Whited, 1996; Guiso and Parigi, 1999; Henley et al., 2003; Bloom et al., 2007), there is little empirical research that investigates the role of *different* types of uncertainty on R&D decisions. In this paper, we develop a generalized version of the model of Lukach et al. (2007) that contains many aspects of real-life R&D decisions within a net present value (NPV) framework. Besides entry threat, competition and multi-stage R&D, our model includes demand uncertainty as well as technical uncertainty. We deduct testable hypotheses on the basis of which we empirically analyze the impact of uncertainty on the decision to start an R&D project. The uniqueness of our data lies in the availability of proxies for demand and technical uncertainty as well as perceived entry threat. In sum, the main goal of our analysis is to quantify the effect of demand and technical uncertainty on the likelihood of undertaking R&D.

We model an R&D project as a two-stage game where the incumbent must decide at the first stage to start and at the second stage to continue R&D. The decision to start is influenced by on the one hand demand uncertainty modelled as a lottery between a proportional increase (=good state) and decrease (=bad state) in demand and on the other hand technical uncertainty modelled as a lottery between a decrease (=good state) and increase (=bad state) in the cost to continue R&D. We say that a lottery becomes more divergent when the difference between the outcomes of the lottery increases. We derive under which lottery probabilities more divergent demand and supply lotteries positively or negatively affect the decision to start R&D. For empirical testing, we use data from the fourth Community Innovation Survey (CIS IV) in Germany for about 4000 firms to explain actual and planned R&D investments. Our main results, strongly confirming our model predictions, are that for firms facing lotteries where the good state is more likely to prevail (i) a 10% increase in the degree of divergence of the demand lottery increases the likelihood of undertaking R&D by 1.4% and (ii) a change from a low to a high degree of divergence of the supply lottery increases the likelihood of undertaking R&D by 23.3%. For firms facing a demand lottery where the bad state is more likely to prevail, an increase in the degree of divergence of the demand lottery decreases the probability of undertaking R&D significantly.

We believe that our article contributes to the current state of research on both the theoretical and empirical side. From a theoretical point of view, we model uncertainty as a lottery rather than a stochastic process (e.g. Dasgupta and Stiglitz, 1980; Weeds, 2002) to capture the uncertainty resolving nature of multi-stage R&D. But unlike Lukach et al. (2007), who only consider supply lotteries and quantify differences in uncertainty using the variance-based concept of a mean preserving spread, our analysis allows to study a broader class of both demand and supply lotteries as we relate differences in uncertainty to differences in risk premia (Pratt, 1964). Strongly embedded in expected utility theory, the use of a risk premium to quantify uncertainty is particularly suitable in a NPV framework since the risk premium and the NPV of an R&D project are calculated in a similar way. From an empirical point of view, we believe that exploiting firm heterogeneity in demand and supply lotteries credibly provides empirical evidence of the uncertainty-R&D investment relationship at the firm level. This belief is motivated by the observation that the results about the effect of uncertainty on investment, both quantitatively and qualitatively greatly vary across studies as soon as the analysis is performed using more aggregated data (e.g. Ferderer, 1993; Darby et al., 1999 using country data; Caballero and Pindyck, 1996; Ghosal and Loungani, 1996; Huizinga, 1993 using industry data) or taking less firm heterogeneity regarding uncertainty into account (see references on firm-level investment mentioned above).

The remaining part of the article is organized as follows. Section 2 provides a theoretical analysis of R&D decisions under uncertainty. The comparative statics of Section 3 allow us to derive testable hypotheses on the relation between demand and technical uncertainty and the decision to start R&D. Section 4 presents the empirical analysis. Section 5 concludes.

2 A theoretical analysis of R&D decisions under uncertainty

2.1 The model

The incumbent is producing a homogeneous good at unit cost $c \in [0, P]$, where $P \in [0, 1]$ denotes the normalized output price. A potential entrant is endowed with a superior technology that, for simplicity, allows him to produce at a zero unit cost. He faces an entry cost equal to $\omega \in \mathbb{R}_{++}$. Upon entry, both firms engage in Bertrand competition.

We model an R&D project as a two-stage game where the incumbent must decide at the first (second) stage to start (continue) R&D. This captures more realistically R&D outcomes as a sequence of successive decisions rather than as a result of an irreversible one-shot decision. Furthermore, by allowing the incumbent to abandon the R&D project in the second stage, we are able to study the effect of an abandonment option on optimal investment decisions. In our model, two types of uncertainty, one on the demand side and one on the supply side, influence the decision to start. The incumbent has a time lead over the potential entrant. When the incumbent starts *and* continues R&D, he obtains the same superior technology as the potential entrant before the latter can enter the market. Figure 1 illustrates the game tree.

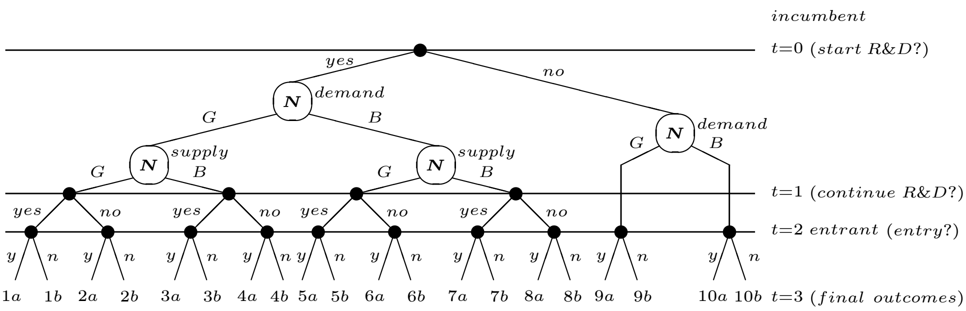

Figure 1 **Game tree.** At t=0, the incumbent decides whether to start R&D. Before t=1, nature (**N**) reveals the good/bad state (G/B) on the demand and supply side (the true state on the supply side is of no influence when the incumbent decides not to start R&D). At t=1, the incumbent decides whether to continue R&D. At t=2, the potential entrant, fully informed about the incumbent's decisions, decides whether to enter. At t=3, final outcomes are realized.

At time zero, the incumbent has to decide whether to start R&D at a *known* cost $I_0 \in \mathbb{R}_{++}$ but under an *unknown* state of the world. There are four possible states of the world, depending on the combination of a good/bad state on the demand and supply side. On the demand side, the good/bad state manifests itself as a proportional increase or decrease in demand, parameterized by $\theta \in [0,1]$. *A priori*, true demand is a lottery, i.e. the inverse market demand function $D(P,\theta)$ equals $(1+\theta)(1-P)$ with probability $p_\theta \in [0,1]$ and $(1-\theta)(1-P)$ with probability $(1-p_\theta)$. On the supply side, the good/bad state manifests itself as a decrease or an increase in a known cost $I_1 \in \mathbb{R}_{++}$ to continue R&D, parameterized by $\lambda \in [0, I_1]$. *A priori*, the true cost to continue R&D is a lottery, i.e. equal to $(I_1 - \lambda)$ with probability $p_\lambda \in [0,1]$ and $(I_1 + \lambda)$ with probability $(1-p_\lambda)$. We assume that all parameters are known beforehand and that both lotteries are independent. Before time one, nature (**N**) reveals the true state of the world.

At time one, the incumbent makes the decision whether to continue R&D.

At time two, the incumbent obtains the superior technology if he continued R&D. Having perfect knowledge about the incumbent's decisions, the potential entrant makes his entry decision. Upon a positive entry decision, the entrant enters the market, producing at a zero unit cost.

At time three, the final market structure is realized and the game ends.

2.2 Optimal entry decision and payoffs

The final market structure is never a duopoly.[1] Indeed, if the incumbent does not possess the superior technology, the potential entrant can push the incumbent out of the market by setting the price slightly under the incumbent's unit production cost, i.e. $P(c) = c - \varepsilon$ with $\varepsilon > 0$. However, entry is only optimal when monopoly profits are higher than or equal to the entry cost ω. If the potential entrant does not enter, the incumbent stays a monopolist who sets $P(c) = \frac{1+c}{2}$. The corresponding profits are $\pi(c) = \frac{(1-c)^2}{4}$ for all $c \in [0, P]$. If the incumbent does possess the superior technology, entry is never optimal. After all, the potential entrant knows that if he would enter, price equals marginal cost in equilibrium ($P(0) = 0$), and hence profits equal zero ($\pi(0) = 0$), which do not cover the entry cost.

In order to characterize the optimal R&D decisions of the incumbent, we present the incumbent's payoffs that correspond with the bottom row outcomes of Figure 1. We ignore the incumbent's monopolistic profits at $t = 0$ and $t = 1$ since they are the same for any outcome of the game and hence do not affect the incumbent's investment decision.

Under scenarios 1, 3, 5 and 7, the incumbent possesses the superior technology and entry is never optimal. Therefore, we only present the incumbent's payoffs under b, which equal:

$$\begin{array}{ll} 1b : (1+\theta)\,\pi(0) - I_0 - (I_1 - \lambda) & 5b : (1-\theta)\,\pi(0) - I_0 - (I_1 - \lambda) \\ 3b : (1+\theta)\,\pi(0) - I_0 - (I_1 + \lambda) & 7b : (1-\theta)\,\pi(0) - I_0 - (I_1 + \lambda) \end{array}$$

Under scenarios 2, 4, 6, 8, 9 and 10, the incumbent does not possess the superior technology. Hence, entry can be optimal. Therefore, we present the incumbent's payoffs valid under a (when entry is optimal ($\pi(0) \geq \omega$)) and b (when entry is not optimal ($\pi(0) < \omega$)).

$$\begin{array}{ll} 2a : -I_0 & 2b : (1+\theta)\,\pi(c) - I_0 \\ 4a : -I_0 & 4b : (1+\theta)\,\pi(c) - I_0 \\ 6a : -I_0 & 6b : (1-\theta)\,\pi(c) - I_0 \\ 8a : -I_0 & 8b : (1-\theta)\,\pi(c) - I_0 \\ 9a : 0 & 9b : (1+\theta)\,\pi(c) \\ 10a : 0 & 10b : (1-\theta)\,\pi(c) \end{array}$$

[1] Since Bertrand competition results in a monopoly in our model, it is not meaningful to distinguish between drastic and non-drastic innovation (contrary to Cournot competition).

2.3 Optimal R&D decisions

We determine the optimal R&D decisions of the incumbent by backward induction. We start at $t = 1$. We denote the four possible states of the world by $\{GG, GB, BG, BB\}$, where the first character reflects the good (G) or bad (B) demand state and the second character reflects the good (G) or bad (B) supply state. Let the incumbent's profit gain from innovation be $\Delta\pi = \pi(0) - \pi(c)$. This profit gain is higher when the entrant enters the market than when the entrant does not enter the market, since $\pi(c) = 0$ for the incumbent in the former case, whereas $\pi(c) > 0$ for the incumbent in the latter case. This immediately clarifies the strategic role of the entrant in our model compared to a monopoly model without entry threat. If the entry cost is low enough to make entry optimal, the incumbent gets additional benefits from investing in the superior technology. This strategic effect is known in the literature as Arrow's replacement effect (Arrow, 1962).

For each possible state of the world $s \in \{GG, GB, BG, BB\}$, we calculate Δ^s_{NPV}, i.e. the difference between the net present value (NPV) of continuing R&D and the NPV of not continuing R&D:

$$\begin{aligned}
\Delta^{GG}_{NPV} &= (1+\theta)\,\Delta\pi - (I_1 - \lambda)\\
\Delta^{GB}_{NPV} &= (1+\theta)\,\Delta\pi - (I_1 + \lambda)\\
\Delta^{BG}_{NPV} &= (1-\theta)\,\Delta\pi - (I_1 - \lambda)\\
\Delta^{BB}_{NPV} &= (1-\theta)\,\Delta\pi - (I_1 + \lambda).
\end{aligned}$$

The incumbent continues R&D if and only if this difference is positive under the true state of the world, taking the entrant's entry decision into account.

Optimal decision to continue R&D: *For each possible state of the world $s \in \{GG, GB, BG, BB\}$, the incumbent continues R&D if and only if $\Delta^s_{NPV} \geq 0$.*

Let $\boldsymbol{\psi} = (\psi_{GG}, \psi_{GB}, \psi_{BG}, \psi_{BB})$, where $\psi_s = 1$ when $\Delta^s_{NPV} \geq 0$ and $\psi_s = 0$ when $\Delta^s_{NPV} < 0$ for all $s \in \{GG, GB, BG, BB\}$, be the vector that comprises the optimal decision to continue R&D under every possible state of the world. Notice that $\Delta^{GG}_{NPV} \geq \Delta^s_{NPV} \geq \Delta^{BB}_{NPV}$ for $s \in \{GB, BG\}$. Therefore $\boldsymbol{\psi} \in \Psi = \{\,(1,1,1,1), (1,1,1,0), (1,1,0,0), (1,0,1,0), (1,0,0,0), (0,0,0,0)\,\}$.

At $t = 0$, for every $\boldsymbol{\psi} \in \Psi$, we calculate $\Delta^{\boldsymbol{\psi}}_{NPV}$, i.e. the difference between the NPV of starting R&D and the NPV of not starting R&D. For every $\boldsymbol{\psi} \in \Psi$, we determine the NPV of starting R&D by calculating the weighted sum of the incumbent's payoffs when starting R&D in every possible state of the world (using the probabilities of a good/bad state on the demand and supply side as weights). We determine the NPV of not starting R&D by calculating the weighted sum of the incumbent's payoffs when not starting R&D (using the probabilities of a good/bad state on the demand and supply side as weights). The NPV of not starting R&D is the same for every $\boldsymbol{\psi} \in \Psi$.

Hence, we get:

$$\begin{aligned}
\Delta^{(1,1,1,1)}_{NPV} &= p_\theta p_\lambda \left[(1+\theta)\,\pi(0) - I_0 - (I_1 - \lambda)\right]\\
&\quad + p_\theta\,(1-p_\lambda)\left[(1+\theta)\,\pi(0) - I_0 - (I_1 + \lambda)\right]\\
&\quad + (1-p_\theta)\,p_\lambda\left[(1-\theta)\,\pi(0) - I_0 - (I_1 - \lambda)\right]\\
&\quad + (1-p_\theta)\,(1-p_\lambda)\left[(1-\theta)\,\pi(0) - I_0 - (I_1 + \lambda)\right]\\
&\quad - \left[p_\theta\left[(1+\theta)\,\pi(c)\right] + (1-p_\theta)\left[(1-\theta)\,\pi(c)\right]\right]\\
&= p_\theta p_\lambda \Delta^{GG}_{NPV} + p_\theta\,(1-p_\lambda)\,\Delta^{GB}_{NPV} + (1-p_\theta)\,p_\lambda \Delta^{BG}_{NPV}\\
&\quad + (1-p_\theta)\,(1-p_\lambda)\,\Delta^{BB}_{NPV} - I_0.
\end{aligned}$$

From this, we calculate:

$$\begin{aligned}\Delta_{NPV}^{(1,1,1,0)} &= \Delta_{NPV}^{(1,1,1,1)} - (1-p_\theta)\,(1-p_\lambda)\,[(1-\theta)\,\pi(0) - I_0 - (I_1+\lambda)] \\ &\quad + (1-p_\theta)\,(1-p_\lambda)\,[(1-\theta)\,\pi(c) - I_0] \\ &= \Delta_{NPV}^{(1,1,1,1)} - (1-p_\theta)\,(1-p_\lambda)\,\Delta_{NPV}^{BB} \\ &= p_\theta p_\lambda \Delta_{NPV}^{GG} + p_\theta\,(1-p_\lambda)\,\Delta_{NPV}^{GB} + (1-p_\theta)\,p_\lambda \Delta_{NPV}^{BG} - I_0.\end{aligned}$$

Similarly, we get:

$$\begin{aligned}\Delta_{NPV}^{(1,1,0,0)} &= p_\theta p_\lambda \Delta_{NPV}^{GG} + p_\theta\,(1-p_\lambda)\,\Delta_{NPV}^{GB} - I_0, \\ \Delta_{NPV}^{(1,0,1,0)} &= p_\theta p_\lambda \Delta_{NPV}^{GG} + (1-p_\theta)\,p_\lambda \Delta_{NPV}^{BG} - I_0, \\ \Delta_{NPV}^{(1,0,0,0)} &= p_\theta p_\lambda \Delta_{NPV}^{GG} - I_0, \\ \Delta_{NPV}^{(0,0,0,0)} &= -I_0.\end{aligned}$$

Clearly, $\Delta_{NPV}^{(0,0,0,0)} < 0$ and the incumbent does not start R&D.

The incumbent starts R&D if and only if there exists a positive $\Delta_{NPV}^{\boldsymbol{\psi}}$ for $\boldsymbol{\psi} \in \Psi \backslash \{\,(0,0,0,0)\,\}$. Note that these $\Delta_{NPV}^{\boldsymbol{\psi}}$'s cannot be ordered. For example, take $\Delta_{NPV}^{(1,1,1,1)}$ and $\Delta_{NPV}^{(1,1,1,0)}$. We can write $\Delta_{NPV}^{(1,1,1,1)} = \Delta_{NPV}^{(1,1,1,0)} + (1-p_\theta)\,(1-p_\lambda)\,\Delta_{NPV}^{BB}$. If $\Delta_{NPV}^{BB} > 0$, then $\Delta_{NPV}^{(1,1,1,1)} > \Delta_{NPV}^{(1,1,1,0)}$ and it is possible to have $\Delta_{NPV}^{(1,1,1,1)} > 0$, while $\Delta_{NPV}^{(1,1,1,0)} < 0$. On the other hand, if $\Delta_{NPV}^{BB} < 0$, then $\Delta_{NPV}^{(1,1,1,1)} < \Delta_{NPV}^{(1,1,1,0)}$ and it is possible to have $\Delta_{NPV}^{(1,1,1,1)} < 0$, while $\Delta_{NPV}^{(1,1,1,0)} > 0$. A similar argument can be made for any other comparison.

Therefore, let $\Phi = \max\{\Delta_{NPV}^{(1,1,1,1)}, \Delta_{NPV}^{(1,1,1,0)}, \Delta_{NPV}^{(1,1,0,0)}, \Delta_{NPV}^{(1,0,1,0)}, \Delta_{NPV}^{(1,0,0,0)}\}$.

Optimal decision to start r&d: *The incumbent starts R&D if and only if* $\Phi \geq 0$.

3 Comparative statics

3.1 Relating divergence to uncertainty

In this section, we investigate how changes in demand and supply lotteries affect the incumbent's decision to start R&D. We therefore assume that entry is not optimal, because if entry were optimal, the entrant would drive the incumbent out of the market (cfr. Section 2.2). Throughout the remaining analysis, we use the following terminology. A lottery is defined to be *favorable (unfavorable)* if the probability of the good state is higher than or equal to (lower than) the probability of the bad state. In comparing two lotteries, a lottery is defined to be *more favorable (more unfavorable)* than another lottery if the probability of the good state of the former is higher (lower) than the probability of the good state of the latter. However, we do not only distinguish between lotteries in terms of probabilities but also in terms of outcomes. In comparing two lotteries with equal probabilities, a lottery is defined to be *more divergent (less divergent)* than another lottery if the difference between the good and the bad state is larger (smaller) in the former than in the latter. In our model, the degree of divergence depends on θ and λ: a demand (supply) lottery becomes more divergent than another demand (supply) lottery when, *ceteris paribus*, θ (λ) increases and a demand (supply) lottery becomes less divergent than another demand (supply) lottery when, *ceteris paribus*, θ (λ) decreases.

Let us first explain how a change in the degree of divergence of the demand (supply) lottery relates to a change in demand (technical) uncertainty. In this paper, we opt to quantify a change in uncertainty by a change in the *risk premium*. We define the risk premium of a demand (supply) lottery as the amount of money the incumbent is willing to pay (or has to receive) to avoid undergoing the lottery. In our model, it equals the difference between obtaining demand equal to $1 - P$ (facing the cost I_1 of continuing R&D) and the expected outcome of undergoing the demand (supply) lottery. One lottery is more uncertain than another lottery if the risk premium of the former is higher than the risk premium of the latter. It is clear that, when comparing a favorable lottery with an unfavorable lottery, the former is less uncertain than the latter irrespective of the degree of divergence of both lotteries. After all, the risk premium of a favorable lottery is negative whereas the risk premium of an unfavorable lottery is strictly positive. Furthermore, favorable lotteries with probability $\frac{1}{2}$ of the good/bad state (i.e. mean-preserving lotteries) are equally uncertain regardless of their degree of divergence since their risk premium always equals 0. When comparing two favorable, non mean-preserving lotteries with equal probabilities, the more divergent lottery corresponds to the less uncertain lottery as the risk premium becomes more negative. Similarly, when comparing two unfavorable lotteries with equal probabilities, the more divergent lottery corresponds to the more uncertain lottery as the risk premium becomes more positive. However, we cannot always conclude that the more divergent lottery corresponds to the less (more) uncertain lottery when the lotteries are favorable (unfavorable) but have unequal probabilities. Whether one lottery is more uncertain than another lottery depends on the trade-off between (i) exactly how much more/less favorable (unfavorable) one lottery is compared to the other and (ii) how much less/more (more/less) divergent one lottery is compared to the other.

Having established the relationship between divergence and uncertainty, it remains to show how a change in the degree of divergence affects the decision to start R&D.

3.2 Relating divergence to the decision to start R&D

In Section 2, we derive that it is optimal for the incumbent to start R&D if and only if $\Phi \geq 0$. This decision depends on the vector of parameters $(c, I_0, I_1, \theta, p_\theta, \lambda, p_\lambda)$. We now focus on how the effect of an increase in θ on the decision to start R&D depends, *ceteris paribus*, on p_θ. A completely similar reasoning, here omitted for reasons of parsimony, holds for how the effect of an increase in λ depends, *ceteris paribus*, on p_λ.

An increase from θ to θ' can, *ceteris paribus*, either have one of the three effects on the decision to start:

(i) a positive effect, i.e. when $\Phi(\theta) < 0$ and $\Phi(\theta') \geq 0$,
(ii) a negative effect, i.e. when $\Phi(\theta) \geq 0$ and $\Phi(\theta') < 0$ or
(iii) no effect, i.e. when $\Phi(\theta) < 0$ and $\Phi(\theta') < 0$ or $\Phi(\theta) \geq 0$ and $\Phi(\theta') \geq 0$.

Our approach aims at comparing $\Phi(\theta)$ and $\Phi(\theta')$ for any $\theta, \theta' \in [0, 1]$ where $\theta < \theta'$. We want to make explicit which effects are found for every $p_\theta \in [0, 1]$, while restricting the parameter space of $(c, I_0, I_1, \lambda, p_\lambda)$ as little as possible.

Ceteris paribus, it is impossible to compare $\Phi(\theta)$ and $\Phi(\theta')$ for any $\theta, \theta' \in [0, 1]$ where $\theta < \theta'$ and never find no effect, since $\Phi(\theta)$ is a continuous function in θ.

Our first two propositions are straightforward. Proposition 1 states that a more divergent demand lottery never positively affects the decision to start R&D when the demand lottery is most unfavorable. After all, for a demand lottery that excludes the good state to happen, an increase in θ corresponds to a worsening of the bad state, which never positively affects the decision to start. Proposition 2 states that a more divergent demand lottery never negatively affects the decision to start R&D when the demand lottery belongs to the set of favorable demand lotteries. After all, for demand lotteries where the good state is more likely to happen than the bad state, an increase in θ *a priori* increases the attractiveness of the R&D project and hence never affects the decision to start negatively. Both Propositions 1&2 hold over the complete parameter space of $(c, I_0, I_1, \lambda, p_\lambda)$. Remember that the same results are obtained by replacing p_θ and θ by p_λ and λ respectively. All proofs are relegated to Appendix A.

Proposition 1: *If* $p_\theta = 0$*, there does not exist a* $\theta, \theta' \in [0, 1]$*, where* $\theta < \theta'$*, such that* $\Phi(\theta) < 0$ *and* $\Phi(\theta') \geq 0$ *for all* $(c, I_0, I_1, \lambda, p_\lambda) \in [0, 1] \times \mathbb{R}^3_{++} \times [0, 1]$.

Proposition 2: *If* $p_\theta \in [\frac{1}{2}, 1]$*, there does not exist a* $\theta, \theta' \in [0, 1]$*, where* $\theta < \theta'$*, such that* $\Phi(\theta) \geq 0$ *and* $\Phi(\theta') < 0$ *for all* $(c, I_0, I_1, \lambda, p_\lambda) \in [0, 1] \times \mathbb{R}^3_{++} \times [0, 1]$.

It remains to show how more divergent demand lotteries affect the decision to start R&D when the demand lottery is unfavorable. From Proposition 1, the open question is from which value of p_θ on, it is possible to find a positive effect. Similarly, from Proposition 2, the question remains from which value of p_θ on, it is not possible to find a negative effect. In other words, we aim at extending Propositions 1&2 by respectively finding the minimal values $x \in (0, 1]$ and $y \in [0, \frac{1}{2}]$ such that the following results hold:

If $p_\theta \in [0, x)$, there does not exist a $\theta, \theta' \in [0, 1]$, where $\theta < \theta'$, such that $\Phi(\theta) < 0$ and $\Phi(\theta') \geq 0$.

If $p_\theta \in [y, 1]$, there does not exist a $\theta, \theta' \in [0, 1]$, where $\theta < \theta'$, such that $\Phi(\theta) \geq 0$ and $\Phi(\theta') < 0$.

The additional question becomes over which domains these extensions of Propositions 1&2 hold. Necessary conditions to obtain a positive (negative) effect are that, *ceteris paribus*, there exists a $\theta \in [0, 1]$ such that $\Phi(\theta) \geq (<) 0$. Obviously, these necessary conditions cannot be fulfilled over the complete parameter space of $(c, I_0, I_1, \lambda, p_\lambda)$. The intuition is that if the total cost of undertaking the R&D project —which depends on $(I_0, I_1, \lambda, p_\lambda)$— exceeds by far (is much smaller than) the total gain of the R&D project —which depends on (c, θ, p_θ)—, then Φ will always be negative (positive).

We impose two assumptions on the model, relating (in the absence of technical uncertainty) the cost of starting R&D to the cost of continuing R&D and the total cost of the R&D project to the profit gain. We assume that (i) the two cost components of R&D would be the same in the two periods when $\lambda = 0$ and (ii) the total cost of R&D would equal the profit gain of R&D when $\lambda = 0$.

Assumption 1: $I_0 = I_1 = I$.
Assumption 2: $I_0 + I_1 = \Delta\pi$.

Our results hold over the complete parameter space of (c, λ, p_λ). Indeed, in relating different demand lotteries to the decision to start the R&D project, we deliberately do not want to restrict the set of lotteries on the supply side. In other words, in determining x and y, we choose from the total set of supply lotteries (i) that particular lottery for which we obtain the smallest interval $p_\theta \in [0, x)$ of demand lotteries for which a more divergent demand lottery cannot positively affect the decision to start R&D and (ii) that particular lottery for which we obtain the smallest interval $p_\theta \in [y, 1]$ of demand lotteries for which a more divergent demand lottery cannot negatively affect the decision to start R&D. Larger intervals than $[0, x)$ and $[y, 1]$ would be obtained if one excluded these particular supply lotteries from the total set. All results also hold for any strictly positive value of c. When c equals zero, the incumbent never starts the R&D project. A completely similar exercise is performed to relate changes in λ and values of p_λ to changes in Φ under the complete parameter space of (c, θ, p_θ).[2]

Under Assumptions 1-2, we obtain Propositions 3a&3b for the minimal values x, v and Proposition 4 for the minimal values y, w respectively; all proofs are relegated to Appendix A:[3]

Proposition 3a: *Under Assumptions 1-2, if* $p_\theta \in [0, \frac{1}{4})$*, there does not exist a* $\theta, \theta' \in [0,1]$*, where* $\theta < \theta'$*, such that* $\Phi(\theta) < 0$ *and* $\Phi(\theta') \geq 0$ *for all* $(c, \lambda, p_\lambda) \in [0,1] \times \mathbb{R}_{++} \times [0,1]$.

Proposition 3b: *Under Assumptions 1-2, if* $p_\lambda \in [0, 0.28)$*, there does not exist a* $\lambda, \lambda' \in [0,1]$*, where* $\lambda < \lambda'$*, such that* $\Phi(\lambda) < 0$ *and* $\Phi(\lambda') \geq 0$ *for all* $(c, \theta, p_\theta) \in [0,1]^3$.

Proposition 4: *Under Assumptions 1-2, Proposition 2 is not extended: both* y *and* w *equal* $\frac{1}{2}$ *for all* $(c, \lambda, p_\lambda) \in [0,1] \times \mathbb{R}_{++} \times [0,1]$ *and for all* $(c, \theta, p_\theta) \in [0,1]^3$ *respectively.*

From Proposition 3a it follows that for the subset of unfavorable demand lotteries with $p_\theta \in [0, \frac{1}{4})$, a more divergent demand lottery never positively affects the decision to start R&D. From the determination of y in Proposition 4 we learn that for all unfavorable demand lotteries, we can not exclude that a more divergent demand lottery negatively affects the decision to start R&D. From Proposition 3b it follows that for the subset of unfavorable supply lotteries with $p_\lambda \in [0, 0.28)$, a more divergent supply lottery never positively affects the decision to start R&D. From the determination of w in Proposition 4 we learn that for all unfavorable supply lotteries, we cannot exclude that a more divergent supply lottery negatively affects the decision to start R&D.

Propositions 3a, 3b and 4 provide important additional insight in the relation between demand (supply) uncertainty and the decision to start R&D. Let us focus on demand uncertainty. Propositions 3a and 4 demonstrate that, for the set of unfavorable demand lotteries with $p_\theta \in [\frac{1}{4}, \frac{1}{2})$ and depending *inter alia* on the supply lottery the incumbent faces, an increase in demand uncertainty can either positively or negatively affect the decision to start R&D. Especially the fact that an increase in demand uncertainty of an unfavorable demand lottery can positively affect the decision to start R&D deserves some explanation. We obtain this result because of the abandonment option that the incumbent possesses. As we show in the proof of Proposition 3a in Appendix A, an increase in θ positively affects the decision to start R&D when $\Phi = \Delta_{NPV}^{(1,1,0,0)}$. Exactly in this case the R&D project is started under the assumption that the project will be completed when the good state on the demand side occurs (although it is more likely that the bad state on the demand side occurs since the demand lottery is unfavorable). In other words, the incumbent completely ignores the downside risk of the R&D project when the bad state on the demand side occurs exactly because it can abandon the project when this happens. Hence,

[2] More specifically, we aim at finding respectively the minimal values $v \in (0,1]$ and $w \in [0, \frac{1}{2}]$ such that the following results hold:

If $p_\lambda \in [0, v)$, there does not exist a $\lambda, \lambda' \in [0,1]$, where $\lambda < \lambda'$, such that
$\Phi(\lambda) < 0$ and $\Phi(\lambda') \geq 0$ for all $(c, \theta, p_\theta) \in [0,1]^3$.

If $p_\lambda \in [w, 1]$, there does not exist a $\lambda, \lambda' \in [0,1]$, where $\lambda < \lambda'$, such that
$\Phi(\lambda) \geq 0$ and $\Phi(\lambda') < 0$ for all $(c, \theta, p_\theta) \in [0,1]^3$.

[3] We performed a sensitivity analysis on Assumptions 1&2. We relax Assumption 1, setting $I_1 = aI_0$, where $a \in \mathbb{R}_{++}$. We find that the higher (lower) the cost of continuing R&D compared to the cost of starting R&D, the smaller (larger) the subset of unfavorable demand (supply) lotteries for which a more divergent demand (supply) lottery never positively affects the decision to start R&D. Furthermore, for all unfavorable demand/supply lotteries, we cannot exclude that a more divergent demand/supply lottery negatively affects the decision to start R&D, whatever the relative importance of the two cost components I_0 and I_1. We relax Assumption 2 by expressing the total cost of R&D as a proportion $b \in \mathbb{R}_{++}$ of the profit gain of R&D when $\lambda = 0$, i.e. $I_0 + I_1 = b\Delta\pi$. We find that the lower the profit gain of the R&D project compared to the total cost, the more favorable the demand/supply lottery has to become in order to start R&D. For reasons of parsimony, we omit the detailed results which are available upon request.

under the good state on the demand side, an increase in θ improves the profitability of the R&D project, which explains the result. If there were no abandonment option, an increase in demand uncertainty would never positively affect the decision to start R&D.[4]

4 An empirical analysis of the optimal decision to undertake R&D under uncertainty

4.1 Data

To test the propositions derived in the previous section, we use data from the 2005 official innovation survey in the German manufacturing and services industries which constitute the German part of the European-wide harmonized fourth Community Innovation Surveys (CIS IV).[5] The CIS data provide rich information on firms' innovation behavior. The target population consists of all legally independent firms with at least 5 employees and their headquarters located in Germany.[6] The survey is drawn as a stratified random sample and is representative of the corresponding target population. The stratification criteria are firm size (8 size classes according to the number of employees), industry (22 two-digit industries according to the NACE Rev.1 classification system) and region (East and West Germany). The survey is performed by mail and in 2005 data on 4776 firms were collected (*total sample*), corresponding to a response rate of about 20%.[7] In order to control for a response bias in the net sample, a non-response analysis was carried out collecting data on 4000 additional firms. A comparison shows that the innovation behavior of respondents and non-respondents does not differ significantly. The share of innovators is 63.9% in the former group and 62.2% in the latter group.[8]

For estimation purposes we exclude firms with incomplete data for any of the relevant variables (which are discussed in Section 4.2), ending up with a sample of 3681 firms. As illustrated in Table B.1 in Appendix B, our sample (*full sample*) reflects total-sample distributional characteristics very well and does not give any obvious cause for selectivity concerns. About 53.8% of the observed firms are in manufacturing.

4.2 Econometric model and testable hypotheses

Econometric model

In our theoretical model, the incumbent has to decide whether to undertake an R&D project which aims at obtaining the same superior production technology as the potential entrant.[9] The optimal decision to undertake R&D depends, *ceteris paribus*, on the degree of divergence of the demand and supply lotteries. Empirically, we operationalize this optimal decision as follows.

[4] If the incumbent is forced to complete the R&D project once the project is started, he only will start the project when $\Delta_{NPV}^{(1,1,1,1)} > 0$. Note that $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial \theta} = (2p_\theta - 1)\,\Delta\pi$ which is positive for all favorable demand lotteries and strictly negative for all unfavorable demand lotteries. This explains the result, given the relation between divergence and uncertainty (cfr. Section 3.1).

[5] The innovation surveys are conducted by the Centre for European Economic Research (ZEW), Fraunhofer Institute for Systems and Innovation Research (ISI) and infas Institute for Applied Social Sciences on behalf of the German Federal Ministry of Education and Research (BMBF). A detailed description of the data is given in Peters (2008).

[6] A firm is defined as the smallest combination of legal units operating as an organizational unit producing goods or services.

[7] This rather low response rate is not unusual for surveys in Germany and is due to the fact that participation is voluntary.

[8] The p-value of the Fisher-test on equal shares in both groups amounts to 0.108.

[9] In what follows, the notions *firm* and *incumbent* are used interchangeably.

Let y_i^* denote firm i's maximal difference between the NPV of undertaking R&D and the NPV of not undertaking R&D, which cannot be observed. Exploiting the firm heterogeneity in our unique dataset, we assume that for firm i this difference depends on θ_i and λ_i, some other observable characteristics summarized in the row vector $\mathbf{x}_i$ and unobservable factors captured by ϵ_i:

$$y_i^* = \alpha\theta_i + \gamma\lambda_i + \mathbf{x}_i\boldsymbol{\beta} + \epsilon_i \tag{1}$$

In Section 2.3, we derive that it is optimal for incumbent i to undertake R&D if and only if y_i^* is larger than or equal to zero:

$$y_i = \begin{cases} 1 & \text{if} \quad y_i^* \geq 0 \\ 0 & \text{if} \quad y_i^* < 0 \end{cases} \tag{2}$$

where y_i denotes the observed binary endogenous variable. We estimate equation (2) using the probit estimator.

Testable hypotheses

Table 1 gives the descriptive statistics of all variables used in the econometric analysis and Table B.2 in Appendix B provides detailed definitions of all variables. We proxy the observed binary endogenous variable (y_i) by two variables. The first proxy indicates whether the firm has performed R&D in the period 2002-2004 (*R&D*). Table 1 shows that 48% of the firms in the full sample undertook R&D projects. However, over the same period, we observe θ_i and λ_i, our measures reflecting uncertainty on the demand and the supply side respectively. Due to the short time-span of our data, we cannot use lagged values as instruments for the uncertainty measures to encounter the possible endogeneity problem. Instead, we employ as an alternative proxy an *expected* decision, indicating whether the firm plans to introduce a new production technology in the next year 2005 ($PROCESS$). We find that 46% of the firms in the full sample planned to introduce a process innovation.

<Insert Table 1 about here>

In our theoretical model, demand uncertainty stems from the two components in the lottery on the demand side: the degree of divergence (represented by θ) and the probability (p_θ) of facing a good demand state. The variable θ is measured by the average of the absolute percentage change in sales over the last two years 2002-2003 and 2003-2004 ($THETA$).[10] Table 1 reveals that the absolute change in sales was on average about 14 % in the last two years. In our benchmark estimations, we assume that p_θ is the same for all firms. Our dataset enables us to relax this assumption later on.

Similarly, technical uncertainty is represented by the two components in the lottery on the supply side: the degree of divergence (parameterized by λ) and the probability (p_λ) of facing a good supply state. For the full sample, λ can only be proxied by a dummy variable $LAMBDA1$. $LAMBDA1$ equals 1 if an innovation project was extended due to high innovation costs in the period 2002-2004. The motivation for using this information is that an unexpected delay of an innovation project is presumably associated with *unexpected* higher costs. Hence, $LAMBDA1$ partitions the set of firms into a subset of firms with a low degree of divergence and a subset of firms with a high degree of divergence. Around 19% of the firms belong to the latter. Alternatively, we use a second proxy for λ ($LAMBDA2$) which is defined as the absolute deviation between on the one hand the R&D expenditures for 2004 expected in 2003 and on the other hand the realized R&D expenditures in 2004. The virtue of this measure is that it more closely corresponds to the way we model λ in our theoretical analysis. The defect is that we can apply

[10] We implicitly assume that firms expect sales to stay constant over the short time-span under consideration.

it only to a subset of enterprises since we have to use the prior wave of the innovation survey to construct this variable.[11] However, this subsample is representative for the full sample as can be inferred from Table B.1 in Appendix B. The average absolute deviation between expected and realized innovation expenditure comes to 2.4 mill. Euro. The deviation turns out to be highly skewed. We therefore use a logarithmic transformation of this variable in the econometric analysis. In all our estimations, we assume that p_λ is the same for all firms. Our dataset does not allow to relax this assumption.

The probabilities p_θ and p_λ are determined as follows. To calculate p_θ using the full sample, we derive that 56.9% of the firms experienced a positive growth in sales between 2002 and 2003 and 61.8% between 2003 and 2004. No information is available to calculate p_λ from the full sample. However, we are able to determine p_λ from the subsample. More specifically, we observe that for 59.4% of the firms, realized innovation expenditure in 2004 turns out to be lower than expected in 2003. Given the representativeness of the subsample, we assume that the calculated p_λ is also valid for the full sample.

Assuming that p_θ and p_λ are the same for all firms and given that p_θ and p_λ are calculated to be larger than $\frac{1}{2}$, we postulate from Proposition 2 the following hypotheses.

Hypothesis 1: The probability of undertaking R&D does not decrease with a more divergent demand lottery.

Hypothesis 2: The probability of undertaking R&D does not decrease with a more divergent supply lottery.

In our theoretical model, the incumbent is challenged by a potential competitor. Our data reveal that about 91% of the firms perceive a threat of its own market position due to the potential entry of new competitors. In the estimations, we therefore control for potential entry by including 3 dummy variables indicating whether the firm perceives a high, medium or low threat.

We also control for the following factors found to be important in the literature. Two main determinants explaining innovation activities go back to Schumpeter (1942), who states that large firms in concentrated markets have an advantage in innovation. Therefore, we include firm size ($SIZE$) and market structure ($NUMCOMP$). Firm size is measured by the logarithm of the number of employees in 2003 and we expect a positive relationship. Market structure is captured by 3 dummy variables indicating the number of competitors. Schumpeter stresses a negative relationship between competition and innovation. His argument is that *ex ante* product market power on the one hand increases monopoly rents from innovation and on the other hand reduces the uncertainty associated with excessive rivalry. Recently, Aghion et al. (2005) find evidence for an inverted U-relationship between competition and innovation. For low initial levels of competition an escape-competition effect dominates (i.e. competition increases the incremental profits from innovating, and, thereby, encourages innovation investments) whereas the Schumpeterian effect tends to dominate at higher levels of competition.

The incentive to engage in R&D may further depend on the type of competition ($COMP$). We include 5 dummy variables indicating whether firms primarily compete in prices, product quality, technological lead, product variety or product design.

The innovation literature stresses that certain firm characteristics —such as the degree of product diversification, the degree of internationalization, the availability of financial resources and

[11] In Germany, the innovation surveys are conducted annually and they are designed as a panel (so called Mannheim Innovation Panel). Unfortunately, the overlap between the 2004 and 2005 survey only amounts to almost 40% due to a major refreshment and enlargement of the gross sample.

technological capabilities— are likewise crucial for explaining innovation (see, e.g., the references cited in Peters, 2008). More diversified firms possess economies of scope in innovation. As they have more opportunities to exploit new knowledge and complementarities among their diversified activities, they tend to be more innovative. We measure product diversification by the share of turnover of the firm's most important product in 2004 ($DIVERS$). Therefore, we expect a negative coefficient since more diversified firms exhibit lower values for this proxy.

The more a firm is exposed to international competition, the more likely the firm engages in R&D activities. The degree to which a firm is exposed to international competition is captured by a dummy variable taking the value of 1 if the firm sells its products to international markets ($EXPORT$).

The availability of financial resources is proxied by an index of creditworthiness ($RATING$). A lower creditworthiness implies less available and more costly external funding to finance R&D projects. Since the index ranges from 1 (best rating) to 6 (worst rating), we expect a negative coefficient for this proxy.

Innovative capabilities are determined by the skills of employees. We take into account the share of employees with a university degree ($HIGHSKILLED$), a dummy variable being 1 if the firm has not invested in training its employees ($NOTRAIN$) and the amount of training expenditure per employee ($TRAINEXP$) if the firm has invested in training. Since information on training expenditure is missing for 11.3% of the firms, we do not drop these observations but rather set the expenditure to zero and include a dummy variable indicating the missing value status ($MVTRAIN$).

We also include variables reflecting whether the firm is located in East Germany ($EAST$) and whether the firm is part of an enterprise group ($GROUP$). *A priori*, the effect of these variables is unclear. Finally, industry dummies are included in all regressions.

4.3 Results

4.3.1 Firms facing equal lottery probabilities

Table 2 reports the marginal effects of the probit estimates for the full sample, assuming that all firms face the same probabilities in the demand and supply lotteries. For each of the two endogenous variables, the first column reports the results for a parsimonious specification —including only $SIZE$ and industry dummies in addition to demand uncertainty, technical uncertainty and entry threat— whereas the second column employs the full set of control variables described in the previous section.

<Insert Table 2 about here>

Hypothesis 1, postulating that the probability of undertaking R&D does not decrease with an increase in θ, is strongly confirmed. Focusing on our preferred specification (*R&D* (2)), our results indicate that a 10% increase in θ increases the likelihood of undertaking R&D by 1.4%.

Hypothesis 2, postulating that the probability of undertaking R&D does not decrease with an increase in λ, is strongly confirmed. This result is robust across the two endogenous variables and holds when additional control variables are incorporated. We estimate that a change from a low to a high degree of divergence increases the likelihood of undertaking R&D by 23.3%.

Entry threat does not significantly influence the decision to undertake R&D. As *R&D* and $THREAT$ are measured over the same period, an endogeneity problem might arise as the decision to perform R&D could reduce the perceived entry threat. This explanation is supported by the fact that entry threat does significantly positively affect the decision to undertake process innovations in the next year.

Regarding the impact of the other control variables, most results are in line with expectations. Firm size exerts a significantly positive impact. Market structure has a non-linear effect on innovation. Firms in oligopoly markets have a higher likelihood of undertaking R&D or introducing new products compared to monopolists or firms with more competitors. Hence, our results support evidence in favor of the inverted U-relationship between competition and innovation as suggested by Aghion et al. (2005). Another striking and robust finding is that firms acting on markets where competition is mainly settled through prices are less likely to innovate. On the contrary, innovation activities are stimulated if competitive advantage can be achieved by technological leadership. Firms being exposed to international competition as well as more diversified firms have a higher likelihood of undertaking R&D and introducing new products. There is, however, no significant impact on process innovation. Finally, the results highlight the important role of innovative capabilities. Firms employing a higher share of high-skilled workers or firms investing in training are likely to be more innovative.

For the subsample, Table 3 presents in columns (2) and (4) the estimates using our preferred measure for technical uncertainty ($LAMBDA2$). For reasons of comparison, columns (1) and (3) show the subsample results employing $LAMBDA1$. In general, the results are very similar to the full sample. Hypothesis 2 is also strongly confirmed using $LAMBDA2$. Since we measure this variable in logarithm, a value of 0.017 implies that an increase in the deviation of expected and actual R&D expenditure by 10% increases the propensity to undertake R&D by 17%.[12]

<Insert Table 3 about here>

4.3.2 Firms facing different demand lottery probabilities

In this section we relax the assumption that the probability of facing a good demand state is the same for all firms. We approximate p_θ by looking at the firms' sales histories in the past two years. We define three groups of firms (see Table B.2 in Appendix B for exact definitions). Group 1 ($G1$) comprises all firms that experienced a decrease in sales in 2002-2003 as well as in 2003-2004. The idea is that these firms face an unfavorable demand lottery reflected by a low value of p_θ. These firms are much more likely to face a bad demand state than a good demand state. Group 2 ($G2$) consists of all firms that experienced one yearly decrease and one yearly increase in sales during the period 2002-2004. On the basis of this observation, we assume that these firms face equal probabilities of a good/bad demand state and therefore have a p_θ around $\frac{1}{2}$. All firms in group 3 ($G3$) experienced an increase in sales in 2002-2003 as well as in 2003-2004. The assumption is that these firms face a favorable demand lottery reflected by a high value of p_θ.

Assuming that firms in group 1 have a p_θ smaller than $\frac{1}{4}$, firms in group 2 have a p_θ around $\frac{1}{2}$ and firms in group 3 have a p_θ larger than $\frac{1}{2}$, we postulate from Proposition 3a and Proposition 2 respectively the following hypotheses.

Hypothesis 3: For firms in group 1, the probability of undertaking R&D does not increase with a more divergent demand lottery.

[12] To test whether multicollinearity between our main independent variables affect our results in Tables 2 and 3, we estimate specifications that include demand uncertainty, technical uncertainty or entry threat separately and specifications that combine demand uncertainty or technical uncertainty with entry threat. The significance as well as the magnitude of the estimated marginal effects are very robust in both the full sample and the subsample (results available upon request).

Hypothesis 4: For firms in group 2 and group 3, the probability of undertaking R&D does not decrease with a more divergent demand lottery.

Table 4 presents the results of distinguishing the effect of a more divergent demand lottery across groups of firms facing different demand lottery probabilities. Confirming hypothesis 3, we find that for firms in group 1 the effect of an increase in θ is significantly negative for $PROCESS$ and negative but not significant for $R\&D$ in the specifications including all control variables. Furthermore, the impact of $THETA$ is significantly different for firms in group 1 compared to firms in group 2 and group 3. Hypothesis 4 is strongly confirmed since the impact of $THETA$ is never significantly negative for firms in group 2 and group 3. Moreover, in all specifications, the effect of a more divergent demand lottery is significantly positive for firms in group 3. Furthermore, the impact is significantly larger for firms in group 3 than for firms in group 2 when process innovations are considered.[13]

<Insert Table 4 about here>

5 Conclusion

This article contributes to the theoretical as well as the empirical literature on R&D decisions under uncertainty.

From a theoretical point of view, we study a two-stage R&D project with an abandonment option. Two types of uncertainty influence the decision to start R&D. Demand uncertainty is modelled as a lottery between a proportional increase and decrease in demand. Technical uncertainty is modelled as a lottery between a decrease and increase in the cost to continue R&D. Both lotteries become more divergent when the difference between the outcomes of the lottery increases. We relate differences in uncertainty to differences in risk premia. This allows us to consider a broader set of demand and supply lotteries than only the subset of lotteries that preserve the mean, as previously studied in the literature. A potential entrant is endowed with a superior technology and threatens to drive the incumbent out of the market. The incumbent has a time lead over the entrant and can obtain the same superior technology by completing the R&D project before the entrant can enter the market. The presence of the entrant in our model provides the incumbent with additional benefits from investing in the superior technology, a strategic effect known as Arrow's replacement effect. In order to deduct testable hypotheses, we derive under which lottery probabilities more divergent demand and supply lotteries positively or negatively affect the decision to start R&D. Under mild assumptions, relating (in the absence of technical uncertainty) the cost of starting R&D to the cost of continuing R&D and the total cost of the R&D project to the profit gain, our most counterintuitive result is that an increase in the degree of divergence of an unfavorable demand lottery can positively affect the decision to start R&D. We obtain this result because of the abandonment option that the incumbent possesses.

From an empirical point of view, we test the hypotheses derived from the theoretical model using data from the fourth Community Innovation Survey (CIS IV) data in Germany. The uniqueness of our data lies in the availability of proxies for demand and technical uncertainty as well as perceived entry threat for about 4000 firms to explain actual and planned R&D investments. Our main results, strongly confirming our model predictions, are that for firms facing lotteries where the good state is more likely to prevail (i) a 10% increase in the degree of

[13] To test the robustness of the results in Table 4, we make a distinction between manufacturing and services. Hypotheses 2, 3 and 4 are confirmed in both samples (results available upon request).

divergence of the demand lottery increases the likelihood of undertaking R&D by 1.4% and (ii) a change from a low to a high degree of divergence of the supply lottery increases the likelihood of undertaking R&D by 23.3%. Using a subsample of firms for which we have a proxy for the degree of divergence of the supply lottery that more closely corresponds to our theoretical model, gives similar estimation results. For firms facing a demand lottery where the bad state is more likely to prevail, an increase in the degree of divergence of the demand lottery decreases the probability of undertaking R&D significantly.

References

[1] Aghion, P., N. Bloom, R. Blundell, R. Griffith and P. Howitt, 2005, Competition and innovation: An inverted U-relationship, Quarterly Journal of Economics, 120(2), 701-728.

[2] Arrow, K., 1962, The economic implications of learning by doing, Review of Economic Studies, 29(3), 155-173.

[3] Bloom, N., S. Bond and J. Van Reenen, 2007, Uncertainty and investment dynamics, Review of Economic Studies, 74(2), 391-415.

[4] Caballero, R. and R.S. Pindyck, 1996, Investment, uncertainty and industry evolution, International Economic Review, 37(3), 641-662.

[5] Darby, J., A.H. Hallett, J. Ireland and L. Piscitelli, 1999, The impact of exchange rate uncertainty on the level of investment, Economic Journal, 109(454), c55-c67.

[6] Dasgupta, P. and J. Stiglitz, 1980, Industrial structure and the nature of innovative activity, Economic Journal, 90(4), 266-293.

[7] Dixit, A. and R.S. Pindyck, 1994, Investment under uncertainty, Princeton, New Jersey: Princeton University Press.

[8] Dorfman, J.H. and D. Heien, 1989, The effects of uncertainty and adjustment costs on investment in the almond industry, The Review of Economics and Statistics, 71(2), 263-274.

[9] Eiser R. and M.I. Nadiri, 1968, Investment behavior and Neo-Classical theory, The Review of Economics and Statistics, 50(3), 369-382.

[10] Farzin, Y., K. Huisman and P. Kort, 1998, Optimal timing of technology adoption, Journal of Economic Dynamics and Control, 22(5), 779-799.

[11] Ferderer, J.P., 1993, The impact of uncertainty on aggregate investment spending: An empirical analysis, Journal of Money, Credit, and Banking, 25(1), 30-48.

[12] Ghosal, V. and P. Loungani, 1996, Product market competition and the impact of price uncertainty on investment: Some evidence from US manufacturing industries, Journal of Industrial Economics, 44(2), 217-228.

[13] Grenadier, S., 2000, Game choices: The intersection of real options and game theory, London: Risk Books.

[14] Grenadier, S. and A. Weiss, 1997, Investment in technological innovations: An option pricing approach, Journal of Financial Economics, 44(3), 397-416.

[15] Guiso, L. and G. Parigi, 1999, Investment and demand uncertainty, Quarterly Journal of Economics, 114(1), 185-227.

[16] Henley, A., A. Carruth and A. Dickerson, 2003, Industry-wide versus firm-specific uncertainty and investment: British company panel data evidence, Economics Letters, 78(1), 87-92.

[17] Huizinga, J., 1993, Inflation uncertainty, relative price uncertainty, and investment in U.S. manufacturing, Journal of Money, Credit, and Banking, 25(3), 521-557.

[18] Jorgenson, D.W., 1963, Capital theory and investment behavior, American Economic Review, 53(2), 247-259.

[19] Leahy, J. and T. Whited, 1996, The effect of uncertainty on investment: Some stylized facts, Journal of Money, Credit and Banking, 28(1), 64-83.

[20] Lukach, R., P.M. Kort and J. Plasmans, 2007, Optimal R&D investment strategies under the threat of new technology entry, International Journal of Industrial Organization, 25(1), 103-119.

[21] Peters, B., 2008, Innovation and firm performance. An empirical investigation for German firms, ZEW Economic Studies, 38, Heidelberg, New York.

[22] Pindyck, R.S., 1993, Investments of uncertain cost, Journal of Financial Economics, 34(1), 53-76.

[23] Pratt, J., 1964, Risk aversion in the small and in the large, Econometrica, 32(1-2), 122-136.

[24] Tyagi, R., 2006, New product introductions and failures under uncertainty, International Journal of Research in Marketing, 23(2), 199-213.

[25] Weeds, H., 2002, Strategic delay in a real options model of R&D competition, Review of Economic Studies, 69(3), 729-747.

Table 1
Descriptive Statistics - Full Sample

Variable	Unit	Mean	SD	Median	Skewness	Min	Max
Dependent variables							
R&D	[0/1]	0.484	0.500	0	–	0	1
PROCESS	[0/1]	0.460	0.498	0	–	0	1
Independent variables							
Demand uncertainty							
THETA	%	0.141	0.152	0.094	2.777	0	1.242
G1	[0/1]	0.199	–	–	–	–	–
G2	[0/1]	0.412	–	–	–	–	–
G3	[0/1]	0.389	–	–	–	–	–
Technical uncertainty							
LAMBDA1	[0/1]	0.193	0.395	0	–	0	1
LAMBDA2	Mill. Euro	2.425	12.769	0.098	10.101	0	207.318
Additional control variables							
THREAT: no	[0/1]	0.093	0.291	0	–	0	1
THREAT: low	[0/1]	0.447	0.497	0	–	0	1
THREAT: medium	[0/1]	0.310	0.463	0	–	0	1
THREAT: high	[0/1]	0.150	0.357	0	–	0	1
SIZE	# Empl.	587.179	5495.988	45	27.059	1	232700
NUMCOMP: 0	[0/1]	0.021	0.142	0	–	0	1
NUMCOMP: 1-5	[0/1]	0.578	0.494	1	–	0	1
NUMCOMP: 6-15	[0/1]	0.211	0.408	0	–	0	1
NUMCOMP: >15	[0/1]	0.191	0.393	0	–	0	1
COMP: PRICE	[0/1]	0.527	0.499	1	–	0	1
COMP: QUAL	[0/1]	0.418	0.493	0	–	0	1
COMP: LEAD	[0/1]	0.110	0.313	0	–	0	1
COMP: VARIETY	[0/1]	0.052	0.222	0	–	0	1
COMP: DESIGN	[0/1]	0.033	0.180	0	–	0	1
DIVERS	[0-100]	71.312	23.427	75	-0.502	0.5	100
EXPORT	[0/1]	0.527	0.499	1	–	0	1
RATING	[1-6]	2.15	0.817	2.19	0.571	1	6
HIGHSKILLED	[0-100]	20.314	24.013	10	1.633	0	100
TRAINEXP	Mill. Euro	0.001	0.001	0	7.852	0	0.025
NOTRAIN	[0/1]	0.120	0.325	0	–	0	1
MVTRAIN	[0/1]	0.113	0.317	0	–	0	1
EAST	[0/1]	0.322	0.467	0	–	0	1
GROUP	[0/1]	0.579	0.494	1	–	0	1

Values for $LAMBDA2$, $SIZE$ and $TRAINEXP$ are not log-transformed. For estimation purposes, however, a log-transformation of these variables is used to take into account the skewness of the distribution.

Table 2
Effect of demand and technical uncertainty on innovation - Full Sample

Dep. variables	R&D		PROCESS	
	(1)	(2)	(3)	(4)
Demand uncertainty				
THETA	0.233***	0.137***	0.124**	0.109**
	(0.048)	(0.047)	(0.055)	(0.055)
Technical uncertainty				
LAMBDA1	0.286***	0.233***	0.215***	0.189***
	(0.019)	(0.020)	(0.021)	(0.021)
Additional control variables				
THREAT: low	0.012	0.009	0.047	0.040
	(0.025)	(0.024)	(0.030)	(0.030)
THREAT: medium	-0.018	-0.004	0.062**	0.065**
	(0.026)	(0.025)	(0.031)	(0.031)
THREAT: high	-0.019	0.033	0.017	0.044
	(0.029)	(0.028)	(0.035)	(0.035)
SIZE	0.061***	0.051***	0.069***	0.061***
	(0.004)	(0.005)	(0.005)	(0.006)
NUMCOMP: 0	–	0.014	–	-0.019
		(0.049)		(0.063)
NUMCOMP: 1-5	–	0.037**	–	-0.008
		(0.018)		(0.022)
NUMCOMP: 6-15	–	-0.009	–	-0.029
		(0.021)		(0.025)
COMP: PRICE	–	-0.057***	–	-0.067***
		(0.018)		(0.019)
COMP: QUAL	–	-0.008	–	-0.006
		(0.017)		(0.018)
COMP: LEAD	–	0.134***	–	-0.000
		(0.026)		(0.027)
COMP: VARIETY	–	-0.044	–	-0.004
		(0.036)		(0.027)
COMP: DESIGN	–	-0.003	–	0.037
		(0.043)		(0.045)
DIVERS	–	-0.115***	–	-0.056
		(0.034)		(0.035)
EXPORT	–	0.142***	–	0.017
		(0.017)		(0.020)
RATING	–	0.003	–	-0.011
		(0.009)		(0.011)
HIGHSKILLED	–	0.003***	–	0.001
		(0.000)		(0.000)
TRAINEXP	–	0.048***	–	0.044***
		(0.007)		(0.008)
NOTRAIN	–	-0.439***	–	-0.419***
		(0.031)		(0.034)
MVTRAIN	–	-0.341***	–	-0.347***
		(0.032)		(0.037)
EAST	–	0.036**	–	0.014
		(0.015)		(0.018)
GROUP	–	0.011	–	0.031*
		(0.014)		(0.017)
$LogL$	-1952.6	-1773.7	-2066.4	-2010.6
R^2_{MF}	0.234	0.304	0.096	0.121
R^2_{MZ}	0.424	0.520	0.196	0.241
Count R^2	0.730	0.764	0.654	0.670
LM_{het} (p-value)	0.034	0.001	0.537	0.555
LM_{norm} (p-value)	0.398	0.641	0.552	0.579
# Obs.	3681	3681	3314	3314

Average marginal effects of the probit estimations are reported. Robust standard errors in parentheses.
***Significant at 1%; **Significant at 5%; *Significant at 10%. Industry dummies are included but not reported.
$LogL$: log likelihood value of the model with regressors. R^2_{MF} (likelihood ratio index): McFadden (1974) Pseudo R^2, comparing the likelihood of an intercept-only model to the likelihood of the model with regressors. R^2_{MZ}: McKelvey and Zavoina (1976) R^2, measuring the proportion of variance of the latent variable accounted for by the model. Count R^2: proportion of accurate predictions. LM_{het}: Davidson and MacKinnon (1984) test statistic for heteroskedasticity. LM_{norm}: Shapiro and Wilk (1965) test statistic for normality.

Table 3
Effect of demand and technical uncertainty on innovation - Subsample

Dep. variables	R&D		PROCESS	
	(1)	(2)	(3)	(4)
Demand uncertainty				
THETA	0.274**	0.224**	0.219*	0.154
	(0.106)	(0.101)	(0.130)	(0.122)
Technical uncertainty				
LAMBDA1	0.196***	–	0.161***	–
	(0.034)		(0.040)	
LAMBDA2	–	0.017***	–	0.023***
		(0.002)		(0.003)
Additional control variables				
THREAT: low	-0.022	0.001	0.002	0.018
	(0.039)	(0.039)	(0.054)	(0.054)
THREAT: medium	-0.016	-0.000	0.014	0.031
	(0.041)	(0.040)	(0.056)	(0.055)
THREAT: high	-0.024	0.004	-0.064	-0.042
	(0.047)	(0.045)	(0.065)	(0.065)
SIZE	0.057***	0.042***	0.058***	0.037***
	(0.009)	(0.009)	(0.011)	(0.011)
NUMCOMP: 0	-0.085	-0.037	0.083	0.118
	(0.083)	(0.073)	(0.127)	(0.110)
NUMCOMP: 1-5	0.074**	0.052*	-0.023	-0.040
	(0.030)	(0.030)	(0.042)	(0.040)
NUMCOMP: 6-15	0.022	0.024	0.010	0.009
	(0.035)	(0.034)	(0.048)	(0.047)
COMP: PRICE	-0.098***	-0.089***	-0.090**	-0.079**
	(0.033)	(0.026)	(0.034)	(0.034)
COMP: QUAL	-0.047*	-0.061***	-0.012	-0.024
	(0.032)	(0.023)	(0.034)	(0.033)
COMP: LEAD	0.131***	0.096***	-0.002	-0.026
	(0.037)	(0.033)	(0.048)	(0.045)
COMP: VARIETY	-0.018	-0.052	-0.017	-0.061
	(0.058)	(0.046)	(0.076)	(0.074)
COMP: DESIGN	-0.064	-0.097*	0.134	0.092
	(0.061)	(0.055)	(0.089)	(0.085)
DIVERS	-0.051	-0.051	-0.023	-0.022
	(0.049)	(0.046)	(0.066)	(0.065)
EXPORT	0.164***	0.162***	-0.016	-0.031
	(0.032)	(0.032)	(0.041)	(0.040)
RATING	-0.002	-0.001	-0.015	-0.016
	(0.015)	(0.015)	(0.019)	(0.018)
HIGHSKILLED	0.002***	0.002***	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
TRAINEXP	0.067***	0.049***	0.048***	0.029*
	(0.012)	(0.012)	(0.015)	(0.015)
NOTRAIN	-0.544***	-0.476***	-0.422***	-0.302***
	(0.039)	(0.067)	(0.063)	(0.104)
MVTRAIN	-0.458***	-0.388***	-0.417***	-0.339***
	(0.042)	(0.070)	(0.053)	(0.085)
EAST	0.006	0.002	-0.011	-0.027
	(0.025)	(0.024)	(0.036)	(0.034)
GROUP	0.014	-0.001	0.071**	0.056*
	(0.025)	(0.023)	(0.034)	(0.033)
LogL	-401.2	-376.8	-557.5	-531.9
R^2_{MF}	0.423	0.458	0.132	0.172
R^2_{MZ}	0.658	0.685	0.259	0.326
Count R^2	0.807	0.825	0.666	0.695
LM_{het} (p-value)	0.733	0.472	0.961	0.929
LM_{norm} (p-value)	0.288	0.679	0.343	0.417
# Obs.	1009	1009	929	929

Average marginal effects of the probit estimations are reported. Robust standard errors in parentheses.
***Significant at 1%; **Significant at 5%; *Significant at 10%. Industry dummies are included but not reported.
For notes on goodness-of-fit and specification tests: see Table 2.

Table 4
Effect of demand uncertainty on innovation across groups of firms facing a different p_θ - Full Sample

Dep. variables	**R&D**		**PROCESS**	
	(1)	(2)	(3)	(4)
Demand uncertainty				
THETA*G1	-0.185	-0.136	-0.402***	-0.312**
	(0.125)	(0.119)	(0.143)	(0.144)
THETA*G2	0.214***	0.149**	0.009	0.021
	(0.062)	(0.059)	(0.072)	(0.072)
THETA*G3	0.293***	0.155***	0.281***	0.233***
	(0.060)	(0.058)	(0.069)	(0.068)
$\alpha_{\theta*G1} >= \alpha_{\theta*G2}$ (p-value)	0.001	0.009	0.002	0.011
$\alpha_{\theta*G1} >= \alpha_{\theta*G3}$ (p-value)	0.000	0.008	0.000	0.000
$\alpha_{\theta*G2} >= \alpha_{\theta*G3}$ (p-value)	0.139	0.463	0.001	0.005
Technical uncertainty				
LAMBDA1	0.283***	0.231***	0.211***	0.188***
	(0.019)	(0.020)	(0.021)	(0.021)
Entry threat				
THREAT: low	0.011	0.009	0.046	0.041
	(0.025)	(0.024)	(0.030)	(0.030)
THREAT: medium	-0.020	-0.005	0.058*	0.062**
	(0.026)	(0.025)	(0.031)	(0.031)
THREAT: high	-0.014	0.035	0.022	0.047
	(0.029)	(0.028)	(0.035)	(0.035)
$LogL$	-1944.1	-1770.1	-2052.4	-2001.7
R^2_{MF}	0.237	0.306	0.102	0.125
R^2_{MZ}	0.428	0.522	0.207	0.248
Count R^2	0.734	0.767	0.653	0.674
LM_{het} (p-value)	0.008	0.000	0.511	0.678
LM_{norm} (p-value)	0.224	0.702	0.330	0.962
# Obs.	3681	3681	3314	3314

Average marginal effects of the probit estimations are reported. Robust standard errors in parentheses. ***Significant at 1%; **Significant at 5%; *Significant at 10%. In columns (1) and (3) $SIZE$ and industry dummies are included as control variables but not reported. In columns (2) and (4) the full set of control variables including industry dummies is used but not reported (see Table 2). For notes on goodness-of-fit and specification tests: see Table 2.

Appendix A: Proofs

Proof of Propositions 1 & 2: Consider the partial derivatives of the five arguments of Φ with respect to θ: $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial \theta} = (2p_\theta - 1)\,\Delta\pi$, $\frac{\partial \Delta_{NPV}^{(1,1,1,0)}}{\partial \theta} = (p_\theta - p_\lambda + p_\theta p_\lambda)\,\Delta\pi$, $\frac{\partial \Delta_{NPV}^{(1,1,0,0)}}{\partial \theta} = p_\theta \Delta\pi$, $\frac{\partial \Delta_{NPV}^{(1,0,1,0)}}{\partial \theta} = p_\lambda\,(2p_\theta - 1)\,\Delta\pi$, $\frac{\partial \Delta_{NPV}^{(1,0,0,0)}}{\partial \theta} = p_\theta p_\lambda \Delta\pi$. All five partial derivatives are either negative or equal to zero when $p_\theta = 0$ for all $p_\lambda \in [0,1]$. This is a sufficient condition to obtain Proposition 1. All five partial derivatives are either positive or equal to zero when $p_\theta \in [\frac{1}{2}, 1]$ for all $p_\lambda \in [0,1]$. This is a sufficient condition to obtain Proposition 2. ■

Proofs of Propositions 3a, 3b & 4: Before we prove Propositions 3a, 3b & 4 consequently, we introduce Lemma 1 and Lemma 1'. Lemma 1 identifies Φ for different ranges of the parameters θ and λ. Lemma 1 holds over the complete parameter space of (c, p_θ, p_λ).

Lemma 1:
(1) $\Phi = \Delta_{NPV}^{(1,1,1,1)}$ *for all* $\theta \in [0, \frac{1}{2}]$ *and for all* $\lambda \in [0, \left(\frac{1}{2} - \theta\right)\Delta\pi]$.
(2) $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ *for all* $\theta \in [0, \frac{1}{2}]$ *and for all* $\lambda \in [\left(\frac{1}{2} - \theta\right)\Delta\pi, \lambda_{\max}]$.
(3) $\Phi = \Delta_{NPV}^{(1,1,0,0)}$ *for all* $\theta \in [\frac{1}{2}, 1]$ *and for all* $\lambda \in [0, \left(\theta - \frac{1}{2}\right)\Delta\pi]$.
(4) $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ *for all* $\theta \in [\frac{1}{2}, 1]$ *and for all* $\lambda \in [\left(\theta - \frac{1}{2}\right)\Delta\pi, \lambda_{\max}]$.

Proof of Lemma 1: From Assumptions 1-2, it follows that $\lambda_{\max} = I = \frac{\Delta\pi}{2}$. Then, $\Delta_{NPV}^{GB} \geq 0$ when $\left(\frac{1}{2} + \theta\right)\Delta\pi \geq \lambda$. Therefore, $\Delta_{NPV}^{GB} \geq 0$ for all $\theta \in [0,1]$ and $\lambda \in [0, \lambda_{\max}]$. As a result, also $\Delta_{NPV}^{GG} \geq 0$ for all $\theta \in [0,1]$ and $\lambda \in [0, \lambda_{\max}]$ (cfr. Section 2.3). Then, $\Delta_{NPV}^{BG} \geq 0$ when $\left(\theta - \frac{1}{2}\right)\Delta\pi \leq \lambda$. Therefore, $\Delta_{NPV}^{BG} \geq 0$ for all $\theta \in [0, \frac{1}{2}]$ and for all $\lambda \in [0, \lambda_{\max}]$, $\Delta_{NPV}^{BG} \leq 0$ for all $\theta \in [\frac{1}{2}, 1]$ and for all $\lambda \in [0, \left(\theta - \frac{1}{2}\right)\Delta\pi]$ and $\Delta_{NPV}^{BG} \geq 0$ for all $\theta \in [\frac{1}{2}, 1]$ and for all $\lambda \in [\left(\theta - \frac{1}{2}\right)\Delta\pi], \lambda_{\max}]$. Then, $\Delta_{NPV}^{BB} \geq 0$ when $\left(\frac{1}{2} - \theta\right)\Delta\pi \geq \lambda$. Therefore, $\Delta_{NPV}^{BB} \geq 0$ for all $\theta \in [0, \frac{1}{2}]$ and for all $\lambda \in [0, \left(\frac{1}{2} - \theta\right)\Delta\pi]$, $\Delta_{NPV}^{BB} \leq 0$ for all $\theta \in [0, \frac{1}{2}]$ and for all $\lambda \in [\left(\frac{1}{2} - \theta\right)\Delta\pi, \lambda_{\max}]$ and $\Delta_{NPV}^{BB} \leq 0$ for all $\theta \in [\frac{1}{2}, 1]$ and for all $\lambda \in [0, \lambda_{\max}]$. Lemma 1 follows from noting that $\Phi = \Delta_{NPV}^{(1,1,1,1)}$ when $\Delta_{NPV}^{GB} \geq 0$, $\Delta_{NPV}^{BG} \geq 0$ and $\Delta_{NPV}^{BB} \geq 0$, that $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ when $\Delta_{NPV}^{GB} \geq 0$, $\Delta_{NPV}^{BG} \geq 0$ and $\Delta_{NPV}^{BB} \leq 0$ and that $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ when $\Delta_{NPV}^{GB} \geq 0$, $\Delta_{NPV}^{BG} \leq 0$ and $\Delta_{NPV}^{BB} \leq 0$. ■

We use Lemma 1, where λ is expressed as a function of θ, in the determination of x and y. For the determination of v and w, it is useful to rewrite Lemma 1 as Lemma 1' where we express θ as a function of λ. Again, Lemma 1' holds over the complete parameter space of (c, p_θ, p_λ).

Lemma 1':
(1) $\Phi = \Delta_{NPV}^{(1,1,1,1)}$ *for all* $\lambda \in [0, \lambda_{\max}]$ *and for all* $\theta \in [0, \frac{\lambda_{\max} - \lambda}{\Delta\pi}]$.
(2) $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ *for all* $\lambda \in [0, \lambda_{\max}]$ *and for all* $\theta \in [\frac{\lambda_{\max} - \lambda}{\Delta\pi}, \frac{\lambda_{\max} + \lambda}{\Delta\pi}]$.
(3) $\Phi = \Delta_{NPV}^{(1,1,0,0)}$ *for all* $\lambda \in [0, \lambda_{\max}]$ *and for all* $\theta \in [\frac{\lambda_{\max} + \lambda}{\Delta\pi}, 1]$.

Proof of Proposition 3a: We prove that the smallest p_θ for which a positive effect of an increase in θ on the decision to start R&D is found, equals $\frac{1}{4}$ by showing that $\Phi(\theta) = 0$ for $\Phi = \Delta_{NPV}^{(1,1,0,0)}$, $\theta = 1$, $p_\theta = \frac{1}{4}$, $\lambda = \lambda_{\max}$ and $p_\lambda = 1$.

First, consider the partial derivatives of $\Delta_{NPV}^{(1,1,1,1)}$, $\Delta_{NPV}^{(1,1,1,0)}$ and $\Delta_{NPV}^{(1,1,0,0)}$ with respect to θ when $p_\theta \in [0, \frac{1}{2}]$. Note that $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial \theta} = (2p_\theta - 1)\,\Delta\pi \leq 0$, $\frac{\partial \Delta_{NPV}^{(1,1,1,0)}}{\partial \theta} = (p_\theta - p_\lambda + p_\theta p_\lambda)\,\Delta\pi \geq 0$ if and only if $p_\lambda \leq \frac{p_\theta}{1-p_\theta}$ and $\frac{\partial \Delta_{NPV}^{(1,1,0,0)}}{\partial \theta} = p_\theta \Delta\pi \geq 0$. A positive effect due to an increase in θ can only be found when $\frac{\partial \Phi(\theta)}{\partial \theta} \geq 0$ at some subdomain of θ.

Second, from the fact that $\Delta_{NPV}^{GG} \geq \Delta_{NPV}^{s} \geq \Delta_{NPV}^{BB}$ for $s \in \{GB, BG\}$ (cfr. Section 2.3), it follows that $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial p_\theta} = p_\lambda \left(\Delta_{NPV}^{GG} - \Delta_{NPV}^{BG}\right) + (1-p_\lambda)\left(\Delta_{NPV}^{GB} - \Delta_{NPV}^{BB}\right) \geq 0$, $\frac{\partial \Delta_{NPV}^{(1,1,1,0)}}{\partial p_\theta} = p_\lambda \left(\Delta_{NPV}^{GG} - \Delta_{NPV}^{BG}\right) + (1-p_\lambda)\Delta_{NPV}^{GB} \geq 0$ and $\frac{\partial \Delta_{NPV}^{(1,1,0,0)}}{\partial p_\theta} = p_\lambda \Delta_{NPV}^{GG} + (1-p_\lambda)\Delta_{NPV}^{GB} \geq 0$. From these observations and the definition of x, it follows that when $p_\theta = x$, $\Phi(\theta) = 0$ when $\theta = 1$.

Third, from Lemma 1, $\Phi(1) = 0$ holds for $\Phi = \Delta_{NPV}^{(1,1,0,0)}$. Solving $\Delta_{NPV}^{(1,1,0,0)}(1) = 0$ yields $p_\theta = \frac{\frac{1}{2}\Delta\pi}{\frac{3}{2}\Delta\pi + (2p_\lambda - 1)\lambda}$. We find x by solving $\min_{\lambda, p_\lambda} p_\theta$. For $\lambda = \lambda_{\max}$ and $p_\lambda = 1$, $x = \frac{1}{4}$. ■

Proof of Proposition 3b: We prove that the smallest p_λ for which a positive effect of an increase in λ on the decision to start R&D is found, approximately equals 0.28 by showing that $\Phi(\lambda) = 0$ for $\Phi = \Delta_{NPV}^{(1,1,1,0)}$, $\lambda = \lambda_{\max}$, $p_\lambda = 0.28$, $\theta = 1$, and $p_\theta = \frac{p_\lambda}{1-p_\lambda}$.

First, consider the partial derivatives of $\Delta_{NPV}^{(1,1,1,1)}$, $\Delta_{NPV}^{(1,1,1,0)}$ and $\Delta_{NPV}^{(1,1,0,0)}$ with respect to λ when $p_\lambda \in [0, \frac{1}{2}]$. Note that $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial \lambda} = (2p_\lambda - 1)\Delta\pi \leq 0$, $\frac{\partial \Delta_{NPV}^{(1,1,1,0)}}{\partial \lambda} = (-p_\theta + p_\lambda + p_\theta p_\lambda)\Delta\pi \geq 0$ if and only if $p_\theta \leq \frac{p_\lambda}{1-p_\lambda}$ and $\frac{\partial \Delta_{NPV}^{(1,1,0,0)}}{\partial \lambda} = p_\theta(2p_\lambda - 1)\Delta\pi \leq 0$ for all $p_\theta \in [0,1]$. A positive effect due to an increase in λ can only be found when $\frac{\partial \Phi(\lambda)}{\partial \lambda} \geq 0$ at some subdomain of λ.

Second, $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial p_\lambda} = p_\theta \left(\Delta_{NPV}^{GG} - \Delta_{NPV}^{GB}\right) + (1-p_\theta)\left(\Delta_{NPV}^{BG} - \Delta_{NPV}^{BB}\right) \geq 0$ and $\frac{\partial \Delta_{NPV}^{(1,1,0,0)}}{\partial p_\lambda} = p_\theta \left(\Delta_{NPV}^{GG} - \Delta_{NPV}^{GB}\right) \geq 0$. Also, $\frac{\partial \Delta_{NPV}^{(1,1,1,0)}}{\partial p_\lambda} = p_\theta \left(\Delta_{NPV}^{GG} - \Delta_{NPV}^{GB}\right) + (1-p_\theta)\Delta_{NPV}^{BG} \geq 0$ if and only if $\Delta_{NPV}^{BG} \geq 0$. This is the case when $\Phi = \Delta_{NPV}^{(1,1,1,0)}$. From these observations and the definition of v, it follows that when $p_\lambda = v$, $\Phi(\lambda) = 0$ when $\lambda = \lambda_{\max}$.

Third, from Lemma 1', $\Phi(\lambda_{\max}) = 0$ holds for $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ when $\theta \in [0,1]$. Solving $\Delta_{NPV}^{(1,1,1,0)}(\lambda_{\max}) = 0$ yields $p_\lambda = \frac{\frac{1}{2} - p_\theta\theta}{1 - \theta + p_\theta\theta}$. We find v by solving $\min_{\theta, p_\theta} p_\lambda$ subject to $p_\theta \leq \frac{p_\lambda}{1-p_\lambda}$. For $\theta = 1$ and $p_\theta = \frac{p_\lambda}{1-p_\lambda}$, $v = 0.280776 \approx 0.28$. ■

Proof of Proposition 4: We first prove that the lowest p_θ for which no negative effect of an increase in θ on the decision to start R&D can be found, equals $\frac{1}{2}$ by showing that $\Phi(\theta) = 0$ for $\Phi = \Delta_{NPV}^{(1,1,1,0)}$, $\theta = \frac{1}{2}$, $p_\theta = \frac{1}{2}$, $\lambda = 0$ and $p_\lambda = 1$.

First, a negative effect due to an increase in θ can only be found when $\frac{\partial \Phi(\theta)}{\partial \theta} \leq 0$ at some subdomain of θ. Hence, Φ has to be equal to $\Delta_{NPV}^{(1,1,1,1)}$or $\Delta_{NPV}^{(1,1,1,0)}$ when $p_\lambda \geq \frac{p_\theta}{1-p_\theta}$ at some subdomain of θ.

Second, from the observation that $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial p_\theta} \geq 0$,$\frac{\partial \Delta_{NPV}^{(1,1,1,0)}}{\partial p_\theta} \geq 0$ and $\frac{\partial \Delta_{NPV}^{(1,1,0,0)}}{\partial p_\theta} \geq 0$ (cfr. proof of Proposition 3a) and from the definition of y, two possibilities arise. Either, $\Phi(\theta) = 0$ for $\theta = \frac{1}{2}$ and $p_\theta = y$, when (i) for $\theta \in [0, \frac{1}{2}]$, $\Phi = \Delta_{NPV}^{(1,1,1,1)}$ or $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ and $p_\lambda \geq \frac{p_\theta}{1-p_\theta}$ and for $\theta \in [\frac{1}{2}, 1]$, $\Phi = \Delta_{NPV}^{(1,1,0,0)}$ or when (ii) for $\theta \in [0, \frac{1}{2}]$, $\Phi = \Delta_{NPV}^{(1,1,1,1)}$ and for $\theta \in [\frac{1}{2}, 1]$, $\Phi = \Delta_{NPV}^{(1,1,0,0)}$or $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ and $p_\lambda \leq \frac{p_\theta}{1-p_\theta}$. Or $\Phi(\theta) = 0$ for $\theta = 1$ and $p_\theta = y$ when, for $\theta \in [0, \frac{1}{2}]$, $\Phi = \Delta_{NPV}^{(1,1,1,1)}$ or $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ and $p_\lambda \geq \frac{p_\theta}{1-p_\theta}$ and for $\theta \in [\frac{1}{2}, 1]$, $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ and $p_\lambda \geq \frac{p_\theta}{1-p_\theta}$.

Third, from Lemma 1, $\Phi(\frac{1}{2}) = 0$ holds for $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ for all $\lambda \in [0, \lambda_{\max}]$ when $p_\lambda \geq \frac{p_\theta}{1-p_\theta}$. Solving $\Delta_{NPV}^{(1,1,1,0)}(\frac{1}{2}) = 0$ yields $p_\theta = \frac{\frac{1}{2}\Delta\pi - p_\lambda \lambda}{\Delta\pi - (1-p_\lambda)\lambda}$. We find y by solving $\max_{\lambda, p_\lambda} p_\theta$ subject to $p_\lambda \geq \frac{p_\theta}{1-p_\theta}$. For $\lambda = 0$ and $p_\lambda = 1$, $y = \frac{1}{2}$. Since y cannot exceed $\frac{1}{2}$ (cfr. Proposition 2), the result follows.

We now prove that the lowest p_λ for which no negative effect of an increase in λ on the decision to start R&D can be found, equals $\frac{1}{2}$ by showing that $\Phi(\lambda) = 0$ for $\Phi = \Delta_{NPV}^{(1,1,1,0)}$, $\lambda = \lambda_{\max}$, $p_\lambda = \frac{1}{2}$, $\theta = 0$ and $p_\theta = 1$.

First, a negative effect due to an increase in λ can only be found when $\frac{\partial \Phi(\lambda)}{\partial \lambda} \leq 0$ at some subdomain of λ. Hence, in order to find a negative effect, Φ has to be equal to $\Delta_{NPV}^{(1,1,1,1)}$, $\Delta_{NPV}^{(1,1,0,0)}$ or $\Delta_{NPV}^{(1,1,1,0)}$ when $p_\theta \geq \frac{p_\lambda}{1-p_\lambda}$ at some subdomain of λ.

Second, from the observation that $\frac{\partial \Delta_{NPV}^{(1,1,1,1)}}{\partial p_\lambda} \geq 0, \frac{\partial \Delta_{NPV}^{(1,1,1,0)}}{\partial p_\lambda} \geq 0$ and $\frac{\partial \Delta_{NPV}^{(1,1,0,0)}}{\partial p_\lambda} \geq 0$ (cfr. proof of Proposition 3b) and from the definition of w, it follows that when $p_\lambda = w$, $\Phi(\lambda) = 0$ when $\lambda = \lambda_{\max}$.

Third, from Lemma 1', $\Phi(\lambda_{\max}) = 0$ holds for $\Phi = \Delta_{NPV}^{(1,1,1,0)}$ for all $\theta \in [0, 1]$ when $p_\theta \geq \frac{p_\lambda}{1-p_\lambda}$. Solving $\Delta_{NPV}^{(1,1,1,0)}(\lambda_{\max}) = 0$ yields $p_\lambda = \frac{\frac{1}{2} - p_\theta \theta}{1 - \theta + p_\theta \theta}$. We find w by solving $\max_{\theta, p_\theta} p_\lambda$ subject to $p_\theta \geq \frac{p_\lambda}{1-p_\lambda}$. For $\theta = 0$ and $p_\theta = 1$, $w = \frac{1}{2}$. Since w cannot exceed $\frac{1}{2}$, the result follows. ■

Appendix B: Statistical annex

Table B.1
Distribution of the Total Sample, Full Sample and Subsample

Distribution by:	Total Sample	Full Sample	Subsample
Industry			
Food/tobacco	3.16	3.10	2.68
Textiles	2.97	2.77	2.48
Paper/wood/print	6.7	6.82	6.05
Chemicals	4.1	4.13	4.96
Plastic/rubber	3.62	3.83	3.67
Glass/ceramics	2.14	2.25	2.97
Metal	8.35	8.53	10.31
Machinery	5.99	6.38	7.04
Electrical engineering	4.88	5.22	6.54
Medical, precision and optical instruments	4.92	5.51	6.64
Vehicles	2.66	2.53	2.58
Furniture	2.62	2.69	2.28
Wholesale	4.38	4.18	4.06
Retail	2.35	2.06	2.18
Transport/storage/post	8.46	8.10	5.55
Banks/insurances	5.05	4.48	3.87
Computer/telecommunication	4.59	4.75	4.66
Technical services	8.79	8.88	9.81
Consultancies	3.77	3.59	2.97
Other business related services	7.06	6.93	5.95
Real estate/renting	2.07	1.98	2.28
Media	1.38	1.28	0.50
Size (Number of employees)			
0-4	4.65	3.75	3.47
5-9	14.24	13.34	13.78
10-19	16.52	15.62	13.88
20-49	18.68	19.02	21.01
50-99	13.13	13.61	13.88
100-199	14.07	14.72	14.47
200-499	7.96	8.69	8.42
500-999	4.98	5.35	5.35
1000+	5.78	5.90	5.75
Region			
West Germany	66.86	67.81	64.42
East Germany	33.14	32.19	35.58
Innovation activities			
Non-innovators	36.12	33.14	31.71
Innovators[a]	63.88	66.86	68.29
# Obs.	4776	3681	1009

[a] Innovators are defined as firms having introduced product or process innovations in the period 2002-2004.

Table B.2
Variable definitions

Variable	Type	Definition
Dependent variables		
R&D	0/1	1 if the firm undertook R&D activities in the period 2002-2004.
PROCESS	0/1	1 if the firm planned to undertake process innovations in 2005.
Independent variables		
Demand uncertainty		
THETA	c	Average of the absolute percentage change in sales over the last two years (2002/2003 and 2003/2004).
G1	0/1	1 if the firm experienced two negative demand shocks in the past two years, i.e. a decrease in sales in 2002/2003 as well as in 2003/2004.
G2	0/1	1 if the firm experienced a positive and a negative demand shock in the past two years, i.e. one decrease and one increase in sales within the last two years.
G3	0/1	1 if the firm experienced two positive demand shocks in the past two years, i.e. a positive growth in sales in 2002/2003 as well as in 2003/2004.
Technical uncertainty		
LAMBDA1	0/1	1 if high innovation costs were of high- to medium-size importance and led to an extension of innovation projects in the period 2002-2004.
LAMBDA2	c	Absolute deviation between in year 2003 expected R&D expenditure for 2004 and realized R&D expenditure in 2004, in log.
Additional control variables		
THREAT	0/1	3 dummy variables indicating whether the firm perceived a high/medium/low threat of its own market position due to the potential entry of new competitors (reference group: firms with no entry threat).
SIZE	c	Number of employees in 2003, in log.
NUMCOMP	0/1	3 dummy variables indicating the number of competitors: 0, 1-5 or 6-15 (reference group: more than 15 competitors).
COMP	0/1	5 dummy variables indicating the most important factors of competition: price, quality, technological lead, product variety or product design (multiple factors allowed).
DIVERS	0-100	Share of turnover of most important product in 2004.
EXPORT	0/1	1 if the firm sold its products to international markets in the period 2002-2004.
RATING	c	Credit rating index of the firm in year 2003, ranging between 1 (highest) and 6(worst creditworthiness).
HIGHSKILLED	0-100	Share of employees with a university or college degree in 2003.
NOTRAIN	0/1	1 if the firm did not invest in training in 2004.
TRAINEXP	c	Training expenditure per employee (in log.) if NOTRAIN=0, otherwise 0.
MVTRAIN	0/1	1 if the information on training expenditure is missing in the data.
EAST	0/1	1 if the firm is located in East Germany.
GROUP	0/1	1 if the firm belongs to a group.

$0/1$ indicates a binary variable, c a continuous variable and $0-100$ describes a continuous variable with range of 0 to 100.