TI 2009-044/3
Tinbergen Institute Discussion Paper

Starting an R&D Project under Uncertainty

Sabien Dobbelaere[1]
Roland Iwan Luttens[2]
Bettina Peters[3]

[1] *Faculty of Economics & Business Administration, VU University Amsterdam, Tinbergen Institute, IZA Bonn, Ghent University;*
[2] *SHERPPA, Ghent University, and CORE, Université Catholique de Louvain;*
[3] *Centre for European Economic Research (ZEW).*

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

An analysis of the determinants of starting an R&D project*

(previously circulated as "**Starting an R&D project under uncertainty**")

Sabien Dobbelaere[†] Roland Iwan Luttens[‡] Bettina Peters[§]

September, 2010

Abstract

We study different determinants of real-life R&D decisions within a net present value framework. Besides entry threat, Bertrand competition and multi-stage R&D with an abandonment option, our model includes demand uncertainty and technical uncertainty, both modelled as a lottery. Each 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. Using CIS IV data for about 2400 German firms, we find 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 1.7 percentage points 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 14.1 percentage points. For firms facing a demand lottery where the bad state is most likely to prevail, a 10% increase in the degree of divergence of the demand lottery decreases the likelihood of undertaking R&D by 4.9 percentage points. Having the option to abandon R&D projects significantly increases the likelihood of undertaking R&D.

JEL classification : D21, D81, L12, O31.
Keywords : Multi-stage R&D, demand uncertainty, technical uncertainty, entry threat, abandonment option.

*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, Charles Bos, Bart Hobijn, Boris Lokshin, Jacques Mairesse, Pierre Mohnen, 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), the European Association for Research in Industrial Economics (EARIE) Conference (Toulouse, 2008) and seminars at the University of Strasbourg (2009), VU University Amsterdam (2009) and UNU-MERIT (2010) for helpful comments and suggestions. All remaining errors are ours.

[†]VU University Amsterdam, Tinbergen Institute, IZA Bonn. 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), University of Zurich.

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 they are highly uncertain. The factors that influence the firm's decision to undertake R&D activities have attracted the attention of policy makers, business leaders and researchers for a long time. From the theoretical and empirical literature, we can classify these factors into (i) firm characteristics, (ii) industry characteristics and (iii) project characteristics. The first category includes firm size (Cohen and Levin, 1989), corporate variables such as product diversification (Gabrowski, 1968), absorptive capacity (Cohen and Levinthal, 1989), appropriability (Cassiman and Veugelers, 2002) and technological advantage (Segerstrom, 2007), and financial variables such as financial constraints (Fazzari and Athey, 1987) and internal cashflow (Jorgenson, 1963; Eiser and Nadiri, 1968). The second category includes market structure which is determined by market power (Levin et al., 1985), competition (Grenadier, 2000; Huisman et al., 2005) and entry pressure (Etro, 2006; Acemoglu, 2008; Aghion et al., 2009), and general industry conditions (Acs and Audretsch, 1987; Dorfman, 1987). The third category includes different types of options such as a timing option (Dixit and Pindynck, 1994; Trigeorgis, 1996) and an abandonment option (Myers and Majd, 1990; Berger et al., 1996), and different types of uncertainty such as input cost uncertainty (Pindynck, 1993), technical uncertainty (Pindyck, 1993) and market uncertainty (Tyagi, 2006).

Technical uncertainty and market uncertainty are of particular interest for our analysis. Technical uncertainty implies that, although the input prices are known, the firm does not know at the beginning the amount of 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. 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 (Dorfman and Heien, 1989; Leahy and Whited, 1996; Bell and Campa, 1997; Guiso and Parigi, 1999; Henley et al., 2003; Bulan, 2005; Bloom et al., 2007), there is little empirical research that investigates the role of *different* types of uncertainty on R&D decisions. In particular, the existing empirical evidence only focuses on market uncertainty. In general, market uncertainty reduces R&D investments. However, Czarnitzki and Toole (2009, 2010) show that the negative effect is mitigated when firms receive R&D subsidies or patent their innovations.

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, Bertrand competition and multi-stage R&D with an abandonment option, our model includes demand (market) uncertainty as well as supply (technical) uncertainty. We deduct testable hypotheses on the basis of which we empirically analyze the non-traditional factors driving the decision to start an R&D project. The uniqueness of our data lies in the availability of proxies for demand and supply uncertainty, the abandonment option as well as perceived entry threat.

We model an R&D project as a multi-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. 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 2400 firms to explain the decision to start R&D. 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 1.7 percentage points and (ii) a change from a low to a high degree of divergence of the supply lottery (captured by a shift in the value of a dummy variable) increases the likelihood of undertaking R&D by 14.1 percentage points. For firms facing a demand lottery where the bad state is most likely to prevail, a 10% increase in the degree of divergence of the demand lottery decreases the likelihood of undertaking R&D by 4.9 percentage points. For a subset of firms that are more likely to face a bad demand state than a good demand state, having the option to abandon R&D projects significantly increases the likelihood of undertaking R&D.

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 (Dasgupta and Stiglitz, 1980; Weeds, 2002) to capture the uncertainty resolving nature of multi-stage R&D. Lukach et al. (2007) only consider supply lotteries and use the variance-based concept of a mean preserving spread to distinguish lotteries in terms of uncertainty. Our analysis studies a broader class of both demand and supply lotteries. An increase in the degree of divergence still symmetrically affects the good/bad state but the probability that the good/bad state occurs can take any value between 0 and 1. As a result, our set of lotteries cannot be ordered completely in terms of uncertainty. Instead, we order lotteries in terms of lottery premia. A lottery premium equals the amount of money that the incumbent is willing to pay (or has to receive) to undergo the lottery. The use of a lottery premium is particularly suitable in a NPV framework since the lottery 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 is motivated by the observation that the determinants of real-life R&D decisions greatly vary across studies as soon as the analysis is performed using more aggregated data (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 the determinants of R&D decisions. The comparative statics of Section 3 allow us to derive testable hypotheses on the relation between a change in the degree of divergence of demand and supply lotteries and the decision to start R&D. Section 4 presents the empirical analysis. Section 5 concludes.

2 A theoretical analysis of the determinants of R&D decisions

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 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 multi-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.

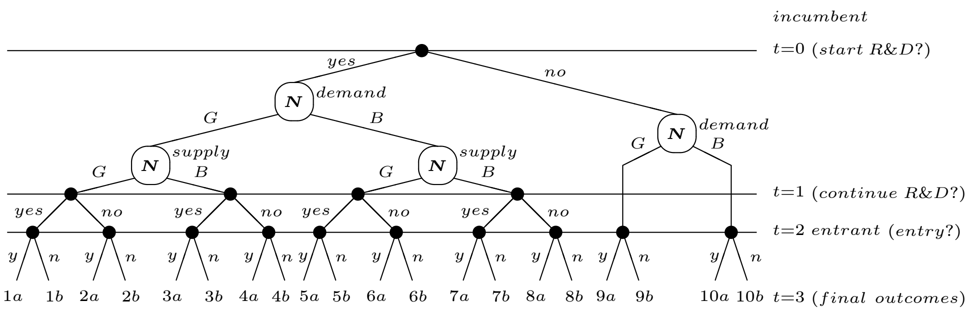

Figure 1 **Game tree.** At t=0, the incumbent decides whether to start R&D. Before t=1, nature (**N**) reveals the good/bad state (G/B) on the demand and supply side (the true state on the supply side is of no influence when the incumbent decides not to start R&D). At t=1, the incumbent decides whether 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 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 ($\pi(0) < \omega$)).

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

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:

[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).

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

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 $\Delta_{NPV}^{s} \geq 0$.*

Let $\boldsymbol{\psi} = (\psi_{GG}, \psi_{GB}, \psi_{BG}, \psi_{BB})$, where $\psi_s = 1$ when $\Delta_{NPV}^{s} \geq 0$ and $\psi_s = 0$ when $\Delta_{NPV}^{s} < 0$ 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 $\Delta_{NPV}^{GG} \geq \Delta_{NPV}^{s} \geq \Delta_{NPV}^{BB}$ 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 $\Delta_{NPV}^{\boldsymbol{\psi}}$, 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:

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

From this, we calculate:

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

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

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.

3.1 Relating divergence to lottery premia

Let us first explain how a change in the degree of divergence of the demand (supply) lottery relates to a change in the lottery premium. We define the lottery premium of a demand (supply) lottery as the amount of money the incumbent is willing to pay (or has to receive) to undergo the lottery. In our model, it equals the difference between the expected outcome of undergoing the demand (supply) lottery and obtaining demand equal to $1-P$ (facing the cost I_1 of continuing R&D). The lottery premium of a favorable lottery is positive whereas the lottery premium of an unfavorable lottery is strictly negative, irrespective of the degree of divergence of both lotteries. The lottery premium of favorable lotteries with probability $\frac{1}{2}$ of the good/bad state (i.e. mean-preserving lotteries) is equal to 0, irrespective of their degree of divergence. When comparing two favorable, non mean-preserving lotteries with equal probabilities, the more divergent lottery entails a more positive lottery premium. Similarly, when comparing two unfavorable, non mean-preserving lotteries with equal probabilities, the more divergent lottery entails a more negative lottery premium. However, we cannot always conclude that the more divergent lottery entails a more positive (a more negative) lottery premium when the lotteries are favorable (unfavorable) but have unequal probabilities. It 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 lottery premia, it remains to show how a change in the lottery premium affects the decision to start R&D.

3.2 Relating lottery premia 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 $\theta < \theta'$ without finding 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. In other words, a decrease in the demand lottery premium never positively affects the decision to start R&D for these lotteries. 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. In other words, an increase in the demand lottery premium never negatively affects the decision to start R&D for these 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; see footnote 2 for definitions of v and w. All proofs are relegated to Appendix A.[3]

[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.

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 (= a decrease in the demand lottery premium) 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 (= a decrease in the demand lottery premium) 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 (= a decrease in the supply lottery premium) 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 (= a decrease in the supply lottery premium) negatively affects the decision to start R&D.

Propositions 3a, 3b and 4 provide important additional insight in the relation between demand (supply) lotteries and the decision to start R&D. Let us focus on demand lotteries. 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, a decrease in the demand lottery premium can either positively or negatively affect the decision to start R&D. Especially the fact that a decrease in the demand lottery premium 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, 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, a decrease in the lottery premium of an unfavorable demand lottery would never positively affect the decision to start R&D.[4]

4 An empirical analysis of the determinants of R&D decisions

4.1 Data

To test the propositions derived in the previous section, we mainly use data from the 2005 official innovation survey in the German manufacturing and services industries which constitute

[4]If the incumbent is forced to complete the R&D project once the project is started, he will only 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 lottery premia (cfr. Section 3.1).

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 data on 4776 firms were collected in 2005 (*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]

All explanatory variables which are explained in Section 4.2 are taken from the 2005 survey. In order to investigate how they affect the decision to start R&D, we merge information on R&D from the 2006 survey.[9] Combining the two surveys reduces the number of observations by 40.3 percent. For estimation purposes we further exclude firms with incomplete data for any of the relevant variables, ending up with a sample of 2411 firms. As illustrated in Table B.1 in Appendix B, our estimation sample (*full sample*) reflects *total sample* distributional characteristics very well and does not give any obvious cause for selectivity concerns.

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

5.1 Data

To test the propositions derived in the previous section, we mainly 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).[10] 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.[11] 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 data on 4776 firms were collected in 2005 (*total sample*), corresponding to a response

[5] The innovation surveys are annually 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).

[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 Germany, the innovation surveys are conducted annually and they are designed as a panel (Mannheim Innovation Panel).

[10] The innovation surveys are annually 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).

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

rate of about 20%.[12] 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.[13]

All explanatory variables which are explained in Section 4.2 are taken from the 2005 survey. In order to investigate how they affect the decision to start R&D, we merge information on R&D from the 2006 survey.[14] Combining the two surveys reduces the number of observations by 40.3 percent. For estimation purposes we further exclude firms with incomplete data for any of the relevant variables, ending up with a sample of 2411 firms. As illustrated in Table B.1 in Appendix B, our estimation sample (*full sample*) reflects *total sample* distributional characteristics very well and does not give any obvious cause for selectivity concerns.

5.2 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.[15] 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.

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.

Specification and 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 variable indicates whether the firm performed R&D in 2005 ($R\&D$). Table 1 shows that 45% of the firms in the full sample undertook R&D projects. Our theoretical model is essentially about cost-reducing process innovations. One drawback of $R\&D$ is that R&D activities cannot be divided into product and process innovations. Therefore, we employ as an alternative proxy a variable indicating whether the firm

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

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

[14] In Germany, the innovation surveys are conducted annually and they are designed as a panel (Mannheim Innovation Panel).

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

planned in the period 2002-2004 to introduce a new production technology in 2005 ($PROCESS$). We find that 46% of the firms in the full sample planned to introduce a process innovation.[16]

In our theoretical model, demand uncertainty stems from the two components in the demand lottery: the degree of divergence (determined by θ) and the probability (p_θ) of facing a good demand state. The variable θ is measured by the average of the absolute values of the absolute changes in real sales over the last two years 2002-2003 and 2003-2004 ($THETA$).[17] Table 1 reveals that the absolute value of the absolute change in real sales was on average about 13% in the last two years. To calculate p_θ using the full sample, we derive that 53.1% of the firms experienced a positive growth in sales between 2002 and 2003 and 57.4% between 2003 and 2004. 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 (determined 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 postponed 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.[18] 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 value of the deviation between on the one hand the innovation expenditures for 2004 expected in 2003 and on the other hand the realized innovation 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 it only to a subset of enterprises since we have to use the prior wave of the innovation survey to construct this variable.[19] However, this subsample is representative for the full sample as can be inferred from Table B.1 in Appendix B. The average absolute value of the deviation between expected and realized innovation expenditure comes to 2.9 mill. Euro. The deviation turns out to be highly skewed. We therefore use a logarithmic transformation of this variable in the econometric analysis. As $LAMBDA2$ measures the absolute value of the deviation between expected and realized innovation costs, we encounter the problem that $LAMBDA2$ equals zero for firms with no prior innovation activities. To filter out this effect, we additionally include $DLAMBDA2$ in the estimations. $DLAMBDA2$ is a binary variable which is one if the firm had no prior innovation activities. 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 60.1% of the firms, realized innovation expenditure in 2004 turned 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. In all our estimations, we assume that p_λ is the same for all firms. Our dataset does not allow to relax this assumption.

[16]In 71% of the observations, *R&D* and $PROCESS$ coincide. This implies a correlation between the two dependent variables of about 0.28 (significant at the 1% level).

[17]We use producer price indices at the 3-digit industry (NACE) level as a deflator. Furthermore, we implicitly assume that firms expect sales to stay constant over the short time-span under consideration.

[18]It might have been that the unexpected delay led to R&D activities in 2005. Though we cannot completely rule out this mechanical effect, we are confident that this is of minor importance in our data set. The delay took place somewhere in the period 2002-2004. Despite the delay the firm could still have finished the innovation project in that period. As a robustness check, we use an alternative R&D indicator. In particular, we only account for R&D activities in 2005 if the firm had no ongoing process innovation activities at the end of year 2004. This leaves the results nearly unchanged. Results are not reported here but are available upon request.

[19]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.

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: A higher demand lottery premium does not decrease the probability of undertaking R&D.

Hypothesis 2: A higher supply lottery premium does not decrease the probability of undertaking R&D.

Besides the importance of lottery premia, we demonstrate in Section 3.2 that having an abandonment option limits the downside risk of the R&D project. Unfortunately, we do not directly observe in our data whether a firm has the option to abandon an R&D project. However, we observe whether the firm abandoned innovation projects in the past three years 2002-2004. We use this information to construct the variable $ABAN$ which captures whether the firm has the option to abandon R&D projects in the following way. $ABAN$ equals 1 if a firm abandoned any innovation project in the past. For all other firms, it is the predicted probability derived from a probit regression.[20]

In our theoretical model, the incumbent is challenged by a potential competitor. Our data reveal that about 91% of the firms perceived 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 perceived a high, medium or low threat.

Our main explanatory variables, i.e. proxies for the degree of divergence in the demand and supply lotteries, abandonment option and entry threat, belong to the categories of respectively project characteristics and industry characteristics that are discussed in Section 1. We also control for the following factors found to be important in the literature. These can be mapped into the categories of firm characteristics and industry characteristics.

Among the *firm characteristics*, we include firm size ($SIZE$), corporate variables such as the degree of product diversification ($DIVERS$), innovative capabilities ($HIGHSKILLED$, $TRAINEXP$, $NOTRAIN$, $MVTRAIN$), the type of competition ($COMP$) and the degree of internationalization ($EXPORT$), and financial variables such as the availability of financial resources ($RATING$).

Firm size ($SIZE$) is measured by the logarithm of the number of employees in 2004. We expect a positive relationship between firm size and the decision to undertake an R&D project.

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.

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 did not invest in training its employees ($NOTRAIN$) and the amount of training expenditure per employee ($TRAINEXP$) if the firm invested in training. Since information on training expenditure is missing for 9.6% of the firms, we do not drop these observations but

[20] We explain the probability of abandoning innovation projects by industry dummies and firm characteristics such as firm size, share of high-skilled employees, degree of internationalization, training expenditure and company group (see Table B.3 in Appendix B).

rather set the expenditure to zero and include a dummy variable indicating the missing value status ($MVTRAIN$).

The incentive to engage in R&D may further depend on the type of competition ($COMP$). We include 3 dummy variables indicating whether firms primarily competed in prices, product quality or technological lead.

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 was exposed to international competition is captured by the export intensity in 2004, i.e. the ratio of exports to sales ($EXPORT$).

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.

We also include a variable reflecting whether the firm was located in East Germany ($EAST$). *A priori*, the effect of $EAST$ is unclear.

Among the *industry characteristics*, we include market structure ($NUMCOMP$) and general industry conditions. Market structure is captured by 3 dummy variables indicating the number of competitors. Schumpeter (1942) 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.

Finally, we control for general industry conditions by including industry dummies in all regressions.

Table 1
Descriptive Statistics - Full Sample

Variable	Unit	Mean	SD	Median	Skewness	Min	Max
Dependent variables							
R&D	[0/1]	0.454	0.498	0	–	0	1
PROCESS	[0/1]	0.460	0.498	0	–	0	1
Independent variables							
Demand lottery premium							
THETA	%	0.132	0.138	0.090	2.831	0.006	1.254
G1	[0/1]	0.129	0.335	0	–	0	1
G2	[0/1]	0.295	0.456	0	–	0	1
G3	[0/1]	0.371	0.483	0	–	0	1
G4	[0/1]	0.204	0.403	0	–	0	1
Supply lottery premium							
LAMBDA1	[0/1]	0.192	0.394	0	–	0	1
LAMBDA2[a]	Mill. Euro	2.887	14.448	0.117	9.013	0	207.318
DLAMBDA2[a]	[0/1]	0.229	0.420	0	–	0	1
Abandonment option							
ABAN	[0-1]	0.263	0.295	0.146	1.923	0.019	1
Additional control variables							
THREAT: no	[0/1]	0.093	0.295	0	–	0	1
THREAT: low	[0/1]	0.457	0.498	0	–	0	1
THREAT: medium	[0/1]	0.304	0.460	0	–	0	1
THREAT: high	[0/1]	0.146	0.353	0	–	0	1
SIZE	# Empl.	692.798	6505.776	46	24.251	1	232700
NUMCOMP: 0	[0/1]	0.024	0.153	0	–	0	1
NUMCOMP: 1-5	[0/1]	0.590	0.492	1	–	0	1
NUMCOMP: 6-15	[0/1]	0.204	0.403	0	–	0	1
NUMCOMP: >15	[0/1]	0.182	0.386	0	–	0	1
COMP: PRICE	[0/1]	0.526	0.499	1	–	0	1
COMP: QUAL	[0/1]	0.427	0.495	0	–	0	1
COMP: LEAD	[0/1]	0.107	0.310	0	–	0	1
DIVERS	[0-1]	0.709	0.234	0.739	-0.476	0	1
EXPORT	[0-1]	0.152	0.311	0.000	13.140	0	9.804
RATING	[1-6]	2.137	0.794	2.180	0.526	0	6.000
HIGHSKILLED	[0-100]	20.824	24.579	10	1.576	0	100
TRAINEXP	Mill. Euro	0.001	0.001	0	4.468	0	0.010
NOTRAIN	[0/1]	0.112	0.315	0	–	0	1
MVTRAIN	[0/1]	0.096	0.294	0	–	0	1
EAST	[0/1]	0.339	0.473	0	–	0	1

[a] Values refer to subsample of 767 firms.
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.

5.3 Results

5.3.1 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 of a parsimonious specification —including only $SIZE$ and industry dummies in addition to demand uncertainty, technical uncertainty, abandonment option and entry threat— whereas the second column employs the full set of control variables described in the previous section.

Hypothesis 1, implying that the probability of undertaking R&D does not decrease with an increase in θ, is confirmed. The effect is positive but only significant in $PROCESS$ (column (3)) implying that a 10% increase in $THETA$ increases the likelihood of undertaking R&D by 1.3 percentage points. The result that $THETA$ is only slightly significant may well reflect the fact that our assumption of equal demand lottery probabilities is not fulfilled in our sample. We elaborate on this point in Section 4.3.2.

Hypothesis 2, implying 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 (captured by a shift in the value of $LAMBDA1$) increases the likelihood of undertaking R&D ($R\&D$, column (2)) by 14.6 percentage points.

Having the option to abandon R&D projects significantly increases the likelihood of undertaking R&D. Focusing on column (2), the marginal effect amounts to 12 percentage points.

We do not find a significant effect of entry threat on the decision to undertake R&D. Regarding the impact of the other control variables, firm size exerts a significantly positive impact. Firms being exposed to international competition as well as more diversified firms have a higher likelihood of undertaking R&D. There is, however, no significant impact on process innovation. Highlighting the important role of innovative capabilities, we find that firms employing a higher share of high-skilled workers or firms investing in training are likely to be more innovative. Innovation activities are stimulated if competitive advantage is achieved by technological leadership. Our estimates do not confirm an impact of market structure on innovation.

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 $LAMBDA1$ which is still significantly positive. In general, the results are very similar to the full sample. Hypothesis 2 is also confirmed using $LAMBDA2$. Since we measure this variable in logarithm, a value of 0.007 implies that an increase in the absolute value of the deviation between expected and realized innovation expenditure by 10% increases the propensity to undertake R&D by 7%.[21][22]

[21] To test whether multicollinearity between our main independent variables affect our results in Tables 2 and 3, we estimate specifications that include demand lottery premium, supply lottery premium or abandonment option separately and specifications that combine demand lottery premium or supply lottery premium with abandonment option. 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).

[22] To test whether multicollinearity between our main independent variables affect our results in Tables 2 and 3, we estimate specifications that include demand lottery premium, supply lottery premium or abandonment option separately and specifications that combine demand lottery premium or supply lottery premium with abandonment option. 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).

Table 2
Effect of demand and supply lottery premium on innovation - Full Sample

Dep. variables	R&D		PROCESS	
	(1)	(2)	(3)	(4)
Demand lottery premium				
THETA	0.082	-0.009	0.128*	0.113
	(0.063)	(0.061)	(0.073)	(0.071)
Supply lottery premium				
LAMBDA1	0.179***	0.146***	0.206***	0.186***
	(0.025)	(0.024)	(0.027)	(0.027)
Abandonment option				
ABAN	0.199***	0.120***	0.153***	0.119***
	(0.035)	(0.032)	(0.038)	(0.038)
Additional control variables				
THREAT: low	-0.038	-0.029	0.007	-0.008
	(0.031)	(0.030)	(0.036)	(0.036)
THREAT: medium	-0.051	-0.029	0.022	0.018
	(0.032)	(0.032)	(0.038)	(0.038)
THREAT: high	-0.081**	-0.035	-0.049	-0.026
	(0.036)	(0.036)	(0.042)	(0.043)
SIZE	0.062***	0.059***	0.060***	0.060***
	(0.006)	(0.006)	(0.006)	(0.007)
NUMCOMP: 0	–	-0.055	–	-0.079
		(0.054)		(0.066)
NUMCOMP: 1-5	–	0.032	–	0.003
		(0.022)		(0.027)
NUMCOMP: 6-15	–	0.003	–	-0.004
		(0.026)		(0.032)
COMP: PRICE	–	-0.024	–	-0.083***
		(0.019)		(0.023)
COMP: QUAL	–	0.017	–	0.000
		(0.018)		(0.022)
COMP: LEAD	–	0.113***	–	-0.014
		(0.028)		(0.033)
DIVERS	–	-0.083**	–	-0.032
		(0.035)		(0.043)
EXPORT	–	0.233***	–	-0.047
		(0.041)		(0.032)
RATING	–	0.001	–	-0.009
		(0.010)		(0.013)
HIGHSKILLED	–	0.002***	–	0.000
		(0.000)		(0.001)
TRAINEXP	–	0.058***	–	0.058***
		(0.008)		(0.010)
NOTRAIN	–	-0.461***	–	-0.471***
		(0.026)		(0.027)
MVTRAIN	–	-0.336***	–	-0.400***
		(0.032)		(0.031)
EAST	–	0.064***	–	0.028
		(0.018)		(0.022)
$LogL$	-1235.1	-1127.1	-1360.4	-1311.4
R^2_{MF}	0.239	0.296	0.085	0.108
R^2_{MZ}	0.453	0.554	0.212	0.273
Count R^2	0.747	0.769	0.660	0.678
LM_{het} (p-value)	0.339	0.457	0.338	0.621
LM_{norm} (p-value)	0.131	0.899	0.976	0.877
# Obs.	2411	2411	2201	2201

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 supply lottery premium on innovation - Subsample

Dep. variables	**R&D**		**PROCESS**	
	(1)	(2)	(3)	(4)
Demand lottery premium				
THETA	0.070	0.117	0.105	0.177
	(0.112)	(0.107)	(0.140)	(0.116)
Supply lottery premium				
LAMBDA1	0.106**	–	0.084*	–
	(0.041)		(0.046)	
LAMBDA2	–	0.007*	–	-0.005
		(0.004)		(0.004)
DLAMBDA2	–	-0.218***	–	-0.516***
		(0.071)		(0.040)
Abandonment option				
ABAN	0.121**	0.102*	0.150**	0.111*
	(0.061)	(0.055)	(0.069)	(0.062)
Additional control variables				
THREAT: low	-0.003	0.006	-0.091	-0.109*
	(0.052)	(0.050)	(0.058)	(0.059)
THREAT: medium	0.010	0.008	-0.055	-0.082
	(0.054)	(0.051)	(0.062)	(0.063)
THREAT: high	-0.064	-0.053	-0.192***	-0.191***
	(0.062)	(0.059)	(0.067)	(0.070)
SIZE	0.051***	0.040***	0.061***	0.050***
	(0.010)	(0.010)	(0.013)	(0.012)
NUMCOMP: 0	-0.112	-0.067	0.022	0.148
	(0.087)	(0.091)	(0.127)	(0.107)
NUMCOMP: 1-5	0.095**	0.074**	-0.058	-0.077*
	(0.038)	(0.037)	(0.048)	(0.046)
NUMCOMP: 6-15	0.052	0.039	0.026	-0.001
	(0.042)	(0.040)	(0.055)	(0.052)
COMP: PRICE	-0.028	-0.026	-0.099**	-0.085**
	(0.033)	(0.031)	(0.040)	(0.036)
COMP: QUAL	0.011	-0.007	0.012	-0.005
	(0.031)	(0.029)	(0.038)	(0.035)
COMP: LEAD	0.063	0.021	0.020	-0.013
	(0.046)	(0.041)	(0.056)	(0.050)
DIVERS	-0.090	-0.055	0.024	0.034
	(0.060)	(0.057)	(0.073)	(0.069)
EXPORT	0.319***	0.302***	-0.026	-0.031
	(0.078)	(0.073)	(0.038)	(0.075)
RATING	-0.010	-0.008	-0.010	-0.009
	(0.017)	(0.016)	(0.022)	(0.020)
HIGHSKILLED	0.001*	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
TRAINEXP	0.081***	0.060***	0.067***	0.046***
	(0.014)	(0.014)	(0.017)	(0.016)
NOTRAIN	-0.563***	-0.516***	-0.478***	-0.356***
	(0.025)	(0.051)	(0.043)	(0.098)
MVTRAIN	-0.419***	-0.360***	-0.461***	-0.423***
	(0.046)	(0.074)	(0.034)	(0.059)
EAST	0.051*	0.059**	0.005	0.011
	(0.030)	(0.029)	(0.040)	(0.038)
$LogL$	-339.6	-316.4	-412.2	-362.9
R^2_{MF}	0.360	0.404	0.157	0.258
R^2_{MZ}	0.601	0.645	0.304	0.495
Count R^2	0.787	0.816	0.686	0.721
LM_{het} (p-value)	0.583	0.057	0.957	0.978
LM_{norm} (p-value)	0.060	0.330	0.101	0.259
# Obs.	767	767	707	707

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.

5.3.2 Firms facing different demand lottery probabilities

In Section 3.2, we show that the effect of an increase in θ on the decision to start R&D depends, *ceteris paribus*, on p_θ. So far, we assumed that p_θ is the same for all firms. In this section, we relax this assumption. We approximate p_θ by looking at the firms' sales histories in the past three years. We define four 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, in 2003-2004 as well as in 2004-2005. The idea is that these firms always face an unfavorable demand lottery and have a p_θ around 0. Group 2 ($G2$) consists of all firms that experienced two negative and one positive demand shock during the period 2002-2005. We assume that these firms have a p_θ around $\frac{1}{3}$. All firms in group 3 ($G3$) experienced one negative and two positive demand shocks during the period 2002-2005. The assumption is that these firms are more likely to face a favorable demand lottery reflected by a p_θ around $\frac{2}{3}$. Group 4 ($G4$) consists of all firms that experienced three consecutive increases in sales during the period 2002-2005. The idea is that these firms always face a favorable demand lottery and have a p_θ around 1.

We postulate from Proposition 3a and Proposition 2 respectively the following hypotheses.

Hypothesis 3: For firms in $G1$, a lower demand lottery premium does not increase the probability of undertaking R&D.

Hypothesis 4: For firms in $G3$ and $G4$, a higher demand lottery premium does not decrease the probability of undertaking R&D.

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 $G1$ the effect of a lower demand lottery premium (= an increase in θ) is significantly negative for $R\&D$ and negative but not significant for $PROCESS$ in the specifications including all control variables. Focusing on column (2), our results indicate that a 10% increase in $THETA$ decreases the likelihood of undertaking R&D by 4.9 percentage points. Furthermore, the impact of $THETA$ is significantly different for firms in $G1$ compared to firms in $G2$, $G3$ and $G4$. Hypothesis 4 is strongly confirmed since the impact of a higher demand lottery premium is never significantly negative for firms in $G3$ and $G4$. Moreover, in all specifications, the effect of a higher demand lottery premium is significantly positive for firms in $G4$. Focusing on column (2), an increase in $THETA$ by 10% increases the probability of undertaking R&D by 1.7 percentage points for firms in $G4$.

From Proposition 3a, Proposition 4 and footnote 4, it follows that for the set of unfavorable demand lotteries with $p_\theta \in [\frac{1}{4}, \frac{1}{2})$, a decrease in the lottery premium *can* positively affect the decision to start R&D because of the abandonment option that the firm possesses. We therefore now consider the firms in $G2$ in isolation. We postulate the following hypothesis.

Hypothesis 5: For firms in $G2$, having an abandonment option does not decrease the probability of undertaking R&D.

Confirming hypothesis 5, having the option to abandon R&D projects significantly increases the likelihood of undertaking an R&D project when using $R\&D$ as the dependent variable. Focusing on column (2), the marginal effect amounts to 21 percentage points (see Table 5).

Table 4
Effect of demand and supply lottery premium on innovation across groups of firms facing a different p_θ - Full sample

Dep. variables	**R&D**		**PROCESS**	
	(1)	(2)	(3)	(4)
Demand lottery premium				
THETA*G1	-0.474**	-0.494**	-0.247	-0.112
	(0.214)	(0.225)	(0.234)	(0.228)
THETA*G2	-0.109	-0.158	-0.023	0.023
	(0.099)	(0.098)	(0.124)	(0.121)
THETA*G3	0.068	0.019	0.021	0.011
	(0.073)	(0.070)	(0.084)	(0.083)
THETA*G4	0.313***	0.171*	0.616***	0.544***
	(0.111)	(0.101)	(0.137)	(0.135)
$\alpha_{\theta*G1} >= \alpha_{\theta*G2}$ (p-value)	0.045	0.068	0.176	0.282
$\alpha_{\theta*G1} >= \alpha_{\theta*G3}$ (p-value)	0.005	0.016	0.124	0.292
$\alpha_{\theta*G1} >= \alpha_{\theta*G4}$ (p-value)	0.000	0.002	0.000	0.004
$\alpha_{\theta*G2} >= \alpha_{\theta*G3}$ (p-value)	0.048	0.089	0.368	0.537
$\alpha_{\theta*G2} >= \alpha_{\theta*G4}$ (p-value)	0.001	0.004	0.000	0.001
$\alpha_{\theta*G3} >= \alpha_{\theta*G4}$ (p-value)	0.019	0.040	0.000	0.000
Supply lottery premium				
LAMBDA1	0.173***	0.141***	0.201***	0.183***
	(0.024)	(0.024)	(0.027)	(0.027)
Abandonment option				
ABAN	0.200***	0.123***	0.151***	0.118***
	(0.035)	(0.032)	(0.038)	(0.038)
$LogL$	-1225.7	-1120.3	-1347.6	-1302.1
R^2_{MF}	0.243	0.298	0.092	0.112
R^2_{MZ}	0.461	0.560	0.227	0.283
Count R^2	0.754	0.768	0.666	0.684
LM_{het} (p-value)	0.358	0.412	0.506	0.778
LM_{norm} (p-value)	0.224	0.068	0.867	0.282
# Obs.	2411	2411	2201	2201

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) $SIZE$, $THREAT$ 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.

Table 5
Effect of demand and supply lottery premium on innovation - Subsample of firms belonging to $G2$

Dep. variables	**R&D**		**PROCESS**	
	(1)	(2)	(3)	(4)
Demand lottery premium				
THETA	-0.029	-0.137	0.040	0.111
	(0.122)	(0.119)	(0.152)	(0.149)
Supply lottery premium				
LAMBDA1	0.171***	0.125***	0.282***	0.264***
	(0.047)	(0.045)	(0.049)	(0.050)
Abandonment option				
ABAN	0.316***	0.210***	0.120*	0.071
	(0.060)	(0.055)	(0.070)	(0.073)
$LogL$	-349.2	-308.0	-372.5	-352.3
R^2_{MF}	0.199	0.257	0.085	0.099
R^2_{MZ}	0.459	0.625	0.297	0.381
Count R^2	0.759	0.807	0.686	0.707
LM_{het} (p-value)	0.425	0.056	0.035	0.070
LM_{norm} (p-value)	0.367	0.738	0.000	0.013
# Obs.	715	715	646	646

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) $SIZE$, $THREAT$ 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.

6 Conclusion

The novelty of this article lies in combining a theoretical and empirical analysis on the determinants of R&D decisions.

From a theoretical point of view, we develop a model that contains many aspects of real-life R&D decisions within a net present value framework. Besides entry threat, Bertrand competition and multi-stage R&D with an abandonment option, our model includes demand uncertainty as well as technical uncertainty, both modelled as a lottery. Each lottery becomes more divergent when the difference between the outcomes of the lottery increases. We relate differences in the degree of divergence to differences in lottery 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. The presence of a potential entrant in our model provides the incumbent with additional benefits from undertaking R&D, a strategic effect known as Arrow's replacement effect. 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, we derive under which lottery probabilities more divergent demand and supply lotteries positively or negatively affect the decision to start R&D. Using CIS IV data for about 2400 German firms, we find 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 1.7 percentage points 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 14.1 percentage points. For firms facing a demand lottery where the bad state is most likely to prevail, a 10% increase in the degree of divergence of the demand lottery decreases the likelihood of undertaking R&D by 4.9 percentage points. A striking result of our theoretical analysis is that a decrease in the lottery premium of an unfavorable demand lottery can positively affect the decision to start R&D due to the abandonment option that the incumbent possesses. We estimate that having the option to abandon R&D projects significantly increases the likelihood of undertaking an R&D project. The marginal effect amounts to 21 percentage points.

Our analysis can be extended in several promising ways. An obvious research avenue is to replace the monopolist threatened by entry in our model by an oligopolistic market structure. The distinction can be important since an oligopolistic setting makes the analysis of R&D incentives more involved. Hence, Cournot competition should be considered and the distinction between drastic and non-drastic innovation should be studied. Furthermore, it would be interesting to investigate how sensitive our results are to differences in the degree of entry threat. Also, a welfare analysis of the social desirability of undertaking R&D in our setting can be conducted. Now, our model is essentially about cost-reducing process innovations. Another research avenue is to consider the development of a new product. This would necessitate an analysis of a differentiated product setting. The current availability of data on product innovations in the CIS surveys would straightforwardly allow an empirical justification. However, we should be aware of the defect that CIS data are related to firms and not to specific R&D projects. Ideally, we would obtain a closer match between our theoretical model and our empirical analysis when project-specific data are available.

References

[1] Acemoglu, D., 2008, Oligarchic versus democratic societies, Journal of the European Economic Association, 6(1), 1-44.

[2] Acs, Z.J. and D.B. Audretsch, 1987, Innovation, market structure and firm size, Review of Economic and Statistics, 69(4), 567-575.

[3] 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.

[4] Aghion, P., R. Blundell, R. Griffith, P. Howitt and S. Prantl, 2009, The effects of entry on incumbent innovation and productivity, Review of Economics and Statistics, 91(1), 20-32.

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

[6] Bell, G. and J. Campa, 1997, Irreversible investment and volatile markets: A study of the chemical processing industry, Review of Economics and Statistics, 79(1), 79-87.

[7] Berger, P., E. Ofek and I. Swary, 1996, Investor valuation of the abandonment option, Journal of Financial Economics, 42(2), 257-287.

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

[9] Bulan, T., 2005, Real options, irreversible investment and firm uncertainty: New evidence from U.S. firms, Review of Financial Economics, 14(3-4), 255-279.

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

[11] Cassiman, B. and R. Veugelers, 2002, R&D cooperation and spillovers: Some empirical evidence from Belgium, American Economic Review, 92(4), 1169-1184.

[12] Czarnitzki D. and A.A. Toole, 2009, R&D investment under uncertainty: The role of R&D subsidies and patent policy, in: Foray, D. (Ed.), The economics of technology policy, Cheltenham : Edward Elgar.

[13] Czarnitzki, D. and A.A. Toole, 2010, Patent protection, market uncertainty, and R&D investment, Review of Economics and Statistics, forthcoming.

[14] Cohen, W.M. and R.C. Levin, 1989, Empirical studies of innovation and market structure, in: Schmalensee, R., Willig R. (Eds.), Handbook of Industrial Organization, vol. 2, Amsterdam: North Holland.

[15] Cohen, W.M and D.A. Levinthal, 1989, Innovation and learning: The two faces of R&D, Economic Journal, 99(397), 569-596.

[16] 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.

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

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

[19] Dorfman, N.S., 1987, Innovation and market structure: Lessons from the computer and semiconductor industries, Cambridge, M.A.: Ballinger.

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

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

[22] Etro, F., 2006, Aggressive leaders, RAND Journal of Economics, 37(1), 46-54.

[23] Fazzari, S.M. and M.J. Athey, 1987, Asymmetric information, financing constraints, and investment, Review of Economics and Statistics, 69(3), 481-487.

[24] 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.

[25] Gabrowski, H., 1968, The determinants of industrial research and development: A study of the chemical, drug, and petroleum industries, Journal of Political Economy, 76(2), 292-306.

[26] 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.

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

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

[29] 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.

[30] Huisman, K.J.M., P.M. Kort, G. Pawlina and J.J.J. Thijssen, 2005, Strategic investment under uncertainty: A survey of game theoretic real option models, Journal of Financial Transformation, 13, 111-118.

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

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

[33] 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.

[34] Levin R.C., W.M. Cohen and D.C. Mowery, 1985, R&D appropriability, opportunity, and market structure: New evidence on some Schumpeterian hypotheses, American Economic Review, 75(2), 20-24.

[35] 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.

[36] Myers S.C. and S. Majd, 1990, Abandonment value and project life, Advances in Futures and Options Research, 4(1), 1-12.

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

[38] Schumpeter, J.A., 1942, Capitalism, socialism and democracy, New York: Harper.

[39] Segerstrom, P., 2007, Intel economics, International Economic Review, 48(1), 247-280.

[40] Trigeorgis, L., 1996, Real options, Cambridge, M.A.: MIT Press.

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

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

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[a]	Full Sample[b]	Subsample[c]
Industry			
Food/tobacco	3.16	2.78	2.60
Textiles	2.97	2.70	2.34
Paper/wood/print	6.7	6.35	5.33
Chemicals	4.1	3.94	5.07
Plastic/rubber	3.62	3.94	3.77
Glass/ceramics	2.14	2.57	3.12
Metal	8.35	8.67	9.36
Machinery	5.99	6.93	7.28
Electrical engineering	4.88	5.23	6.24
Medical, precision and optical instruments	4.92	5.47	6.11
Vehicles	2.66	2.65	2.60
Furniture	2.62	2.86	2.21
Wholesale	4.38	3.98	4.16
Retail	2.35	2.20	2.34
Transport/storage/post	8.46	8.05	5.72
Banks/insurances	5.05	4.02	3.90
Computer/telecommunication	4.59	5.02	5.20
Technical services	8.79	9.79	11.18
Consultancies	3.77	3.32	3.38
Other business related services	7.06	6.72	5.33
Real estate/renting	2.07	1.91	2.47
Media	1.38	0.91	0.26
Size (Number of employees)			
0-4	4.65	3.36	3.38
5-9	14.24	12.44	12.22
10-19	16.52	16.18	14.30
20-49	18.68	19.58	21.33
50-99	13.13	13.15	13.52
100-199	14.07	14.52	13.91
200-499	7.96	8.75	8.84
500-999	4.98	5.14	5.33
1000+	5.78	6.89	7.15
Region			
West Germany	66.86	66.11	63.07
East Germany	33.14	33.89	36.93
Innovation activities			
Non-innovators	36.12	33.26	31.47
Innovators[d]	63.88	66.74	68.53
# Obs.	4776	2411	769

[a] Total sample refers to the net sample of the 2005 survey.
[b] Full sample denotes the estimation sample which is based on a merge of the 2005 and 2006 survey, excluding firms with missing values.
[c] Subsample marks the estimation sample of firms which have answered the 2004, 2005 and 2006 survey, excluding firms with missing values.
[d] 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 year 2005.
PROCESS	0/1	1 if the firm planned to undertake process innovations in year 2005.
Independent variables		
Demand lottery premium		
THETA	c	Average of the absolute values of the absolute changes in real sales over the last two years (2002/2003 and 2003/2004)
G1	0/1	1 if the firm experienced three negative demand shocks in the past three years, i.e. a decrease in sales in 2002/2003, 2003/2004 and 2004/2005.
G2	0/1	1 if the firm experienced two negative and one positive demand shock in the past three years, i.e. two times a decrease and one increase in sales within the last three years.
G3	0/1	1 if the firm experienced one negative and two positive demand shocks in the past three years, i.e. one decrease and two times an increase in sales within the last three years.
G4	0/1	1 if the firm experienced three positive demand shocks in the past three years, i.e. a positive growth in sales in 2002/2003, 2003/2004 and 2004/2005.
Supply lottery premium		
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 value of the deviation between in year 2003 expected innovation expenditure for 2004 and realized innovation expenditure in 2004, in log.
DLAMBDA2	0/1	1 if a firm did not observe uncertainty in R&D costs, i.e. if a firm expected zero innovation expenditure in 2003 for 2004 and realized zero innovation expenditure in 2004. This corresponds to LAMBDA2=0 (without taking logs).
Abandonment option		
ABAN	0-1	Variable proxying whether a firm had the option to abandon R&D projects. It equals 1 if a firm abandoned innovation projects in the past three years 2002-2004. For all other firms it is the predicted value of a probit regression explaining the probability of abandoning innovation projects in the past three years (detailed regression results are given in Table B.3).
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 2004, 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	3 dummy variables indicating the most important factors of competition: price, quality and technological lead (multiple factors allowed).
DIVERS	0-100	Degree of product diversification, measured as the share of turnover of most important product in 2004.
EXPORT	0-1	Export intensity, measured as ratio of exports to sales in 2004.
RATING	c	Credit rating index of the firm in year 2004, ranging between 1 (highest) and 6 (worst creditworthiness).
HIGHSKILLED	0-100	Share of employees with a university or college degree in 2004.
NOTRAIN	0/1	1 if the firm did not invest in training in 2004.
TRAINEXP	c	Training expenditure per employee in 2004 (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 was located in East Germany.

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

Table B.3
Probability of having abandoned an innovation project in years 2002-2004 - Probit estimation

Dep. variables	**ABANDON**	
	marginal effect	(s.e.)
SIZE	0.030***	(0.005)
HIGHSKILLED	0.0008*	(0.0004)
TRAINEXP	0.005*	(0.003)
NOTRAIN	-0.059	(0.036)
GROUP	0.031*	(0.019)
EXPORT	0.038	(0.024)
Industry dummies (reference: food/tobacco)		
Textiles	0.032	(0.076)
Paper/wood/print	-0.090*	(0.037)
Chemicals	-0.024	(0.052)
Plastic/rubber	0.020	(0.064)
Glass/ceramics	-0.113**	(0.034)
Metal	-0.092**	(0.036)
Machinery	-0.076	(0.040)
Electrical engineering	-0.084*	(0.038)
Medical, precision and optical instruments	-0.079	(0.040)
Vehicles	-0.051	(0.050)
Furniture	-0.106**	(0.037)
Wholesale	-0.092*	(0.041)
Retail	-0.056	(0.064)
Transport/storage/post	-0.119***	(0.030)
Banks/insurances	-0.138***	(0.024)
Computer/telecommunication	-0.082	(0.041)
Technical services	-0.109***	(0.035)
Consultancies	-0.043	(0.057)
Other business related services	-0.078	(0.041)
Real estate/renting	-0.100	(0.056)
Media	0.077	(0.122)
$LogL$	-815.5	
R^2_{MF}	0.066	
R^2_{MZ}	0.116	
Count R^2	0.825	
# Obs.	1870	

ABANDON is a dummy variable which equals 1 if a firm with innovation activities has abandoned innovation projects in the three years 2002-2004. Based on this regression the predicted value ABAN is constructed as described in Table B.2.
Average marginal effects of the probit estimations are reported. Robust standard errors in parentheses.
***Significant at 1%; **Significant at 5%; *Significant at 10%.
For notes on goodness-of-fit and specification tests: see Table 2.