Essays on Delegated Search and Temporary Work Agencies

Tomas Raattamaa

Department of Economics
Umeå School of Business and Economics
Umeå University
Doctoral thesis 2016
To my family
Abstract

**Paper [I]** models a game, where two temporary work agencies (TWAs) compete to fill a vacancy at a client firm (CF). They simultaneously choose how much effort to expend, based on their expectation of how good their opponents best candidate will be. I then show that this will make the TWAs overconfident, as the rational way of judging your own probability of winning is not looking at the opponents expected best, but comparing how much effort your opponent will expend.

**Paper [II]** examines the misaligned incentives in the temporary work agency sector, where we first look at pure recruiting contracts, that either require payment on delivery, or payment on some specified point in time. We then look at the incentives of recruit-and-rent contracts, where the worker is leased to the client firm. We assume that the better the worker, the higher the probability that the client firm is going to want to hire him/her. If that happens then the TWA will no longer get revenues from said worker, incentivizing the TWA to not always deliver the first match it finds, if it is too good. Lastly we look at how competition can dampen this perverse incentive.

**Paper [III]** models the waiting behavior that can occur if a TWA is contracted to find a worker for a specific time far in the future; the TWA will postpone effort. This behavior is modeled for two types of TWAs; one that is rational and plans ahead, and another that does not plan ahead at all, but instead only looks at the immediate future. I find that the one that only looks at the immediate future starts exerting effort earlier than the planner. After looking at optimal contracts under perfect monitoring and hidden action I provide two extensions. I first show that for the principal to want to delegate search to a rational TWA, the agent has to be better than the CF, by some factor, as it has to make up in efficiency what the principal loses in moral hazard, when the agent waits longer than the principal would like it to. Lastly I prove that it is profit maximizing for the principal to contract one agent and give it a deadline earlier than when the principal would need the worker, and then replace that agent with a competitor if the first one has not succeeded by that earlier deadline.

**Paper [IV]** estimates at the effect of family experience on relative transition probability into the temporary work agency sector. Using register data for all of Sweden we run a bias-reduced logistic regression, where we include various factors that affect the probability of young adults (aged 18-34) entering the sector. This paper ties in to the literature on occupational inheritance, as well as the literature on changing social norms. We find that having had a parent, sibling or partner in the TWA sector increases your probability of entering the sector yourself.

**Keywords:** delegated search, principal-agent, matching, transition probability, temporary work agencies
Acknowledgments

My sincerest appreciation and gratitude goes to my supervisor Magnus Wikström and co-supervisor Thomas Aronsson. I could not have asked for a better team to guide me through the writing process. This thesis has also benefited greatly from the comments and suggestions received from discussants Tomas Sjögren and Emma Zetterdahl. Thank you to Urban Lindgren and Morgan Westéus, who co-authored one and two of the papers respectively.

I am grateful to Ulf and Kenneth for teaching me how to teach, you are surely great role models in this respect – and many others – and Niklas for then trusting me enough to let me try my own wings. Morgan, Erik G and Christian, since you left I have missed our discussions on and off the clock. Usually off. Our contributions to science and humanity shall not be forgotten, if you remember them. Morgan, who is a co-author on two of the papers is the reason this thesis turned out the way it did, as he nudged me in this direction. It was a pleasure working with you, and your drive inspired me more than you’ll ever know. Erik B, Umeå is not the same without Swenglish adventures, late night NFL games, and terrifying live in-the-flesh basketball; it was all great fun – even though I believe we both still stick to our positions on N=1 experiments. The guys from freshman year; I can’t believe that after all these years, we still talk about fine dining establishments when we meet. André, I believe it has been mutually beneficial for us to take walks, with the objective of consuming both sunlight and caffeine. To Ulf, Sofia, Mathilda and Christian all I have to say is that I still dont restrict my vocabulary when describing wine, but I use different words now. Co-philosopher Calle, I feel privileged that you have said that your door is always open. And if not... I have your spare key. Thank you to Jocke and Stefan for supplying everything from 100% action to miniature golf. Many thanks to Linda and Elin, for always helping me and answering my recurring questions, and thanks to everyone else at the department, past and present, that have been along for the ride. Catia, I still have that Dutch Christmas poem you wrote, pinned to the board in my office. Once all of this has been submitted, I will read it again.

Mum and dad, thank you for always letting me do my own thing. In addition, I would like to thank my friends outside academia. You know who you are. Last but not least, words cannot express how grateful I am to have Carro in my life. Putting things into perspective, I found the following quote in a book on the philosophy of economics:

“It seems likely that economics has nothing to do with most of the universe. It applies only where some systems husband energy into some stores that they selectively spend in pursuit of outcomes that bring their current states or prospects into closer conformity with their preferences.” (D. Ross, 2014)

I am happy to be part of such a system, and make a dent in it, thanks to all of you.

Umeå, August 2016
Tomas
Contents

This thesis consists of an introduction and the following self-contained papers on the topic of delegated search and temporary work agencies:

Paper [I]

Paper [II]

Paper [III]

Paper [IV]
1 Introduction

“Social behavior, particularly in small groups, is more complex, and norms of behavior that are culturally induced or developed over time play a huge role in shaping societies. However, it would be foolish not to recognize the role of private incentives in motivating behavior in addition to these cultural phenomena.” Laffont and Martinort (2002)

This thesis contains three theoretical papers and one empirical; Paper [IV] uses register-based data to analyze which factors determine if a young adult in Sweden enters the temporary work sector, with a focus on the implications of family experience, and Papers [I]-[III] all present new game-theoretical models on the topic of delegated search, which is the business that temporary work agencies are in. While the theoretical papers are generalizable to a number of other settings, the common theme is that they use the temporary work sector as the main example.

The first studies on games in economic literature were those of Cournot (1838), Bertrand (1883) and Edgeworth (1897), but it was not until Von Neumann and Morgenstern (1944) that a general theory of games was introduced (Fudenberg & Tirole, 1991). A few years later Nash (1950) submitted a paper on what we now know as a Nash Equilibrium. John Nash’s life, and his discovery, has been made into a movie titled A Beautiful Mind, and it has also been the subject of a book titled A Beautiful Math (Siegfried, 2006), which focuses more on the development of the theory.

However, when most people think about economics they do not think about games; instead they think about supply and demand, and may even visualize a diagram where the two curves cross. This is called an equilibrium, as a market for one good strives towards that point. General equilibrium models (that extend to many markets that all strive for equilibrium) were proved in the 1950’s and 1960’s, and interaction in these models happens through the price mechanism. This was soon realized to be limited, and so the economics of information emerged.¹ The paper “Information and the Change in the Paradigm in Economics” (Stiglitz, 2002), based on the lecture Joseph E. Stiglitz delivered when he received the Bank of Sweden Prize in Economic sciences in Memory of Alfred Nobel in 2001, is an easy read on the evolution of the role of information in economics. In short, he (like Salanié, 2005) argues that we have gone from a competitive paradigm to the information paradigm.

One class of problems to solve, that sprung up as an area of research due to this paradigm shift, is agency problems, where we have two actors with conflicting goals. One actor, called the principal, contract another actor, the agent, to perform a task. In the (economic) relationship between the principal and the agent there is often asymmetric information and so the agent can use his private information

¹Salanié (2005)
to his advantage.\textsuperscript{2} This information asymmetry consists of either the principal not knowing the agent’s type (e.g. how good is it at the task at hand?) or it cannot verify the agent’s actions (e.g. will it do exactly what it it contracted to do?). The principal is therefore sometimes called the uninformed party, and the agent the informed party, to make this distinction in the amount of information they possess. Examples of principal-agent (PA) relationships are patient - doctor, voter - politician, house buyer - real-estate agent and client firm - temporary work agency.

The interaction between these two actors is typically modeled as a Stackelberg game where the principal is the leader and the agent is the follower, and they interact by entering into a contractual agreement.

The principal-agent relationship then dictates that there is no bargaining, as the principal offers the agent a contract it can either accept or reject. This “take or leave it” offer is a simplification of the bargaining process that we expect to take place in the real world, because if the agent rejects the contract then the parties will have no further interaction as the game would end. The goal of the principal is therefore to draft a contract that is as favorable as possible for its own ends, with the restriction that the agent must be willing accept said contract.

When this theory evolved we needed to forsake the general equilibrium models, in favor of game theory, and thus the theory of contracts started to form. The move away from general equilibrium was partly due to economists having to model the strategic interaction between the actors, but it was hoped that these new studies could later be integrated into (a better) theory of general equilibrium (Salanié, 2005).

Agency theory originated in the 1970’s and S. A. Ross (1973) is often credited for introducing the terminology, but for longer exposes of the history of Agency theory see Eisenhardt (1989) and Mitnick (2011). Today searching for “agency theory” on Google Scholar\textsuperscript{3} yields about 100,000 results, and “principal-agent” about 170,000, so a lot of research has been done in the last 40+ years.

Time in these models can either be modeled as discrete or as continuous, and actions can be taken either simultaneously or sequentially. Each of the three theoretical papers in this thesis has a different setup, but all have temporary work agencies (TWAs) as agents, and the client firm (CF) as the principal; Paper [I] is a simultaneous game where two TWAs compete to provide the best worker, Paper [II] uses discrete time where two or more TWAs search for workers and stop when one has presented a qualified worker to the client firm, and Paper [III] is set in continuous time where the TWA is contracted to supply a qualified worker by a given deadline.

Good introductions to PA theory can be found in Fudenberg and Tirole (1991), Laffont and Martinort (2002) and Salanié (2005) and for a good introduction to continuous-time PA models see Cvitanić and Zhang (2013).

Why should we model the PA relationship?

\textsuperscript{2}For a history see Löfgren, Persson, and Weibull (2002).

\textsuperscript{3}https://scholar.google.com/
“These informational problems prevents society from achieving the first-best allocation of resources that would be possible in a world where all information would be common knowledge” Laffont and Martinot (2002)

This means that in order to minimize the loss of welfare, stemming from inefficient contracts, or misaligned incentive structures, we need to look at the private incentives, as was proposed in the introductory quote. That quote, however, also alluded to something else, namely the complexity of human behavior interaction. While we of course would like to capture as much as possible of human decision making, it is an undertaking of enormous proportions, and the way economists tackle the problem is by constructing models.

2 Motivating the Theoretical Contributions

The goal of this section is to motivate the use of PA models in this thesis, and their specific contributions, in order to give some context for the modeling choices in Papers I-III, we will in a sense need to go back to basics and describe the foundations of microeconomic modeling.

Positive economics, as opposed to normative, attempts to describe economic phenomena and claims to describe what is, and not what ought; how things are and not how they should be. In order to reach, or at least strive for the objective of living up to the positive claim we need to constantly improve theories and models.

Modeling does, however, come at a cost, and that cost is assumptions. During introductory microeconomics we teach the case of perfect competition, which has some underlying assumptions about the structure of the economy. No economist actually believes that we have perfect competition anywhere, but it is a useful place to start when modeling an economy. It is therefore used as a stepping-stone when giving students some economic intuition of markets, before moving on to the polar opposite market structure monopoly, and later other forms. This incremental approach to knowledge can be argued as necessary for learning, as you need to learn to walk before you can run.

The models used and produced in economics today are built using mathematics, but it hasn’t always been that way.

“Econ developed as a form of philosophy and then added math later, becoming basically a form of mathematical philosophy. […] In other words, econ is now a rogue branch of applied math.” (Smith, 2015)

4 For a longer discussion see Gabbay, Thagard, Woods, and Mäk (2012)
5 Note that when I talk about models and modeling in this section I only refer to theoretical microeconomic modeling, and not macroeconomics or econometrics, although much of the same reasoning would apply.
6 With a few exceptions of e.g. computer simulations using agent based modeling, that may gain traction in the future.
It is often argued, by economists, that mathematics is a tool which is used to describe and analyze problems. It is seen as a language used to express relationships and draw logical conclusions about the outcome of interaction, either through a price system on a market or interaction between actors that e.g. form contracts where one party agrees to perform a task for the other, as is the case in this thesis. These mathematical models often must be constrained by assumptions in order to work, and these assumptions, while often reasonable can be very restrictive. Such was the case of the necessary conditions for the First-Order Approach (FOA) to PA problems, which is used in Paper [III]. The FOA replaces incentive compatibility constraints with first order conditions, but when first formulated by Mirrlees (1976) and Rogerson (1985), concerns were raised as to how realistic the necessary conditions were:

“Since Mirrlees (1976) it has been clear that the so-called first-order approach to solving principal agent problems is generally not valid. In spite of this, it seems that the convenience of the approach has often outweighed any reservations as to its validity.” Jewitt (1988)

Jewitt (1988) then replaced some assumptions with other conditions which extended the use of FOA\(^7\), and more work has been done since to further develop this approach, to make it more general (see e.g. Alvi, 1997; Conlon, 2009; LiCalzi & Spaeter, 2003; Moroni & Swinkels, 2014; Williams, 2008). I will not go into detail on the FOA in this introduction, as my point is only to show that while some models at present appear to be restrictive, it is possible that they can be extended or altered in the future to provide better descriptions of reality. Just like we teach basic concepts to novice students, we too, as researchers, must start somewhere, and walk in order to be able to run.

Now that we have looked at contract theory, and specifically the PA relationship, and the evolution of assumptions in some continuous-time PA models let’s take a step back and examine the way that the actors are most often modeled in this literature, and the rest of neo-classical economics.

In the field of economics there are actors (people, firms, governments etc.) that each maximize what is called an objective function; firms are assumed to maximize profits and individuals maximize utility, or happiness if you will. The individuals that economists model are often what has been called Homo Economicus.\(^8\) This “economic man” is assumed to be rational, and can make calculations as to take the actions needed to maximize his or her utility, no matter the complexity of the calculations needed. This way of modeling has provided us insights, but has also been criticized for over half a century.

“Because of psychological limits of the organism (particularly with respect to computational and predictive ability), actual human rationality-striving can at best be an extremely crude and simplified approxima-

---

\(^7\)This includes removing the assumption of convexity of the distribution function condition.

\(^8\)For an early history see Persky (1995).
tion of the kind of global rationality that is implied, for example, by game-theoretical models.” Simon (1955)

While, on one hand we could argue that the models in microeconomics are simplifications of reality, they also make bold assumptions about our cognitive ability. As the mathematical revolution in economics\(^\text{9}\) started taking off in the 1940’s, there was an unplanned side-effect:

“The IQ of Homo Economicus became bounded only by the IQ of the smartest economic theorist!” Thaler (2000)

Simon (1955), who was quoted above, questioned whether this way of modeling was “a suitable foundation on which to erect a theory”. He asked if it was supposed to be a theory of how actors do, or how they should behave, both because they are often assumed to possess a wealth of information about their surroundings and also because they are assumed to be able to carry out very complex calculations.

Do we act as this person, as modeled in neoclassical economic theory, or is Homo Economicus a mere imaginary creature, as posited in Yamagishi, Li, Takagishi, Matsumoto, and Kiyonari (2014)? Sometimes we behave in the way that current theory predicts, and other times not. Yamagishi et al. (2014) found that only 30 of 446 residents of relatively wealthy Tokyo suburbs met the behavioral definition of Homo Economicus and Gintis (2000) lists several examples from experimental economics and covers results from laboratory studies of games “against nature and ourselves”. These studies are divided into experiments in individual choice behavior and in strategic interaction, and the models in this thesis contain ingredients from both of these categories. PA models, by definition, contain strategic interaction, as each player’s decision depends on the decision of the other, so Papers [I]-[III] all model this type of interaction.

The individual choice behaviors are time inconsistency, choice under uncertainty, loss aversion and status quo bias, of which the first two are part of models in this thesis; Paper [I]-[III] all model choice under uncertainty, as the search process for a qualified worker is stochastic – you never know exactly when you are going to find a match (Paper [II] and [III]) – and under competition you do not know if the best worker you expect will find will be better than your competitor’s best (Paper [I]).

Choice under uncertainty usually involves testing logic and heuristics, which are decision rules that are meant to be approximations of the real optimal decision. As we already know that people do not make decisions exactly as HE, we should try to understand how our decisions differ, and why. This means (i) using experiments to find differences between our models and real-world behavior and (ii) adjust our models accordingly.

Are we predictably irrational, as the title of behavioral economist Dan Ariely’s bestselling book\(^\text{10}\) alludes to? If so, what kinds of mistakes do we make? If we

\(^9\) See Weintraub (2002) for an extensive history.

\(^{10}\) Ariely (2008)
can find these consistencies in mis-judgement, then we have a a fair chance of correcting them. This is how Paper [I] should be seen; as an attempt to describe a possible heuristic that does not work, despite sounding like a good idea. The presented heuristic instead produces overconfidence, consistent with observations in similar laboratory experiments (cf. Sheremeta, 2013). In my model, the use of the heuristic leads to a sub-optimal outcome for the agent, so this contribution could be valuable, because in order to correct bad decisions, we need to understand their mechanics.

Gigerenzer and Brighton (2009) suggests that, in order to make our modeled individuals more realistic, we should replace Homo Economicus with Homo Heuristicus, who would ignore information. The motivation is however not only increasing realism, but also efficiency, since the full calculation of various decisions would require information and time, it may in fact be more efficient to use a simple heuristic. This is an idea that has made it into the mainstream, because in the best-selling book Thinking fast, thinking slow Kahneman (2013) argued that we have two systems of decision making; one that is calculating and one that is instinctive. ¹¹

Paper [III] also introduces a heuristic, but also touches on another part of decision making; decisions over time. Somewhat related to cognitive biases, experiments have shown that we do not act rationally when it comes to inter-temporal choice, i.e. decisions over time; we can make choices today that we regret tomorrow, even though we knew the outcome all along. No matter if it comes to partying and staying out late, or not saving for retirement, we know the consequences, yet we still make decisions that we will regret in the future. Economists have been able to model this behavior using something called hyperbolic discounting, which makes an individual care more about the present than the future (see Laibson, 1997). This has led to a large literature on procrastination (Asheim, 2007; Jain, 2009; O’Donoghue & Rabin, 1999a, 1999b, 2001, 2008; Reuben, Sapienza, & Zingales, 2007; Ylmaz, 2015), but there are also other approaches (e.g. Akerlof, 1991).

A paper entitled *Read this paper later*, where Fischer (2001) develops a model with time-consistent procrastination, where work intensity increases as we get closer to a deadline. In light of this, Paper [III] introduces an agent that uses a heuristic for deciding when to start expending search effort, with the goal of finding a qualified worker to fill a vacancy by a given deadline. In this model there are two types of agents; one which is rational and immediately makes the full calculation required for an optimal choice, and another that continuously uses a simpler heuristic. As it turns out, the heuristic is does not produce optimal behavior from the agent’s point of view. Paper [III] is, like Fischer (2001), in essence also about time-consistent procrastination, but it in addition also includes a heuristic and is set in a principal-agent framework. ¹²

¹¹For textbook introductions to heuristics see Gilovich, Griffin, and Kahneman (2002) and Gigerenzer and Selten (2002).

¹²This also entails some assumptions different those in Fischer (2001).
3 The Temporary Work Agency Sector

According to the 2016 Annual Report of CIETT (2016), the employment and recruitment industry enabled work for 71.9 million people in 2014, making the global penetration rate 1.6%. In 2007 there were 59,400 TWA workers in Sweden (Arrowsmith, 2008), corresponding to 1.3% of the total population (Andersson-Joona & Wadensjö, 2010), and this has been steadily growing since. The latest estimate from Bemanningsföretagen (2015) is that there were 74,400 employed in the TWA sector in 2015.

In Sweden, TWAs are used by both the private and public sector, either for pure recruiting, or recruit-and-rent contracts. It is a relatively young sector, however, as there prior to 1993 was a state monopoly on employment mediation. As a result of Sweden ratifying International Labour Office (ILO) convention no. 34 (the Fee-Charging Employment Agencies Convention) in 1934, TWAs were forbidden by law in the 1935 Employment Mediation Act (SFS1935:113). In addition, Sweden ratified ILO’s 1949 revised convention no. 96. in 1950. Until the deregulation, the law therefore strictly forebode for-profit labor exchanges, which was, in essence, a ban on employment services.

Up until the law was abolished there were a number of public inquiries, none of which resulted in any major changes to the 1935 law, as TWAs were seen as a threat to the Swedish Model, and were feared to erode unionization and the possibilities of executing effective labor policies (Bergström, Håkansson, Isidorsson, & Walter, 2007; Johnson, 2010).

The deregulation started with the enactment of the 1991 Private Mediation Act. Sweden then revoked the ILO convention no. 96 the following year and in 1993 the deregulation was completed through Government bill Prop.1992/93:218. Notably, today the Swedish TWA sector is unique in that there is a collective agreement, which most employees are covered by, and states that TWA employees are guaranteed a wage, even if they are not placed at a client firm. (Bergström et al., 2007; Walter, 2012)

According to Bergström et al. (2007), the reason for the sector’s existence, most often put forth by its representatives, is to satisfy the employers’ need for flexibility (see e.g. CIETT, 2000; Eurociett, 2007), but other reasons cited include the need to adapt to the business cycle, globalization and increased competition. Andersson and Wadensjö (2004) provide additional explanations for the existence of a TWA sector. Firstly, combining smaller tasks at various companies to full time employment. The flip side is that the TWA take the one-time cost of finding

---

13 A trade organization representing the interests of the employment and recruitment industry around the world.
14 For share of TWA workers of total population in various European countries in 2012 see ILO (2015).
15 A Swedish trade organization with over 500 member firms, at time of writing.
16 For graphs of the development over time see Larsson (2014) and Konjunkturinstitutet (2012).
17 Later amended in 1942 (SFS 1942:209) to clarify the definition of private employment services.
18 For a more detailed history see Walter (2012) and Westéus (2014).
a worker, but may be able to spread out the client firm premium over multiple client firms, the consequence being indirect cost sharing among client firms, which may make contracting a TWA an attractive option, especially for client firms that need a worker for only a short period of time.

There is also the assumption that the TWA may be better at finding a qualified worker, and this comes from arguments of economies of scale and specialization. Lastly, as shown in Autor (2001), since the TWA does the screening, client firm’s risk of making a bad hiring decision is reduced, since there is less uncertainty as to the quality of the person it is getting.

Cost reduction, however, is less likely in Sweden compared to other countries, since collective agreements dictate that the wage of TWA workers cannot be lower than that of regular employees.

As stated in Bergström et al. (2007), indirect costs may be reduced if the client firm can avoid missing its production goal, or a delivery deadline, but there may however be an offset of direct costs in the recruit-and-rent case where the client firm does not need to expend resources on the recruitment process, in the form of announcing the vacancy, interviewing and subsequent administration. Gibelman (2005), however, points out that there, on the other hand, is a contracting cost for negotiating with the TWA, which must also be taken into consideration. Gibelman (2005) also states that there is a lack of empirical evidence that there are cost savings to be had.

So far, all arguments except the one about combining smaller tasks into a full time position have been demand side, i.e. from the client firm’s point of view. The argument put forth for the supply side, i.e. the worker’s point of view, is that TWA work could be a stepping stone into regular employment.19 This has been investigated for Sweden by Hveem (2013), who, contrary to popular opinion, found no stepping stone effect. In fact, he found a negative regular employment effect, which slowly faded away over a couple of years.

Paper [IV] in this looks at the transition probability into the temporary work sector, with a special focus on family experience. This is tied in with the literature on occupational inheritance, because it is known that it used to be common for farmer’s children to “inherit” the occupation of their parents, but this effect has not been investigated for temporary work agencies. While the paper does not focus on people whose parents have been long-time employees in the sector, it looks at parents, siblings and partners participation in the sector as possible explanatory variables for why individuals may enter the sector themselves.

Bergström et al. (2007) argues that an increase in demand for TWA workers can be explained by the factors mentioned earlier, when the sector is an established part of the economy. If not, then one needs to consider the social and institutional context in which the sector is growing. They argue that the growth of the sector can then be a result of an increased social acceptance for this type of employment. In the Swedish context, they argue, you need to see this in light of e.g. the norm of permanent posts. Paper [IV] can therefore be seen in the context of family

---

19Industry reports stating this include Bemanningsföretagen (2012) and Silva and Hylander (2012).
connections possibly increasing social acceptance for TWA sector work.

4 Summary of Papers

Paper [I]: Why Beating the Expected Best is a Bad Idea

In Paper [I] I construct a Principal-Agent model, in the form of a game where two temporary work agencies (the agents) compete to provide the best worker for a vacancy at a client firm (the principal).

The agents move simultaneously, meaning that they present their best candidate at the same time. Prior to presenting their respective candidate the agents have to decide how much effort they should expend screening potential hires. They face a trade-off, where additional screening increases their costs, but also their chance of presenting the best worker, thereby winning the contract and securing the prize the principal will pay the winner.

When entering any form of competition all participants should size up their opponents to assess their own chance of winning, so when modeling a game such as this a central concept is that of the contest success function (CSF); it describes the agent’s expectation of its probability of winning. The CSF is thus of vital importance for the agent’s behavior and thus of the actual final outcome. In the paper I present a new CSF called the Fixed Expectation Contest Success Function (FE-CSF); a feasible heuristic which consists of the probability of beating the opponent’s expected best candidate.

There is strategic interaction in this game, as one agent’s decision will depend on the decision of the other. When one of the agents decides how much effort to expend it calculates how good it thinks the best worker its competitor will present, and then decides how effort to expend in order to find someone better.

While this may seem like a reasonable way of thinking, it is not rational, as Tullock (1980) showed that the way you judge your chance of winning is by looking at the effort of the opponent, and then comparing it to yours. I show that if instead of focusing on effort, i.e. the opponents sweat and blood, the agent focuses on the opponent’s best expected worker, then the agent will become overconfident. Then, looking at how this overconfidence affects behavior, i.e. how many workers are screened, I show that the level of effort in symmetric equilibrium will increase, compared to rational Tullock agents.

This result is interesting, as overconfidence is a well-documented phenomenon in economic game experiments (see Sheremeta, 2013, for a review), and the biggest selling point of this paper is that the overconfidence derived (and its size) comes not from an arbitrary term added to the agent’s rational expectations, but from the functional form resulting from the aforementioned heuristic.
This paper looks at the incentive structure in the temporary work sector by modeling a Principal-Agent relationship, when a risk-neutral CF (the principal) contracts one or more risk-neutral TWAs (the agents) to fill a vacancy. Unlike Paper [I], where the agents presented their best worker simultaneously, this model is set in discrete time and search is carried out sequentially. The CF requires a worker with some minimum level of productivity, so in the pool of workers to be screened there are some who are qualified for the position, while others are not. The game ends when a TWA presents a sufficient worker to the CF.

Two types of contracts are analyzed; pure recruiting, where the TWA is hired to find a match which will be hired by the CF, and recruit-and-rent, where the CF will rent the worker from the TWA for some amount of time.

Looking at the pure recruitment contract, we analyze three payment schemes. The first being continuous payment until a match has been found, or the contract expires at some predetermined point in time. As we assume hidden action, there is no incentive for the TWA to exert any effort at all with a contract such as this, as it will be paid no matter what and screening workers to try to find a match will yield no additional revenue.

The second payment scheme we look at is payment on delivery, which ensures that the agent starts searching immediately upon signing the contract. Here we conclude that the TWA will never attempt to find an alternative candidate after having found the first match.

The last payment scheme, which we only briefly look at, is payment at a pre-specified point in time, conditioned on delivery. Our conclusion is that if the TWA cannot, because of liquidity constraints, or does not want to search for the total amount of time until the pre-specified point when it will be rewarded, then it will defer search, as any given search effort will have a lower present value cost the later it is expended. This intuition is only sketched out, and will be the topic of Paper [III].

The second type of contract we analyze is the recruit-and-rent contracts. Here we first look at the incentives when contracting a single TWA, and assume that the better the worker, the higher the chance of him/her transitioning into regular employment at the CF. This means that if the TWA finds a good enough worker, it will may make the decision to spend resources searching for a worse worker, which can be assumed to be rented for a longer period of time.

The last thing we do in the paper is then to look at how incentives are affected by increased competition among TWAs. While a single TWA would ideally like to find a worker as close as possible to the minimum productivity level stipulated by the principal, when competition increases, the TWAs probability of winning with a lower (but still sufficient) productivity worker will decline. We show that the TWAs’ preferred worker gets progressively better as competition increases.
Paper [III]: Think About the Future and Wait

Paper [III] examines the waiting behavior that can occur if a TWA is contracted to find a worker for a specific time far in the future; the TWA will postpone effort, i.e. wait to start searching for a qualified worker.

In the paper I present two types of TWAs; one that is rational and plans ahead, and another that does not plan ahead at all, but instead only looks at the immediate future. I find that the one that only looks at the immediate future starts exerting effort earlier than the planner.

I then proceed to characterize the optimal contract between the CF and the TWA, because the CF will have to find the optimal amount to promise as prize money if the TWA successfully supplies a sufficient worker; if the CF pays to little, then the TWA will not want to exert any effort at all, and if the CF pays too much it will not make a profit itself. This contract is examined first under the assumption that the CF can monitor the TWA, and then under the assumption that it cannot, where the latter is more realistic.

Under the strong assumption of perfect monitoring, the CF can not only decide the prize if the TWA is successful, but also dictate when the agent should start searching, as it can see if the TWA exerts the promised effort. This makes it possible to achieve an efficient contract.

If there is hidden action on the other hand, meaning that the CF cannot monitor the TWA, then writing an efficient contract with a rational TWA is not possible. Even though the CF could pay exactly the amount it would like the TWA to use for search effort, the TWA will find it profit maximizing to wait a little longer, as not to have to expect to spend all of its expected revenue on search.

I then show that for the principal to want to delegate search to a rational TWA, the agent has to be better than the CF, by some factor, as it has to make up in efficiency what the CF loses in the TWAs moral hazard – when the TWA waits longer than the CF would like it to. Lastly I show that it is profit maximizing for the CF to contract one TWA and give it a deadline earlier than when the CF would need the worker, and then replace that TWA with a competitor if the first one has not succeeded by that earlier deadline.

Paper [IV]: Young Adults in the Swedish Temporary Agency Sector: Implications of Family Experience

A persons first experience of working life is not the individuals actual first job, but rather the perception conveyed by his or her family and other reference groups. This paper ties in to the literature on occupational inheritance, as well as the literature on changing social norms, by investigating the implications of family experience on young adults’ relative probability of transitioning into the Temporary Agency Sector. The reason for looking at young people is that the sector has only been deregulated since 1993, as mentioned above, making parental effects hard to motivate looking at for older cohorts.
We specifically focus on the effect of a family member having been in the temporary agency sector; on one hand one could imagine it would make a person less inclined to enter the sector, as our review of the literature shows some negative effects of TWA employment; a wage penalty, worse working conditions, significantly higher risk of feeling depressed and workers in the sector experiencing the lowest degree of autonomy and overall job satisfaction compared to other types of employees (Andersson-Joona & Wadensjö, 2012; Fabiano, Curro, Reverberi, & Pastorino, 2008; Håkansson, Isidorsson, & Strauss-Raats, 2013; Tijdens, van Klaveren, Houwing, van der Meer, & van Essen, 2006). Despite this, the effect was the opposite; Using Swedish register data on young adults (aged 18-34), and controlling for personal characteristics, we find that individuals with family members or partners with work experience from the temporary agency sector are highly over-represented in the sector. Other variables we control for is gender, age group, highest level of education attained, if the person is a second-generation immigrant, lives in a metropolitan municipality, is a student or has children. The peer-groups previous experience is, in fact, found to be among the most influential variables determining the relative probability that an individual will work in the temporary agency sector.

We also look at two sub-groups separately; gainfully employed and students, as we believe there may be important differences between the groups. The results for gainfully employed workers and for the student group do indeed show that there are some other important differences that have not been captured by previous studies. For instance, there are relatively many temporary agency workers in some of the lower age cohorts in the student sample, whereas the gainfully employed show an almost linear decay in the relative probability of being employed in the agency sector as they grow older. A noteworthy result that is very similar in all samples, but quite different to the findings in 1999 by Joona and Wadensjö (2008), is the relatively high education level among the younger cohorts of the temporary agency sector (cf. also Andersson-Joona & Wadensjö, 2012; Petersson, 2013).

The overall results of this study further establish that individuals with an immigrant background are still over-represented in the sector, but also that, in the younger cohorts, there is a predominance of men employed.
References


Why Beating the Expected Best is a Bad Idea

Tomas Raattamaa

Department of Economics
Umeå School of Business and Economics
Umeå University

Abstract

Overconfidence is a well-documented phenomenon in economic game experiments. This paper modifies the standard Tullock contest by introducing a feasible heuristic which consists of the probability of beating the expected best outcome of the opponent. This new Fixed Expectation Contest Success Function (FE-CSF) is compared to Tullock’s CSF and shown to lead to overconfidence and over-expenditure of resources, from the contestant’s point of view.

Acknowledgments: The author would like to thank Magnus Wikström, Thomas Aronsson, Tomas Sjögren, Chandra Kiran, as well as the seminar participants at the Department of Economics for their comments and suggestions.
1 Introduction

“I’m above average.”
(Most people)

While the above quote certainly may seem humorous, it is never the less not always a poor description of our perception. On the smaller scale we would refer to it as overconfidence, and on a larger scale this has been called the above average effect, or the better than average effect. Some also call it The Lake Wobegon effect, in honor of author Garrison Keillor’s fictitious town "where all the women are strong, all the men are good-looking, and all the children are above average" (quoted in Maxwell & Lopus, 1994).

There is more than anecdotal evidence for the existence of this effect, as it has been shown numerous times in empirical studies (see Dunning, Heath, & Suls, 2004, for some examples) and while overconfidence has been demonstrated in various settings, such as assessing ones driving ability (see Svensson, 1981), this paper will focus on overconfidence in an economic contest.

As noted in L. C. Corchón (2007) the rent seeking studied by Tullock (1980) and Krueger (1974) and lobbying studied by Becker (1983) are the seminal works in the literature on contests. The type of contest we will be looking at in this paper has thus been explored in the literature since the 1980’s and the basic explanation of how the game is setup is as follows; two (or more) contestants compete for a prize by expending costly non-recoverable, non-observable effort, which means that there will be strategic interaction among the players, as one player’s probability of winning is affected by its competitor’s effort level, and vice versa. Often used examples of this type of contest are R&D contests, patent races, lobbying, litigation etc., but the one we will use throughout this paper is a contest where a Client Firm (CF) presents two or more Temporary Work Agencies (TWAs) with a contest where the TWA that supplies the best, i.e. most productive, worker wins. This is a suitable model for the principal-agent relationship between a CF and TWA(s) as it is not uncommon for several TWAs to attempt to fill the same vacancy. Other aspects of the aforementioned principal-agent relationship have been examined in a continuous time model by Raattamaa (2016), demonstrating optimal waiting, and in a discrete time model by Westéus and Raattamaa (2014) demonstrating that the agents may not have incentives to provide the best worker, even in the existence of competition. This paper however will focus on a contest in the form of a one-shot game, reducing the problem not to a decision over time, but only in effort.

When entering a competition all participants should size up their opponents to assess their own chance of winning. How that conclusion is reached is up for debate, but when modeling a game such as this a central concept is that of the contest success function (CSF); it describes the agent’s expectation of the outcome of the game, i.e. its probability of winning. The CSF is thus of vital importance for the agent’s behavior and thus of the actual final outcome. The two main CSFs are those of Tullock (1980) and Hirshleifer (1989) (both axiomatized in Skaperdas
(1996) and the latter also in Cubel and Sanchez-Pages (2015)). The difference-form CSF in Hirshleifer (1989) relies on the relative difference between the contestants’ efforts, and in Tullock (1980) the CSF was defined as the given contestant’s effort divided by the total effort expended by all participants. In a two-player game the Tullock CSF for player $i$ can be written as

$$\rho \left( q_i, q_{-i} \right) = \frac{q_i}{q_i + q_{-i}} \quad (1)$$

where $q_{x \in \{i,-i\}}$ denotes the effort of the respective agent. This CSF has become known as the Tullock ratio and the game as a Tullock contest. The Tullock ratio in its most basic form, as described above, is unbiased, in the sense that the estimated probability is equal to the true probability, and the game has been studied and developed extensively since its inception; it has been extended to include e.g. contest uncertainty (Grossmann, 2014), the role of information (L. Corchón & Dahm, 2010; Einy, Haimanko, Moreno, Sela, & Shitovitz, 2013; Epstein & Mealem, 2013; Polishchuk & Tonis, 2013; Serena, 2014), risk aversion (Cornes & Hartley, 2012), multiple equilibria (Chowdhury & Sheremeta, 2011b) and spillovers (Chowdhury & Sheremeta, 2011a) to name a few. This is quite an active area of research with a number of good literature reviews; see eg. Nitzan (1994), Konrad (2007), Long (2013) and Jia, Skaperdas, and Vaidya (2013). Recently Chowdhury and Sheremeta (2015) introduced equivalence among some Tullock-type contests and for more axiomatizations and other contest success functions see Chakravarty and Maharaj (2014) and Lorentz (2014).

The basic setup though is that any agents’ profit function can be described as $\pi = \rho V - cq$, where $V$ is the value of the contract (i.e. the prize), if won, $c$ is the unit cost of effort and $q$ is the amount of effort expended. $\rho$ is the probability of winning, which depends on all participants’ efforts.

The standard theory cannot explain overconfidence, but it can easily be added to the Tullock ratio; two straight-forward ways are adding a positive term to the CSF, or multiplying the CSF with a term larger than one. The question then becomes how to determine its functional form (and size.) These, and similar methods are however somewhat ad-hoc and as rational behavior has already been demonstrated by Tullock (1980), the goal of this paper is not to derive an optimal behavioral rule, but instead show a plausible heuristic. This heuristic will then be shown to produce overconfidence.

For the Tullock ratio the rationale is that when optimizing the agents look at the (level of) effort of their opponent(s), but one could argue that an equally probable optimization rule or heuristic would be to fix your target, i.e. decide what you think your opponent will present and use that as a metric for your effort expenditure. It is the heuristic equivalent of having a mental picture of the outcome, rather than the opponent’s blood, sweat and tears.

---


2Another way would be to underestimate costs, as was done in Ludwig, Wichardt, and Wickhorst (2011)
This paper departs from the standard Tullock ratio (and its variations) and introduces a new CSF. Our contribution is that we will build a (theoretical) model where agents aim to beat their opponent based on how good they expect their opponent’s best worker to be. While this seems intuitive, Tullock (1980) demonstrated that it is actually rational to set your level of effort in relation to your opponent’s effort. We will show the difference between these two approaches, and also show that fixing your expectation of your opponent’s best worker will lead to you being overconfident in your own probability of winning. This paper’s biggest selling point is that the overconfidence derived (and its size) will come not from an arbitrary term added to the agent’s rational expectations, but from the functional form resulting from the aforementioned heuristic.

Modeling overconfidence, and also more specifically overconfidence in contests is interesting because it is a behavior we humans seem to exhibit. There is a literature on contest experiments and of those summarized in Sheremeta (2013) statistically significant overbidding, meaning effort higher than the risk-neutral Nash equilibrium prediction, was found in 28 of 30 papers published between 1989 and 2013 and the median overbidding rate was 72%. While deriving a novel approach to modeling overconfidence this paper may also offer a piece of the puzzle of human behavior.

We will now go on to presenting the model and results, where we first introduce our agents and then compare them to the rational Tullock agents to show that they will be overconfident. Lastly there will be some concluding remarks.

## 2 Model and Results

The contest will be modeled as a simultaneous game where the principal offers a set $Y$ of $n > 1$ risk-neutral agents a contract promising a payment $V$ to the agent that supplies the best worker for a given vacant position. The agents simultaneously choose how many workers to screen, i.e. any agent $i \in Y$ exerts effort $q_i \in [0, \infty)$ at a unit cost of $c_i$, before sending the best one to the principal for evaluation against their competitor’s best candidate. An agent’s level of effort (here the number of screened workers) is thus chosen based on the value of the contract in question, the unit cost of screening and the agent’s expectation of the competitor’s effort.

Let us now focus on the two-player game, where agent $i$ plays against agent $-i$. The expected profit function of firm $i$, given the effort of firm $-i$, can be formulated as

$$\pi_i = \begin{cases} V - c_i q_i & \text{with probability } \rho(q_i, q_i) \\ -c_i q_i & \text{with probability } 1 - \rho(q_i, q_i) \end{cases}$$

---

3 For another review see Dechenaux, Kovenock, and Sheremeta (2015)
where the (true) probability of winning, \( \rho \), is the aforementioned Tullock ratio from Equation (1).

While the Tullock contest is based on continuous effort, it has been shown to be strategically equivalent to innovation tournaments such as this, where participants draw from a distribution in order to increase their chance of winning (see Baye & Hoppe, 2003). This, so far, is the standard formulation of the contest, and we will now make and explain some assumptions of our setup, before replacing the Tullock ratio with our own CSF.

In our model the interpretation of the continuous effort \( q_i \) and a unit search cost per worker is that effort represents the average number of qualified workers the agent will find. We make a conscious decision not to pay much attention to what happens if effort is less than one, as we want to focus on the mechanics of the game, rather than the intricacies of the matching process.\(^4\)

As stated above, the cost of finding a worker is fixed. This constant unit cost is a common assumption in contests and reasonable in our context if the pool of qualified workers is large enough, and feasible if TWAs are efficient enough not to get economies of scale in the screening process.

We assume that the potential hires are uniformly distributed in terms of productivity, \( x \sim U(0, 1) \), and that any draw will result in a sufficient worker that the principal may accept. The interpretation of \( x \) should thus not be absolute productivity, but productivity above some by the principal required minimum. Another, more general interpretation of \( x \) is that it is the rank of the worker, since this is all that the agents care about; an agent gets paid as long as its supplied worker is better (i.e. has a higher rank) than the competitor’s. That means that we would not need to make any assumption about the distribution of productivity among the workers.

Proceeding with constructing our model, the agents are assumed to have what we call fixed outcome expectations; instead of comparing efforts as in the Tullock game, they will fix their expectation of how well their opponent will do, given the opponent’s expected effort, and then decide how much effort to expend trying to beat that.

This means that we will need to use the order statistics of the uniform distribution. The value of the opponent’s best candidate, given that the competitor \(-i\) evaluates \( q_{-i} \) workers is

\[
\bar{x}_{-i}(q_{-i}) = \max \{ x_1, x_2, ..., x_{q_{-i}} \} \tag{3}
\]

The PDF of the \( n \)th order statistic is from the family of beta distributions, and as we are working with a continuous uniform distribution on the unit interval we know that the expected largest ordered statistic will be

\(^4\)A more realistic assumption would be that if effort is less than what is required for a sufficient worker to be found, then the probability of success should be equal to \( \Pr(x_j | q) = \min(q, 1) \). This would mean that putting in half the effort required for finding a sufficient worker would give the agent a 50% chance of succeeding in finding someone. Adding such an assumption for the special case of \( q < 1 \) only serves to over-complicate the model, which is why it is only mentioned here and not explicitly modeled.
which is the best worker agent $-i$ is expected to obtain after exerting $q_{-i}$ amount of effort. Equation (4) can also be interpreted as the probability of agent $i$ not finding a better worker than $-i$ after screening one worker. This probability raised to the power of $q_i$ is then the probability of agent $i$ not finding a better worker (i.e. losing) after $q_i$ attempts. The expected probability of agent $i$ winning (i.e. not losing) is therefore

$$\hat{\rho} (q_i, q_{-i}) = 1 - E [\bar{x}_{-i} (q_{-i})]^{q_i}$$

which we will refer to as the Fixed Expectation Contest Success Function (FE-CSF) from now on. Replacing the Tullock ratio ($\rho$) in Equation (2) with the FE-CSF ($\hat{\rho}$) we obtain the profit function of the Fixed Expectation (FE) agent:

$$\pi_i^e = \hat{\rho} (q_i, q_{-i}) V - c_i q_i =$$

$$= \left[1 - \left(\frac{q_{-i}}{q_{-i} + 1}\right)^{q_i}\right] V - c_i q_i$$

Now that we have formulated the agents’ profit function we will firstly show that they will be overconfident, then look at the implication for their best response function, i.e. what effect this overconfidence has on behavior, and lastly discuss the implication for the contest holder, the principal.

### 2.1 Overconfidence

Overconfidence in our economic contest is overestimating your own probability of winning against your opponent, so we define overconfidence of agent $i$ as

$$\Theta (q_i, q_{-i}) \equiv \hat{\rho} (q_i, q_{-i}) - \rho (q_i, q_{-i})$$

which translates into the percentage point difference in expected versus true probability of winning the contest.

We will start by looking at overconfidence in the symmetric case, so we therefore assume homogeneous costs and drop the effort variable subscript in order to simplify notation.

**Proposition 1.** In a symmetric two-player game both FE agents will overestimate their probability of winning.

**Proof.** Assuming symmetry implies homogeneous costs which lead to identical efforts in equilibrium, $q = q_i = q_{-i}$, so the CSFs can be written as
\[ \rho(q, q) = \frac{q}{q + q} = \frac{1}{2} \]  
\[
\hat{\rho}(q, q) = 1 - \left( \frac{q}{1+q} \right)^q
\]  
(8)

Now let the starting point be an effort by both firms of \( q = 1 \), so that
\[
\rho(1, 1) = \frac{1}{1+1} = \frac{1}{2}
\]  
\[
\hat{\rho}(1, 1) = 1 - \left( \frac{1}{1+1} \right)^1 = \frac{1}{2}
\]  
(9)

because if both agents exert the same amount of effort, then the true probability of winning is \( \frac{1}{2} \). The true probability of winning in symmetric equilibrium will always be \( \frac{1}{2} \) as is seen in Equation (8) and while the partial derivative of the Tullock CSF w.r.t. effort in is always zero, the partial derivative of the FE-CSF will be positive for any \( q > 1 \), demonstrated below.

\[
\frac{\partial \hat{\rho}(q, q)}{\partial q} = \frac{\partial}{\partial q} \left[ 1 - \left( \frac{q}{q+1} \right)^q \right] = -\left( \frac{q}{q+1} \right)^q \left\{ (q+1) \left[ \frac{1}{q+1} - \frac{q}{(q+1)^2} \right] + \log \left( \frac{q}{q+1} \right) \right\} = -\left( \frac{q}{q+1} \right)^q \left[ \frac{1}{q+1} + \log \left( \frac{q}{q+1} \right) \right] > 0 \text{ for } q > 1
\]  
(10)

will be positive, as the second term above is negative, but monotonically increasing, and approaching zero from below.\(^5\) This proves that the FE agents will overestimate their probability of winning in symmetric equilibrium, and thus are overconfident.

A simpler way of proving this would be to simply show that
\[
\Theta(q, q) = \frac{1}{2} - \left( \frac{q}{1+q} \right)^q > 0 \text{ for } q > 1
\]  
(11)

The above partial derivative showed that as the agents’ efforts increase, so will their overconfidence. This overconfidence will however not be unbounded, but approach
\[
\lim_{q \to \infty} \left[ \frac{1}{q+1} + \log \left( \frac{q}{q+1} \right) \right] = 0^-
\]  
\(^5\)
\[ \lim_{q \to \infty} \Theta(q,q) = \frac{e - 2}{2e} \approx 0.13 > 0 \]  

This is graphically illustrated in Figure 1, where we see that \( \Theta(q,q) \) approaches the upper limit as symmetric effort increases.

While the symmetry case is useful in comparing to earlier literature, as symmetry is a common assumption, this is somewhat restrictive, so to provide some insights into the CSFs and their difference they are plotted in Figure 2. The only negative overconfidence in Figure 2b is found for \( q_i < 1 \) and that region should be ignored, as we do not model this case properly, which we mentioned earlier.

Removing the symmetry restriction yields an overconfidence maximum for agent \( i \) of \( \Theta(q_i, q_{-i}) \approx 0.20 \) percentage points, for very high levels of effort.\(^6\)

While the largest overconfidence is not found in symmetric equilibrium, we can see that it does have a maximum for a given level of one of the agents’ effort. This

\(^6\)This is found numerically at \( \{q_i, q_{-i}\} \approx \{8.0, 3.3\} \cdot 10^8 \), as tolerance does not allow for maximization at higher levels because of machine precision.
can be explained by the CSFs being bounded by \((0, 1)\) for non-negative levels of effort

\[
\lim_{q_i \to \{\infty, 0\}} \hat{\rho} (q_i, q_{-i}) = \lim_{q_i \to \{\infty, 0\}} \rho (q_i, q_{-i}) = \{1, 0\} \quad (15)
\]

\[
\lim_{q_{-i} \to \{\infty, 0\}} \hat{\rho} (q_i, q_{-i}) = \lim_{q_{-i} \to \{\infty, 0\}} \rho (q_i, q_{-i}) = \{0, 1\} \quad (16)
\]

and therefore overconfidence will be eliminated in the limits

\[
\lim_{q_i \to \{\infty, 0\}} \Theta (q_i, q_{-i}) = \{0, 0\} \quad (17)
\]

\[
\lim_{q_{-i} \to \{\infty, 0\}} \Theta (q_i, q_{-i}) = \{0, 0\} \quad (18)
\]

The intuition is that if you put in a lot more effort than your competitor, then both your estimated and the true probability of winning approach 1, eliminating the possibility of overconfidence, but if you on the other hand put in a lot less effort than your competitor, then your probability of winning is so small that your percentage point overestimation disappears. Similarly to how overconfidence disappears in the above limits, competition will have the same effect. For sake of simplifying notation, we go back to the symmetry assumption.\(^7\)

**Proposition 2.** Increased (perfect) competition eliminates this overconfidence.

**Proof.** Generalizing to \(n\) identical competitors we define overconfidence as \(\Theta (q, n) \equiv \hat{\rho} (q, n) - \rho (q, n)\) and the CSFs become

\[
\rho (q, n) = \frac{1}{n} \quad (19)
\]

\[
\hat{\rho} (q, n) = 1 - \left[ \frac{(n - 1) q}{(n - 1) q + 1} \right]^q \quad (20)
\]

both of which will go towards zero in the limit thereby making overconfidence disappear:

\[
\lim_{n \to \infty} \Theta (q, n) = 0 \quad (21)
\]

\(\square\)

This is the standard result in the literature, and not at all surprising. Under perfect competition there will be no overconfidence, i.e. difference between the expectations of firms and the outcomes.

\(^{7}\)The result of Proposition 2 does not, however, rely on the symmetry assumption.
As \( \frac{\partial \Theta(q,n)}{\partial q} = \frac{\partial \hat{r}(q,n)}{\partial q} > 0 \) for \( n \geq 2 \) and \( q \geq 1 \) we can show that the largest overconfidence will be

\[
\Theta(q,n) \equiv \lim_{q \to \infty} \Theta(q,n) = 1 - e^{\frac{1}{1-n}} - \frac{1}{n}
\]  

for which the case \( n = 2 \) corresponds to Equation (14).

We have now shown that the FE agents will be overconfident, but how does this affect their behavior? To investigate that we need to look at the FE agent’s best response function.

### 2.2 Best Response Function and Rent-Seeking

A best response function is the result of optimization and describes what to do, given your competitor’s action. Temporarily treating \( q_i \) as a continuous variable lets us maximize the profit function \( \pi^i \) given in Equation (6) with respect to effort \( q_i \), which gives us a first order condition we can solve for \( q_i \) to obtain the best response of firm \( i \), given the effort of firm \( -i \).

\[
q_i^{BR} (q_{-i}) = \log_s \left[ -\frac{c_i}{V} \log_s (10) \right] = \log_s \left[ \frac{c_i}{V} \right] + \log_s \left[ -\log_s (10) \right]
\]  

(23)

where the base \( s = s(q_{-i}) = \frac{q_{-i}}{q_{-i}+1} \).

We are unfortunately unable to obtain an expression for the equilibrium quantities, even by assuming symmetry, so let us instead turn our attention to the differences between this and the somewhat more well-behaved Tullock game. As we are interested in the difference in behavior stemming from the differences in the CSFs we will assume symmetric firms, i.e. \( c_i = c_{-i} \), and therefore drop the subscript. In order to examine the FE agent we will define a (general) best response function \( q^{BR} \), and as is standard in the literature the best response function can be iterated a number of times in order to approach the true equilibrium value \( q^{BR*} \).

**Definition 3.** Let \( q^n_{BR} (q_{init}) \overset{def}{=} (q_{BR} \circ q^{n-1}_{BR}) (q_{init}) \) be the \( n \)-th iteration of the best-response function \( q_{BR} \), in which \( \circ \) is a function composition.

**Remark 4.** As the best response function is (quasi)concave (implying single-crossing) the best response functions will converge, so it follows from the definition above that \( q^n_{BR} = q^{n-1}_{BR} \) as \( n \to \infty \), and thus \( \lim_{n \to \infty} q^n_{BR} (q_{init}) = q^{BR*} \).

To find the Nash equilibrium numerically we will use the algorithm defined in Appendix B.

---

8See Appendix A.
Setting the unit search cost equal to unity the standard result in the literature is that two Tullock agents competing against each other will each exert $\frac{1}{4}V$ units of effort in equilibrium, leading them to only use half of the value of the contract for search (see eg. Chowdhury & Sheremeta, 2011a; Sheremeta, 2013; Tullock, 1980). Figure 3 shows the total (equilibrium) expenditure when Tullock agents compete and when FE agents compete, given a normalized unit search cost. The contract value is the 45-degree line, and neither FE agents competing or Tullock agents competing will use the full value of the contract for search. The relationship between the contract value and Tullock agents’ expenditure is linear, but there also seems to be a linear relationship between contract value and the FE agents aggregate expenditure. Estimating the slope of the line shows it to be close to $\frac{6}{8}$ (for sufficiently large contracts), indicating that the FE agents will each spend about $\frac{3}{8}V$ on search.

Since we have assumed a normalized unit cost the total expenditure is equal to total effort. This means that we can safely generalize the graphical representation of the FE agents’ best response functions as in Figure 4. There we see, as we estimated, that FE agents expend more effort than Tullock agents in equilibrium. The diagonal line represents the points at which the aggregate effort sums to $V$, and as the FE equilibrium is under this line aggregate effort does not exceed the value of the contract.

---

9Actual numbers used to generate the graph are $V = 10$ and $c = 1$.  

10
Nevertheless, it would be possible for two agents to participate in a contest where they both have lower capacity constraints above half of the value of the contract. Plotting the iso-profit curves in Figure 5 we see that sometimes the FE agents irrationally expect to make a profit even if aggregate expenditure is above the value of the contract and their symmetric expected break-even point is indicated in the contour plot.
Additionally Figure 6 shows the size of the FE agents expenditure share in their expected break-even, and we see that it is increasing in contract value \( (V) \). We start plotting at \( V = 2c \) as the Tullock agent will not participate if the prize is any smaller and we can see that for any \( V \) above \( 2c \) the break-even is overestimated, as the true (Tullock) break-even at \( \frac{1}{2} \) is indicated by the dashed line.

![Figure 6: Expected break-even expenditure share](image)

Moving away from the normalized search cost we use Figure 7, where the y-axis shows the total expenditure share \( \left( \frac{2cq_{\text{BR}}}{V} \right) \) and the x-axis the unit search cost share \( \left( \frac{c}{V} \right) \). There we can see that over-investment in search decreases as the unit search cost approaches the contract value, independent of the value of the contract. Why is this? Higher unit costs lead to the agents being able to screen fewer workers, which in turn leads to lower overconfidence. For the Tullock agents the unit search cost does not affect the total aggregate expenditure share, as they always only spend one fourth of the contract value each, but the FE agents on the other hand will change their behavior somewhat as the ratio between unit cost and prize changes. Along the x-axis we increase the unit search cost, but because we divide it by the contract value we can see that the expenditure share (of the contract value) is the same for any unit cost share (of the contract value). This has been tested for various levels of \( V \) and the results always stay the same. Just as in the previous figure the Tullock agent is not going to want to participate if the unit search cost exceeds \( \frac{1}{2}V \).

3 Conclusions

This paper presented a heuristic that could be used by agents engaged in a one-shot simultaneous contest and the heuristic turned out to lead to overconfidence. Our example was a CF which holds a competition where two (or more) TWAs compete to provide the best worker. If the TWAs use the expected value of the
best worker to be presented by their competitor as a basis on which to form the expected probability of winning, then they will become overconfident and thus both believe to be above average in terms of their ability to win the contest. This means that the FE agents’ expected probability of winning sums to more than unity, which is common when modeling deviations from the Tullock ratio. This overconfidence in turn was shown to lead to increased effort on the agent’s part. As symmetric Tullock agents together spend half of the contract value, i.e. \( \frac{1}{2}V \), on search and the FE agents introduced in this paper spend \( \frac{3}{8}V \), some of the FE agents’ rents dissipate. The FE agents will still make a profit (although not as large as they would expect), but their overconfidence will not (on average) drive them out of business. While this over-expenditure on effort somewhat diminishes the agent’s profits, it would be beneficial for the principal in a contest where the it gets utility from the outcome, e.g. as in our recruiting example. As the principal wants the agents to exert as much effort as possible, then FE agents that are overconfident will be preferred to agents not using this heuristic, because the overconfidence we have found will lead to the CF on average being presented with better workers.

This resulting overconfidence may very well explain some of the results in the experimental literature on contests, where overbidding is frequently observed, but it is too early to draw any conclusions, as this hypothesis needs testing.

References


Epstein, G. S., & Mealem, Y. (2013, sep). Who gains from information asymmetry?


Appendix

A The Two-Player Best Response Function

In order to derive the best response function we take the partial derivative of the expected profit function in Equation (6)

$$\frac{\partial \pi^e}{\partial q_i} = -c_i - V \left( \frac{q_{-i}}{q_{-i} + 1} \right)^{q_i} \log \left( \frac{q_{-i}}{q_{-i} + 1} \right)$$

which we set equal to zero and solve for $q_i$ to obtain the best response function

$$q_i^{BR} (q_{-i}) = \frac{\log \left( -\frac{c_i}{V \log \left( \frac{q_{-i}}{q_{-i} + 1} \right)} \right)}{\log \left( s \right)}$$

$$= \frac{\log \left( -\frac{c_i}{V \log (s)} \right)}{\log (s)}$$

$$= \log_s \left( -\frac{c_i}{V \log (s)} \right)$$

$$= \log_s \left( -\frac{c_i}{V} \log (s) \right)$$

$$= \log_s \left( -\frac{c_i}{V} \log (s) \right)$$

$$= \log_s \left( -\frac{c_i}{V} \log (s) \right)$$

$$= \log_s \left( -\frac{c_i}{V} \log (s) \right)$$

where $s = s (q_{-i}) = \frac{q_{-i}}{q_{-i} + 1}$

B Numerically Finding Symmetric Equilibrium

Because we cannot solve for an explicit best response function we will need to iterate to find an approximate solution. Implementing the algorithm is done in the following steps.

Algorithm 5. The iterated function in Definition 3 is implemented as follows;

1. Set the marginal cost and prize to define the best response function from Equation 23.

2. Set the threshold difference at which the iterative process will be terminated.
3. Because of symmetry, iterate until the difference between two immediately following iterations is less than a pre-specified threshold, \( |q_{BR}^n - q_{BR}^{n-1}| < \text{limit} \).

4. Start the iterative process, where the first reaction will be to \( q_{\text{init}} = \frac{V}{4c} \).

Any initial value can be used, but I have chosen \( q_{\text{init}} = \frac{V}{4c} \) since this is the Tullock equilibrium value. It is closer to the final equilibrium than e.g. 0 or \( V \). Using this speeds up the time the iterative process needs to find equilibrium.
The Misaligned Incentives of Temporary Work Agencies and their Client Firms

Morgan Westéus  Tomas Raattamaa
Department of Economics  Department of Economics
USBE  USBE
Umeå university  Umeå university

Abstract
This paper adds to the theoretical literature on the incentives of Temporary Work Agencies (TWAs). Using a principal-agent model with hidden action to analyse two main types of contracts between a TWA and a Client Firm (CF), the TWA is shown to potentially act against the best interest of the CF when helping to fill a vacant position. The results also suggest that the adverse effect of the incentive misalignment is larger when workers are leased rather than hired by the CF. However, this effect could potentially be offset by introducing a sufficient level of competition among TWAs.

Acknowledgements: The authors wish to thank Thomas Aronsson, Magnus Wikström, Karl-Gustaf Löfgren, Christian Lundström (USBE) and Erik Geijer (SLU/CERE), discussants Tomas Sjögren, Stefanie Heidrich (both USBE), and Mattias Lundbäck (RATIO), as well as the seminar participants at the Department of Economics and at the RATIO institute for a gainful discussion and their insightful comments. The paper has also benefited from comments given by Ann-Sofie Kolm (SU) and the other participants at the Swedish National Conference in Economics 2014.
1 Introduction

The number of workers employed through a temporary work agency (TWA), either as a consultant employed by the TWA, or screened by the TWA before being hired by the client firm (CF), has risen substantially in the last decades (Andersson-Joona & Wadensjö, 2010; Forde & Slater, 2005). While it might seem as though the interests of the CF and TWA are aligned, i.e. one pays to fill a vacant position while the other gets paid to fill said position, we show that this is not necessarily the case (cf. Gibelman, 2005).

By outlining a principal-agent model with hidden action we are able to shed light on some important consequences of the differing incentive structures between the CF (principal) and the TWA (agent); among other things, the CF wants to lease the best possible worker for a given position, while the TWA may want to provide the least productive worker possible who is still sufficiently good at his/her job.

The TWA is usually assumed to have some advantage(s) in the recruitment process through which it is argued to be able to supply a worker more quickly than the CF would be able to, and/or hedge certain liabilities of the CF towards the contingent TWA worker (Autor, 2001, 2003; Baumann, Mechel, & Stähler, 2011; ECORYS-NEI, 2002; Houseman, 2001; Mitlacher, 2007; Neugart, 2005; Westéus, 2014). The CF’s primary use of TWAs is thereby to either delegate search, i.e. identify prospective candidates for a vacant position and assess their productivity, and/or be able to (quickly) lease a worker to fill a vacant position (Beckmann & Kuhn, 2012; Forde, MacKenzie, & Robinson, 2008). Our principal-agent model therefore investigates two primary contract types: Recruitment Contracts and Recruit-and-Rent Contracts.

While not having been used in exactly this context, the principal-agent literature dates back to the 1970’s where Spence and Zeckhauser (1971) developed the first model, focusing on insurance. These types of models have since then been used to answer questions in various fields (see e.g. Miller, 2005, for a review). Bendor, Glazer, and Hammond (2001) offer a basic introduction to delegation, and Lewis (2012) contains a review of recent studies on delegated search. Milner and Pinker (2001) considered two types of problems; the first being when the productivity of a temporary worker is difficult to evaluate and subsequently the TWA is used for screening purposes, and the second one entailing the impossibility of creating a socially optimal labour supply contract between a CF and a TWA under hidden action, when the demanded quantity of sufficient, uniformly productive workers is uncertain when the unit price is negotiated.

Hidden action in this context implies that by only being able to observe the supplied worker(s) and not the process (cf. the findings in Beckmann & Kuhn, 2012; Connell & Burgess, 2002), the CF is only able to assess who they get, but not who the alternative workers were (cf. Gibelman, 2005), as any match is only privately observable by the TWA unless supplied to the CF (cf. Halac, Kartik, & Liu, 2016). This allows the TWA to match workers in a way that might not be optimal for the CF, and these differences have mostly been neglected in previous studies, even
though they are potentially highly influential in determining the type of worker that the TWA will supply. One exception is Postl (2004), who found that when an agent is given two alternatives to evaluate, he may only have the incentive to evaluate one and then lie about the quality of the other, resulting in an efficiency loss, as the principal would base its decision on (possibly) incorrect information. Additionally, this aspect of the matching problem has been overlooked in other studies where the agent is contracted to evaluate and provide information on the quality of an already available alternative, or a stream of alternatives, under hidden action (see e.g. Chade & Kovrijnykh, 2011).

This paper expands on the reasoning of Postl (2004) and related papers, but extends the model by letting the TWA itself search for any number of alternatives and also by removing the assumption of a fixed search cost per alternative. This allows us to carry out a more thorough investigation into the outcome of the incentive misalignment resulting from hidden action.

The present paper adds to the existing literature by providing a complementary (or perhaps even an alternative) and structural explanation for the ability to quickly match a worker to a vacant position, which in previous studies has been assumed to be the result of some superior matching technology (Baumann et al., 2011; Neugart, 2005). The results, similar to Beckmann and Kuhn (2012), suggest that CFs should use TWAs to only screen applicants (which are then employed directly at the CF), rather than continuously leasing the workers. We also discuss whether a logical extension of our theoretical results could provide an alternate (demand side) explanation to the increased level of education among the workers in the Swedish temporary work agency sector between 1999 (Joona & Wadensjö, 2008) and 2007 (Westéus, Raattamaa, & Lindgren, 2016), in contrast to the (supply side) rationale offered by Walter (2012).

Furthermore, these results are independent of the price level of the TWAs’ services (see Baumann et al., 2011; Neugart, 2005; Westéus, 2014), in that they only require the existence of a price that the CF is willing to pay and the TWA is willing to accept. This paper therefore does not need to consider optimal pricing for the TWAs’ services. The results instead rely on the assumptions of asymmetric information and an imperfect labour market where the offered wage is related to the vacant position rather than the productivity of the matched worker.

The body of research on the productivity of temporary (agency) employees uses different productivity measures and yields somewhat inconclusive results (compare e.g. Beckmann & Kuhn, 2012; Kleinknecht, Oostendorp, Pradhan, & Naastepad, 2006) and has not considered the potential incentive misalignment suggested in this paper (see Hirsch & Mueller, 2012; Nielen & Schiersch, 2011). This is why it will be up to future research to measure the importance of our contribution. To our knowledge, there are no studies on the performance of workers on fixed-term contracts employed directly at the CF, relative to temporary agency workers at the same firm who perform the same type of jobs. Our model therefore does not fully

1The CFs believe the TWAs to be superior in some aspect(s), making the CFs willing to use the TWAs’ services.
support or reject any of the results in previous studies from an individual worker productivity aspect, but rather it emphasises the difference in the type of worker the TWA might supply. It has similarly been argued by Walter (2012) that TWAs have an incentive to be able to continuously lease their workers, and therefore they might not only match for the specific traits requested by the CF.

The paper is outlined as follows: section two outlines the model for the two aforementioned main types of activities; a Recruitment Contract, or a Recruit-and-Rent Contract. The Recruitment Contract implies that the CF always employs the matched worker directly at the firm whereas the Recruit-and-Rent Contract allows the CF, at each point in time, to make a choice of either subsequently leasing or directly employing the matched worker. Initially the analysis is concerned with the outcome of different types of contracts between the CF and a single TWA, whereas the last part of the analysis is concerned with the effects when there are several competing TWAs. The main results of the model are thereafter summarised, after which the final chapter contains a longer discussion of the models assumptions and implications.

2 Model

The model consists of two types of risk-neutral actors: a CF and one or more external recruitment agencies (TWAs), where the former is defined as any firm having established that there is a demand for an additional worker (i.e. a vacancy) that will be matched by a TWA.

For any vacancy we assume that there are \( J \) possible applicants and each individual \( j \) is identified by his/her unique productivity level, \( x_j \), where the set of available workers is assumed to follow a uniform distribution: \( X = \{ x_1, ..., x_J \} \sim U \left[ x, \bar{x} \right] \).

For each vacancy we assume that there is an objective (i.e. true) exogenous minimum productivity required: \( x^* \), and that the position pays a fixed wage: \( w \), to the worker once filled. The analysis is therefore delineated to when there is a proper non-empty subset \( \chi = \{ x \in X \mid x \geq x^* \} \) containing \( K < J \) elements where each individual has a unique \( x_k \). This allows us to ignore the special cases when there are either no applicants at all, or when only unqualified workers will apply (mar-

\[ x^* \] The choice to employ the worker is final, and the CF may thereafter not lease the worker.

\[ x \] For simplicity, we model the productivity as a scalar, but it could also be modeled as a multi-dimensional vector. The productivity vector would then consist of all possible traits that a worker may have (e.g. preferences on commute distance, age, education, previous job experience, family situation etc.) with a complete set of marginal rates of substitution between every pair of traits.

\[ U \left[ x, \bar{x} \right] \] This simplification is done to keep the mathematics as simple as possible and not divert from the qualitative implications of the model.

\[ x^* \] We discuss the implications of endogenising this parameter at the end of the section.

\[ w \] We argue that this construct is empirically relevant. Optimal marginal wage-setting requires the employer to be able to estimate the individual’s marginal productivity which, outside of a perfect labour market with either piece-work pay or low-cost alternative employment opportunities that allow the workers to self-select, is often quite hard - or even impossible.
ket of lemons). It also implies that the TWA is expected to (asymptotically) be able to identify a worker with a sufficient productivity; \(x_m \in \chi\), which is denoted as a *match*. We define \(x_L, (x_H)\) as a subsequent match that has a lower (higher) productivity than the first match, while still being sufficient, i.e. \(x^* < x_L < x_m < x_H\).

The shape and size of the distribution of workers and its subset is assumed to be known by the TWA, due to its specialisation in creating matches between vacancies and job seekers, but for the same reasons it is assumed to be unknown by the CF. The CF is also either incapable, or unwilling, to monitor any search effort by the TWA other than the actual output: the productivity of the supplied worker.

The model is outlined in discrete time \(t \in (0, 1, ..., T)\) where only one worker may be evaluated at each sub-period. The model also includes a perfect credit market with interest rate \(r\). Following prior simplifications, the constant probability of finding a match at each point in time becomes \(p = p(x_m \geq x^*) = \frac{K}{T} = \frac{x^* - x_m}{x_H - x_m}\). This allows the cumulative probability that a match will be found within a given amount of search periods to asymptotically approach one.\(^7\) For each time period the TWA searches, it incurs a constant cost \(c\).

Finally, a necessary (but not sufficient) condition for a TWA to accept the assignment is that the expected present value stream of payments from the CF, after netting off the TWA’s expected accumulated search cost, is non-negative. Because the model is not concerned with the explicit pricing of the TWA’s services, this condition will always be fulfilled by assumption, so that we are able to focus on the strategic search behaviour of the TWA.

### 2.1 Recruitment Contract

#### 2.1.1 Continuous Payment

A continuous payment contract stipulates that the CF will pay the TWA a fixed amount \(\phi\) at each time period until a sufficient worker is supplied, or the contract expires at time \(T\). The TWA’s present value decision rule for accepting the recruitment assignment can be written as:

\[
\pi = \sum_{t=0}^{T} \left\{ \frac{\phi}{(1+r)^t} \right\} - e \sum_{t=a}^{b} \left\{ \frac{c}{(1+r)^t} \right\} \geq 0, \tag{1}
\]

where \(e \in (0, 1)\) is the TWA’s discrete decision whether to exert effort or not over the time interval \([a, b] \subset [0, T]\). \(a \in [1, T]\) is when \(x_m\) is presented to the CF, or when the contract expires.

The assumption of hidden action makes the CF unable to monitor the actual (search) activity of the TWA. This removes the incentive for the TWA to carry out any search at all and thus the TWA always chooses \(e = 0\) to maximise its

\(^7\)The search duration will follow a geometric distribution due to the fixed probability of finding a match at each turn. We assume that there are either enough applicants, or that individuals may enter and leave the set which makes the best approximation of the probability of finding a match to be constant over time.
profit. This payment scheme thereby does not create any incentive for the TWA to actually carry out any search. The observant reader might notice that \( \pi \) would be stochastic if the TWA would have incentives to reveal a match prior to \( T \). This however will not be the case, as the profit maximising strategy of the TWA does not include any search activity at all. It is not the sequential nature of the contract that drives this result, but rather that payment is not conditioned on the CF actually being supplied a sufficient worker. Conditioning payment on delivery thereby becomes necessary for any effort to be exerted at all.

2.1.2 Payment on Delivery

In order to incentivise the TWA to search, the CF could offer a contract where a fixed payment \( \phi \) is made when \( x_m \) has been supplied. This payment-on-delivery contract ensures that upon accepting the contract, the TWA will start searching immediately, yet it also implies that the TWA will receive the same payment for supplying any \( x_j \in \chi \). Thus there is no reason for the TWA to continue searching for \( x_H \), after having found \( x_m \), as any additional search effort will both increase the TWA’s costs and decrease its present value revenue.\(^8\) The optimal strategy by the TWA is therefore to immediately deliver its first match, whose expected productivity will be \( E(x_m) = \frac{x^* + \tilde{x}}{2} \).

**Proposition 1.** The TWA will never attempt to find an alternative candidate after having found the first match.

This behaviour could create a different outcome compared to the option that the CF would have preferred, as the TWA will stop searching even if there were resources left to conduct additional search.\(^9\) If the CF had perfect monitoring, and thereby a better ability to enforce continuous effort, then the first match would still have a random (but still sufficient) productivity. However any additional resources could then be spent on finding an even better worker until the expected marginal cost of additional search would surpass its expected additional benefit. We define this as searching for the *marginally most productive worker*, whom would have the expected productivity \( E(\tilde{x}_m) = \frac{x^* + \bar{x}}{2} + \gamma \), where \( \gamma \geq 0 \) is the aforementioned productivity difference which is determined by how much additional search would have been profitable for the CF given the search cost and the residual probability of finding a more productive worker.

---

\(^8\)The model will later be extended to relate the productivity of the worker to the expected revenue of the TWA.

\(^9\)If the TWA (by chance) would find and supply the most productive worker in the population \( (x_m = \bar{x}) \), then matching is efficient even for the CF (as engaging in additional search would have a zero probability of finding a better worker). The same is true when the expected productivity gain of additional search is not expected to surpass its costs, had the CF carried out the search itself with the information and cost structure of the TWA. We will however focus our attention on when the best worker is not (necessarily) found right away and where additional search would have been an option.
The search decision for this type of contract is sequential in nature and denoting $\Xi_t$ as the set of draws prior to $t = \tau$ the two conditions for search in the subsequent time period $\tau$ become:

$$\Xi_t \cap \chi = \emptyset$$

$$p\phi - c (1 + r) \geq 0$$

Equation (2) means that the TWA has not found a sufficient worker in the previous period, and Equation (3) is a marginal condition saying that any additional search effort must have a positive expected profit\(^{10}\).

It does not matter if the TWA expects the CF to terminate the contract at some point\(^ {11}\), because even though this would set an upper bound on search activity, no incentives will be altered as payment is immediate upon delivery. Implications of contractually specifying a fixed point in time when payment will be made, conditioned on delivery, will be discussed in the next subsection.

### 2.1.3 Payment at a Pre-specified Point in Time, $T$, Conditioned on Delivery

If the contract is designed in such a way that payment offered by the CF is made at time $T$, conditional on the TWA supplying $x_m$ at or prior to time $T$, then the TWA must consider its probability of finding $x_m$ over time. The present value at time $t = 0$ of the accumulated cost from searching an arbitrary number of periods $[a, b] \in [0, T]$ can be expressed as\(^ {12}\):

$$C(s) = \sum_{t=a}^{b} \left\{ \frac{c}{(1 + r)^t} \right\},$$

Denoting the fixed payment the TWA will obtain at $T$ if $x_m$ is supplied to the CF as $\phi$, the present value expected revenue at time $t = 0$ from searching up to $s$ time periods is:

$$E[R(s)] = [1 - (1 - p)^s] \left\{ \frac{\phi}{(1 + r)^s} \right\}$$

where the first term signifies the cumulative probability that a match will have been found during the $s$ periods of searching.

If there is some exogenous accumulated-cost constraint making the TWA unable to fund search during all periods $t \in (0, ..., T)$, then searching will be deferred towards $T$, to maximise the constrained probability of delivery. Given a positive interest rate and denoting the maximum number of time periods the TWA will afford to search as $s$ if starting to search immediately, and $\bar{s}$ if searching towards the end, we have that $s \leq \bar{s}$ for any contract where $s < T$\(^ {13}\) as any given search effort will have a lower present value cost the later it is expended. When the

\(^{10}\)This condition will always be fulfilled, as the contract has been accepted.

\(^{11}\)E.g. that the TWA is not supplying a match within a “reasonable” amount of time.

\(^{12}\)The effort parameter $e$, previously used in Equation (1), is dropped in (4) as the present value of accepting a contract, and thereafter exerting no effort, will always be zero.

\(^{13}\)The first inequality is not strict due to the discrete nature of the model.
TWA has the option to choose when to search during the lifetime of the contract \( t \in (0, 1, ..., T) \), it is clear that the TWA will prefer to exert search efforts later rather than sooner in order to maximise its present value expected profit.

**Proposition 2.** The TWA will find it optimal to defer searching due to discounting if \( \bar{s} < T \).

Given that the TWA has found \( x_m \) prior to \( T \), the design of the contract makes the TWA indifferent between delivering \( x_m \) immediately, or waiting an arbitrary amount of time\(^{14}\), since the payment is independent of when \( x_m \) is presented. Moreover, this aspect is not optimal for the CF, as it can hardly be worse off by being presented with the (first and final) match as soon as it has been discovered – especially since the first actual day on the job can be subsequently negotiated between the CF and the matched worker.

The model will now be extended to also include the option for the CF to either employ the worker directly, or lease the worker through the TWA. This follows e.g. Houseman (2001) and ECORYS-NEI (2002) in that the main rationale for utilising TWAs is to screen the productivity of a worker before making the decision to either employ the worker, end the collaboration, or continue to lease the worker. It also follows the transitory pattern between temporary agency workers and workers on standard employment contracts in Westéus (2014), Baumann et al. (2011) and Neugart (2005).

### 2.2 Recruit-and-Rent Contracts

#### 2.2.1 A Single TWA

In this setup we assume that any supplied worker is initially always employed by the TWA and leased to the CF on an open-ended contract. At each subsequent point in time, the CF may either choose to continue to lease the worker, or employ the worker directly at the CF which will end the collaboration with the TWA. This corresponds to the TWA assuming all liabilities when the CF screens the productivity of the worker (see Neugart, 2005; Westéus, 2014).

The probability that a sufficient worker will become employed directly at the CF is defined as the worker’s *transition probability*, and is assumed to be an increasing function of the worker’s residual productivity: i.e., \( f(x_i) = \tilde{f}(x_i - x^*) \in [0, 1] \), with the added simplification that it is assumed to be constant over time. This assumption implies that a more productive worker may leave as a result of obtaining a better offer elsewhere (as argued by Walter, 2005), or will have a greater chance of obtaining employment directly at the CF than a less productive worker. The duration of the open-ended contract is \( T \) periods, where \( T \sim \text{Geo}[f(x_i)] \) due to the assumed constant transition probability. This implies that the expected duration of the contract for a worker with productivity \( x_i \) can be expressed as

\[^{14}\text{But no longer than until } T.\]
\[ T(x_i) = \frac{1}{f(x_i)} = N(x_i) + n(x_i); \text{ where } N(x_i) = \left\lfloor \frac{1}{f(x_i)} \right\rfloor \text{ is the integer part, and} \]
\[ n(x_i) = \frac{1}{f(x_i)} - N(x_i) \text{ is the fractional part.} \]

The accumulated search cost for a TWA searching from \( t = 0 \) to \( \hat{s} \) when the match \( x_m \) is found becomes:
\[ C(\hat{s}) = \sum_{t=0}^{\hat{s}} \left[ c(1+r)^t \right] \]

When the CF leases the worker, the TWA charges a wage-proportional fee \( \sigma \cdot w^* \) where \( \sigma > 1 \). The TWA in turn pays the worker \( \delta \cdot w^* \); \( \delta < \sigma \), resulting in a revenue of \( (\sigma - \delta) w^* \) for the TWA.\(^{15}\) Assuming that the profit from the fractional part is always incurred in the last time period, the expected present value profit for the TWA at time \( t = \hat{s} \) becomes:
\[ E(\pi \mid x_m) = -C(\hat{s}) + \left[ \sum_{t=0}^{\hat{s}} \frac{1}{(1+r)^t} + \frac{n(x_m)}{(1+r)^T(x_m)} \right] (\sigma - \delta) w^* \]

The above definition of the expected duration, \( \bar{T}(x_i) \), of a contract states that a worker \( x_H(x_L) \) is expected to transition to employment directly at the CF faster (slower) than the current match as \( f(x_H) > f(x_m) > f(x_L) \), and consequently \( \bar{T}(x_H) < \bar{T}(x_m) < \bar{T}(x_L) \). This implies that the TWA can only expect to compensate for any additional search costs by finding a worker with a lower productivity: \( x_L \), for which the TWA expects to be able to collect its fee for a longer period of time.\(^{16}\)

In order to provide comparable expression for the difference between the expected values of two non-linear stochastic processes, we utilise the short duration of temporary assignments (see e.g. Forde & Slater, 2005) and apply the limit argument to the non-linear term (i.e., allow the interest rate to approach zero). Furthermore, in order to keep the notation simple we will only explicitly model the case when \( \bar{T}(x_L) - \bar{T}(x_m) \geq 1 \) and \( n(x_\theta) \equiv 0 \) for \( \theta = (L, m) \) as this setup relates to the discrete nature of the model the most.

Since the productivity parameter is assumed to follow a uniform distribution, the conditional probability of finding a less productive match becomes \( q_L = q(x^* \leq x_L < x_m) = \frac{(x_m-x^*)-x+1}{x^*-x+1} \). The linearised expressions for the expected additional revenue and cost for the TWA associated with additional search can thereby be written as:
\[ \lim_{r \to 0^+} E(\Delta R) = \lim_{r \to 0^+} \left[ q_L \sum_{\bar{T}(x_m)} \frac{(\sigma - \delta) w^*}{(1+r)^{\bar{T}(x_m)}} \right] = q_L [\bar{T}(x_L) - \bar{T}(x_m)] (\sigma - \delta) w^* \]

\(^{15}\)A number of studies find \( \delta < 1 \), however the EU Temporary and Agency Workers Directive (2008/104/EC) intends to ensure \( \delta = 1 \) (Westéus, 2014). Following Westéus (2014), Baumann et al. (2011) and Neugart (2005) we further assume that \( \sigma > 1 \) since paying the mark-up corresponds to a liability insurance for the CF.

\(^{16}\)If \( x_L \) is not found, then the TWA may still supply \( x_m \).
$$\lim_{r \to 0^+} \Delta C = \lim_{r \to 0^+} \frac{c}{1 + r} = c$$

Proposition 3. Using the above simplifications, additional search will be profitable for the TWA if

$$ql [\hat{T} (x_i) - \hat{T} (x_m)] (\sigma - \delta) w^* \geq c.$$ However, any additional search will always be for a worker with a lower, but still sufficient, productivity.

The above proposition shows that there are situations where the TWA will have incentives to actively act against the best interest of the CF. We denote the resulting misaligned-incentive induced expected productivity level difference as $\kappa$.

Following Proposition 1, the expectation will go from $E (\tilde{x}_m) = \frac{r + x}{2}$ to $E (\tilde{x}_m) = \frac{\tilde{x} + x^*}{2} - \kappa$, where $\kappa \geq 0$ when using a Recruit-and-Rent Contract instead of a Recruitment Contract.

Assuming that TWAs have superior search capabilities, compared to the CF (similar to Baumann et al., 2011; Neugart, 2005), then this would further increase the size of the difference in expected match quality because better matching technology would allow the TWA to screen a larger number of workers for the same amount of resources.

The difference may be mitigated by introducing an opportunity cost for any TWA that finds a match but does not supply him/her to the CF. This can be done by increasing the number of TWAs competing for the assignment, and stating that only the TWA with the most productive match will get to supply the worker to the CF.

2.2.2 Incentives Caused by Competition Among TWAs

In this setup we assume that the CF has engaged $z$ TWAs to find a match for the vacant position. We also assume that there are significantly fewer TWAs than applicants ($z \ll J$) in order to maintain the delineation to only analyse situations where there is actual strategic search behaviour on behalf of the TWAs when searching for a match to the vacant position ($x^*, w^*$)\(^17\). Each search assignment resembles a Bertrand game to some extent, as we assume that only the first TWA to find a match (and in the case of several TWAs making a match in the same period, then the best match) will get to supply the entire worker demand (fixed at one).

To facilitate the analysis we simplify by assuming that the contract duration for a (sufficient) worker is defined by the linear function: $\hat{T} (x_i | x_i \geq x^*) = T (x^*) - \beta (x_i - x^*)$, where $T (x^*)$ is the maximum duration of the lease contract. We also define $\hat{T} (\bar{x}) = \alpha \geq 1$ and express the integer part and fractional part of the expected duration as before: $\hat{T} (x_i) = N (x_i) + n (x_i)$.

We additionally make the assumption that each TWA makes the assessment that any other TWA will always present any match in the same period that the worker found within one period of search and no strategic behaviour on behalf of the TWA could influence the outcome.

\(^{17}\)As $z \to \infty$ we would expect the best worker in the sample to always be found within one period of search and no strategic behaviour on behalf of the TWA could influence the outcome.
is found, and thus cannot expect any strategic search behaviour from its competitors. This highly restrictive assumption simplifies the model by allowing us to disregard any feedback effects among the competing TWAs. It also minimises the TWA’s incentive not to present a match when found.

Assuming that the accumulation of search costs, and the dynamics of how the TWA expects the CF to lease the supplied worker, follows the outline in the preceding subsection, the expected present value profit at time $t = \hat{s}$ for any of the $z$ TWAs having found a match, $x_m$ at that point in time becomes:

$$E (\pi | x_m, z) = - C (\hat{s}) + g (q_m) \cdot (\sigma - \delta) w^* \cdot \left[ \sum_{t=0}^{N(x_m)} \frac{1}{(1 + r)^t} + \frac{n (x_m)}{(1 + r)^{\hat{T}(x_m)}} \right]$$

(10)

where $g (q_m) = \left( \frac{x_m - \frac{1}{8} \hat{s}}{\hat{s} - \frac{1}{8}} \right)^{-1}$ is the probability that none of the other TWAs have found an even more productive match than $x_m$.

Finding $x_H$ when conducting additional search will decrease the TWA’s expected lease time. However, $x_H$ will also increase the overall chance of winning the contract – which is a necessary (but not sufficient) condition for obtaining any profit at all. As before, finding a less productive worker will increase the expected lease time for the TWA, conditional on winning the contract, but will now also reduce the probability that the given TWA is chosen to supply the worker. This is the main difference when adding competition as the TWA is no longer certain it will be awarded the contract when choosing to supply the matched worker to the CF.

To define when additional search is expected to be profitable for the TWA, we again assume that the expected lease duration difference for any two adjacent sufficient workers is at least one time period and $n (x_k) \equiv 0$, while at the same time we allow the interest rate to approach zero and apply the limit argument (cf. Subsection 2.2.1). We denote the $\lim_{r \to 0^+}$ expected revenue from any given match $x_m$ as

$$E [R (x_m)] = g (q_m) \cdot \hat{T} (x_m) \cdot (\sigma - \delta) w^*$$

and define $\chi_{\psi} = \{ x \in \chi : E [R (x)] \geq E [R (x_m)] \}$ as the set of workers yielding a higher expected revenue than $x_m$. We also define the number of elements in $\chi_{\psi}$ as $\mu_{\psi}$, and $p_{\psi}$ as the probability of finding $x_{\psi} \in \chi_{\psi}$. The linearised expression for the average expected revenue of the workers in $\chi_{\psi}$, given $x_m$ and $\mu_{\psi} > 0$[^18], thereby becomes:

$$E \left[ R \left( x_{\psi} \in \chi_{\psi} \right) \right] = \lim_{r \to 0^+} \sum_{\mu_{\psi}} E \left[ R \left( x_{\psi} \in \chi_{\psi} \right) \right] = \frac{\sum_{x_{\psi} \in \chi_{\psi}} g (q_{\psi}) \hat{T} (x_{\psi})}{\mu_{\psi}} (\sigma - \delta) w^*$$

(11)

[^18]: If $\mu (\chi_m) = 0$, then $x_m$ is the most profitable match and the TWA has no incentives to search for another worker.
where \( g(q_y) \) is defined analogously to \( g(q_m) \) above. The expected revenue of searching for a worker with a higher expected revenue (while still retaining the possibility to provide \( x_m \)) becomes:

\[
E \left[ R (x_y | x_m) \right] = p_y E \left[ R (x_y \in \chi_y) \right] + (1 - p_y) E \left[ R (x_m) \right]
\]

and taking into account that no other TWA supplies a worker in the current period, i.e. \( g(q^*) = \left( x^* - \frac{x^* - 1}{x^*} \right)^{z-1} \), the expected change in revenue from additional search can be expressed as:

\[
E (\Delta R) = g(q^*) \left( E \left[ R (x_y | x_m) \right] - E \left[ R (x_m) \right] \right) =
\]

\[
= g(q^*) p_y \left\{ \frac{\sum \left( E \left[ R (x_y \in \chi_y) \right] \right)}{\mu_y} - E \left[ R (x_m) \right] \right\} =
\]

\[
= g(q^*) p_y \left\{ \frac{\sum_{x_y \in \chi_y} g(q_y) \hat{T}(x_y)}{\mu_y} - g(q_m) \hat{T}(x_m) \right\} \cdot (\sigma - \delta) w^* \quad (13)
\]

The linearised additional search cost follows Equation (9), which enables us to define a condition for additional search.

**Proposition 4.** After introducing competition among \( z \) TWAs within the given framework, additional search will still take place if

\[
g(q^*) p_y \left\{ \frac{\sum_{x_y \in \chi_y} g(q_y) \hat{T}(x_y)}{\mu_y} - g(q_m) \hat{T}(x_m) \right\} \cdot (\sigma - \delta) w^* \geq c.
\]

The resulting expression is quite intuitive; given an initial match, the TWA will conduct additional search if it expects additional revenue to surpass its costs. As previously mentioned, competition introduces an additional trade-off regarding the type of worker that will be matched, compared to the preceding subsection with only one TWA where the only trade-off was between expected contract duration and additional search cost. As increasing the number of competing TWAs decreases all individual TWA’s probability of winning, the contract becomes relatively more important to the TWA than the expected duration of the lease. Therefore the worker with the highest expected revenue for the TWA moves to the right in the distribution – i.e. towards a more productive worker. However, increasing the number of competing TWAs also lowers the expected revenue as the probability of being the TWA with the most productive match decreases accordingly.

To facilitate the analysis we plot the expected revenue curves for \( z = \{2, 3, 5, 10\} \) in Figure 1, while assuming that the number of sufficient workers is no larger than the number of non-sufficient workers. Given these parameters, the worker yielding the highest expected revenue for the TWA in the \( z = 2 \) case is still the \( x^* \) worker (similar to the result in Subsection 2.2.1).

Figure 1 shows that even though the CF would benefit from having several TWAs competing for the contract\(^{19}\), the decreasing expected revenue for each TWA will

\(^{19}\)Indeed, if \( z \) would be very large, then \( \hat{x} \) would most likely be supplied within one search period.
most likely restrict the number of TWAs that are willing to compete. Determining the optimal number of competing TWAs from the CF’s point of view will be left for future research, as the model should then include a hiring cost per TWA and also an opportunity cost for the CF for recruiting in-house and assuming all liabilities, which is outside the scope of this paper.

\[ E[R(x_k)] \]

Figure 1: Expected revenue for varying levels of competition

3 Summary

The model in this paper provides a number of important insights regarding the search behaviour of an external TWA hired by a CF to match a worker with a vacant position when the CF is unable to monitor anything else other than the productivity of the supplied worker. Subsection 2.1 utilises the TWA as a filter to find an appropriate candidate, whereas in Subsection 2.2 the TWA (initially) employs the worker while the CF leases the worker. The established search behaviour of the TWA for each contract type is shown to generally differ from trying to find the marginally most productive worker given the available resources.

The implications of the differing incentives establish that a payment-on-delivery contract is a necessary (but not sufficient) prerequisite for the TWA to carry out any actual search, as the TWA could otherwise merely claim to be searching. Subsection 2.1.3 then shows that any long-term worker supply planning on behalf of the CF (modelled as a fixed delivery date prior to which the CF does not need the matched worker; i.e. when planning vacations, etc.) could very well be negated by the TWA, as any search will occur as close to the last time period as possible. The main result from Subsection 2.1 is nevertheless that the TWA will never have any incentives to provide another worker other than the first sufficient candidate. As this corresponds to a random match from the subset of sufficient workers, it may result in a match with a lower average productivity compared to an in-
stance when any remaining resources would have been spent searching for the marginally most productive worker.

Arguably, this also implies that a vacancy (on average) is likely to be filled more rapidly by a TWA even if it would screen prospective candidates somewhat slower than the CF. The relatively quick vacancy/worker matching by TWAs (Autor, 2001, 2003; Houseman, 2001; Mitlacher, 2007), suggested to be the consequence of better matching technologies (Baumann et al., 2011; Neugart, 2005), could thereby be explained by the differing incentives shown in this paper – either in conjunction with actual differences in the available search technologies, or by the incentive structure itself.

Subsection 2.2 outlines a Recruit-and-Rent contract setup where the TWA can be shown to under certain circumstances (Proposition 3), gain from spending additional resources to find the marginally least productive (but still sufficient) worker after a first (random) match has been found. This will result in an even lower average productivity among the supplied matches than in the recruitment case. The theoretical predictions of our model thereby match the empirical results from a panel data study by Beckmann and Kuhn (2012) which found that firms which only use the TWA for screening purposes are found to be more productive than firms that continuously lease their temporary workforce. However, as the TWA now has incentives to conduct additional search, it is less certain that they, on average, will be able to recruit faster than the CF, as argued in the preceding subsection.

The introduced competition in Subsection 2.2.2 reduces the TWA’s incentives to perform additional search by adding an opportunity cost in that it also provides the other TWAs with another possibility to find an even better match. Here, the results are less clear and ultimately depend on the parameters of the model. At a low level of competition the TWA would like to approach $x^*$, but as the competition increases there is a possibility that the TWA may decide to supply a worker with relatively high productivity in order to win the contract. However, there is also a trade-off in that the number of TWAs that the CF will be able to engage depends on the revenue that each TWA expects to make from being awarded the contract – a variable that decreases with the number of TWAs.

4 Discussion

The principal-agent model with hidden action presented here is a relevant framework for studying the recruitment of labour, since one of the main reasons to use a TWA is disengaged from the recruitment process (and thereby reduce the foregone productivity associated with having to filter all prospective candidates; cf. Beckmann & Kuhn, 2012; Connell & Burgess, 2002). Leasing workers from a TWA is also similar to when the employer signs redundancy insurance, which directly relates back to the first principal-agent model by Spence and Zeckhauser (1971). We further claim that the offered wage is more often associated with the specific position rather than the (maximum) productivity of the worker.
The assumed constant probability of finding a match is arguably less intuitive than allowing the set of remaining applicants to shrink after each screening. However, a constant probability of finding a match is a more restrictive assumption that will not only keep the mathematics more comprehensible, but will also provide more conservative results. We also argue that it mimics, to some extent, a dynamic distribution of applicants where individuals could both enter and leave during the search duration.

Any limited liability for the TWA towards the CF (as the model assumes that there is no penalty for not supplying a worker) could potentially facilitate moral hazard problems by inducing TWAs to accept assignments that they do not expect to complete. The TWA will also always have plausible deniability since the time until a match is found (or not) is stochastic. Determining any suboptimal behaviour would thereby require the CFs to either share information among each other, or to utilise the same TWA for a long succession of similar assignments – which are both highly implausible. Furthermore, the CF will have a hard time proving any suboptimal behaviour of the TWA whenever a worker is supplied since a match is always sufficient by definition.

Every TWA will have incentives to claim that they are able to provide the best match, which will make the screening of available TWAs a delicate task for the CF. Ironically, to some extent, this also corresponds to the underlying problem of choosing the right applicant to fill a vacancy – particularly in the presence of complex pricing, the minimum involvement by the CF, and that a number of verdicts from the Swedish Market Court (Marknadsdomstolen) suggest that no TWA has been able to objectively verify that they have any comparative advantage over their competitors that allows them to supply a better match.

Any advantages for the TWA in screening prospective candidates in the recruitment case would also increase the misaligned-incentive induced expected productivity difference even further. This is because any sufficient relative difference (that could be quite small in absolute terms) would allow the TWA to supply a match faster than the CF, while still rejecting any candidates with a higher productivity than the current least productive match.

With an exogenously determined minimum productivity level (and in the absence of several competing TWAs) the model suggests that while the most productive individuals have the same (random) chance as any other sufficient worker of being matched in the recruitment case, they would suffer the greatest penalty in their probability of being chosen for leased contingent work. However, the CF should arguably be able to mitigate the negative effect on the expected productivity of the supplied worker by endogenising the requested minimum productivity level; \( x^\star \star : x^\star \star > x^\star \). Using formal education as a proxy for the productivity parameter and comparing the situation in Sweden in 1999 (when private employment mediation agencies had only been available for 6-7 years; Joona & Wadensjö, 2008) to 2007 (Westéus et al., 2016), there is evidence

---

that the education level in the temporary work sector has increased considerably, and even surpassed that in the regular sector (i.e. any employment that is not through a TWA. See Andersson-Joona & Wadensjö, 2012; Petersson, 2013; Walter, 2012). This could suggest that the CFs have realised the need to overstate the minimum requirements for a vacant position to be filled with a TWA worker. In this case, the results from our model offer a demand side explanation for the increased education level in the TWA sector that is based on the TWA’s profit maximisation, without the need to introduce supply-side effects such as reputation (see e.g. Walter, 2012).

A worker that is matched to a position claimed to require \( x^{**} \), but that objectively only requires \( x^* \), would arguably also feel overqualified and/or mismatched (cf. Loughlin & Barling, 2001; Petersson, 2013) to a larger extent – especially in combination with the sectors lower wages (Andersson-Joona & Wadensjö, 2012), and adverse working conditions (Håkansson, Isidorsson, & Strauss-Raats, 2013). Evidence of this is found by de Graaf-Zijl (2012) in that agency workers with the highest educational attainment show the largest negative difference in job satisfaction (which relates mostly to the content of their job and to job insecurity only to a lesser extent) compared to what the author denotes regular workers.

This paper contributes to the theoretical literature and leaves it open to future research to measure its importance. The next step could be to empirically analyse the assumptions of limited liabilities when a worker is not supplied, and thereafter aim to de-construct the pricing mechanism of the services of TWAs under the misaligned incentives framework that has been outlined above. Future research could also focus on evaluating the relative productivity of agency workers compared to temporary workers employed directly at the CF to test this model’s theoretical predictions. Another possible direction would be to investigate potential gains from eliminating any incentive misalignment, as they could arguably be significant.

References


Department of Economics, University of Birmingham.
Think About the Future and Wait

Tomas Raattamaa
Department of Economics
Umeå School of Business and Economics
Umeå University

Abstract

Constructing a continuous-time principle-agent model we model two types of agents than strive to complete a project by a deadline set by the principal. Project completion is modeled using a Poisson process, and the agents may postpone effort, as not to accrue costs early on. Both agent types exhibit this behavior, but use different decision rules, and we look at the first-best under perfect monitoring, as well as second-best under hidden action. We then show that a rational agent must be more efficient than the principal, by some factor, in order to be considered for the contract. Lastly we show that the principal may increase its profit under hidden action, by sequentially contracting two different agents, in effect firing the first if it does not finish the job by a deadline set earlier than the hard deadline when the principal needs the project completed.

Acknowledgments: The author would like to thank Magnus Wikström, Thomas Aronsson, Morgan Westéus, Tomas Sjögren, Chandra Kiran, as well as the seminar participants at the Department of Economics for their comments and suggestions.
1 Introduction

The principal-agent (PA) literature deals with the incentive structure and potential problems that arise when someone (the principal) delegates a task to someone else (the agent).\(^1\) In this paper we model aforementioned principal-agent relationship in continuous time, where an agent is contracted by a principal to complete an assignment by a deadline set by the principal. This assignment requires the agent to expend costly effort at a fixed rate, in the form of engaging in search. If the agent successfully completes the assignment, then it will be rewarded (i.e. given a prize) at the time of the deadline.

While the basic trade-off between assignment success probability and total cost is well-known, the exact optimal strategy in continuous time with payment by a given deadline, as opposed to payment on delivery, has not been modeled until now.

The example we will use in this paper is a client firm (CF) contracting a temporary work agency (TWA) to find a qualified worker to fill a vacant position at the CF. The reason for such a contract may be that the CF needs someone for a specific holiday such as Christmas or Thanksgiving, or to fill someone’s spot after his/her planned retirement. This makes the CF want to pay when the worker starts, rather than when the TWA signals that it has found a match for the upcoming vacancy.

Although this paper will use the temporary work sector as the main example, the model is more general and may fit a number of problems observable in the world today where options are evaluated, or experimentation and trial-and-error is required for success. This would include, but is not limited to, aforementioned recruiting, sport or talent scouting and certain types of R&D.

Our agent will engage in search, and that is an activity that has been modeled extensively in economics (see McCall, 1970, for an early example) and the literature on the PA relationship with moral hazard, which started in the 1970’s with Mirrlees (1976), Holmstrom (1979) and Grossman and Hart (1983), has remained a popular area of research up until today. The first continuous-time PA model was introduced in Holmstrom and Milgrom (1987), where the agent controlled the drift rate of a Brownian motion and the more recent literature on principal-agent problems in continuous time also often uses stochastic calculus and optimal control (cf. Djehiche & Helgesson, 2014, 2015; Kadan & Swinkels, 2013a, 2013b; Piskorski & Westerfield, 2009; Sannikov, 2008; Schättler & Sung, 1993; Williams, 2008).

The two papers closest to this one are Mason and Välimäki (2015) and Lewis (2012). Mason and Välimäki (2015) look at the incentives of completing a project, when the principal sets the payment period-by-period, but focuses on the continuous-time limit of the game as the time periods’ duration approach zero. The biggest differences compared to this paper though is that they have payment on delivery and that their agent can control its effort level. Lewis (2012) also looks at search

\(^1\)For a good textbook introduction see Laffont and Martinort (2002).
with a deadline for late discovery, which may be extended as a result of agent investment. The paper emphasizes that early discovery is important in agency search, which is not the case in our model, where all that matters is that the discovery is done prior to the deadline. Lewis (2012) does however mention that if it is increasingly costly to search, then the agent will delay current search if it is possible to search in the future and receive the same reward, but this is never formally shown.

Our main contribution to the literature is that we allow the agent to (rationally and time-consistently) wait, i.e. not engage in search immediately, but instead postpone effort as not to incur interest-bearing costs early on, without it being increasingly costly to search.2

This is both interesting and relevant, because when a CF uses a TWA to fill a vacancy both want the TWA so succeed, but because of the trade-off between probability of success and accumulated cost an optimal contract is not possible under hidden action. This trade-off and result in itself is not new, as it has been shown in a discrete setting by Varian (1992), but to the best of my knowledge, waiting behavior has never been analyzed. In addition, we will look at two types of agents; one that is rational and plans ahead, as is the standard assumption of profit maximizing behavior, and a new type of agent that does not plan, but instead uses a simple decision rule of marginal revenue exceeding marginal cost.

As we work in continuous-time we will employ the First-Order Approach (FOA) (see eg. Jewitt, 1988; Rogerson, 1985), which replaces incentive compatibility constraints with first order conditions.

Three other papers that share some similarities with ours are Cvitanić, Wan, and Zhang (2009) and Cvitanić and Zhang (2007), in that they model optimal lump-sum compensation, and Zhu (2013) that specifically allows the agent to shirk and thereby receive a positive payoff, without the contract having a strict deadline. The difference is that they use stochastic control and we instead model search at a fixed level of intensity and then by including discounting will be able to model waiting behavior.

In order to simplify the analysis and focus on the waiting behavior we will not let the principal simultaneously contract multiple agents; therefore one risk-neutral CF will offer a risk-neutral TWA a contract that it may accept or reject. The risk neutrality assumption is standard for the principal in PA models, and sometimes for the agent. As we will introduce an agent which does not plan ahead and thus has no concept of risk we keep the rational agent type risk-neutral to be able to compare the two.

In a simpler discrete-time PA model (without waiting) the optimal contract has been shown to be setting the prize equal to the value of the match, and then use a participation fee equal to the agent’s expected profit (see e.g. Varian, 1992) and this result would also easily translate over to our model. While this is certainly optimal, it is not reasonable to assume a participation fee when modeling the tem-

---

2While one may refer to it as optimal starting, to relate it to the vast literature on optimal stopping (see eg. Halac, Kartik, & Liu, 2016), we will subsequently refer to this behavior as waiting.
porary work sector, as it would mean that TWAs would have to pay to be offered a recruitment job. Our contribution is that, while we do not allow for a participation fee, we will show that an optimal contract will be possible under perfect monitoring, where the principal decides when the agent should start searching and sets the prize equal to the agent’s expected total cost of search. We will thus cover both on the perfect monitoring case where the principal can observe the actions of the agent (i.e. see if it is searching or not) and on what is called hidden action, where the principal does not have the ability to monitor the agent’s actions. While perfect monitoring will result in a contract that is both profit maximizing for the principal and socially optimal, the point of delegation is not having to participate or monitor the process, as the agent is assumed to do a better job than the principal could have. One may suspect though that there will be moral hazard under hidden action, i.e. that the incentives or interests of the two actors will not be fully aligned, where the principal wants the agent to exert as much effort as possible, but the agent will trade off some probability of success for starting to search later and thereby reducing total (expected) cost. We will show that our agents, when faced with the above trade-off under hidden action, will wait longer than the principal would like, thus making it impossible to write an optimal contract.

A second selling point is also that outsourcing recruiting to a TWA should be better as it is specialized, and in PA models it is usually enough that the agent is better, but under hidden action we will show that the TWA cannot only be relatively more efficient, but must be better by some factor for the CF to want to delegate. This is because what the CF loses in the TWA’s moral hazard must be made up for in the TWA’s relative efficiency.

The paper is organized as follows; first the model is described and two types of agents defined and their behavior examined. The first is rational and profit maximizing, and the second only looks at the immediate future, and does not plan ahead. Then a principal is introduced, first with perfect monitoring and then we allow for hidden action. The principal’s profit maximization problem has been examined we will look at its delegation decision and lastly show that it is profit maximizing to split the previously used contract in two, letting another agent take over if the first one fails. This will be shown to be profit maximizing for the principal, despite the agents being homogeneous and searching at a fixed effort level. Finally, the results are summarized and discussed.

2 Model

The model is set in continuous time, where the principal offers an agent a contract at time $t = 0$, stating that the principal will pay the agent a lump-sum prize $\Theta$ at time $t = T$, contingent on the agent supplying a qualified worker for the given vacancy. We therefore assume that there are two types of workers; qualified and not qualified, and that the proportion (share) of qualified workers in the economy, denoted $\lambda$, is public information. The principal values a qualified worker at $V$,
which sets an upper bound on how much the principal would be willing to pay for the agent’s services. The goal of the principal is therefore to choose a prize \( \Theta \in (0, V) \) that maximizes its profit; if it promises to pay \( V \), then it will make no profit (and the agent gets all the surplus). We will also see that if the principal sets the prize too low, then the agent will not exert any search effort, and if the principal sets the prize high enough the agent will be incentivized to start searching immediately. We therefore define the contract range for agent type \( i \) as

\[
\Theta_i \equiv (\underline{\Theta}_i, \bar{\Theta}_i)
\]

where \( \Theta \) is the prize below which the agent will not exert any effort and \( \bar{\Theta} \) the prize at which it will start searching immediately.

The prize \( \Theta \) ultimately chosen by the principal, before offering the contract to the agent, is constant and we assume that a qualified worker does not yield the agent any revenue in itself, so the agent is therefore dependent on the principal for the payment \( \Theta \) at time \( T \), if it is successful in delivering a match.

In order to attempt to find a match the agent needs to engage in search at a cost of \( c \) per unit of time. The decision of the agent of when to start searching can be summed up by the decision variable; waiting, \( \omega \in (0, T) \). This means that the agent will have time left to engage in search for a total of

\[
\bar{\epsilon}(T, \omega) = T - \omega
\]

before the contract expires, and we will refer to this as maximum effort. Search is then modeled as a Poisson process using the Exponential distribution, which has a probability density function (PDF)

\[
f(\bar{\epsilon}; \lambda) = \lambda e^{-\lambda \bar{\epsilon}}
\]

and a cumulative distribution function (CDF)

\[
F(\bar{\epsilon}; \lambda) = 1 - e^{-\lambda \bar{\epsilon}}
\]

where is \( \lambda \) the arrival rate, i.e. the share of qualified workers mentioned earlier. Even though the expected duration of search until a match is found is \( \lambda^{-1} \), this is only an expected value, as no amount of search can guarantee that the TWA finds a qualified worker. While delivery can never be certain, the cumulative probability of success is concave and approaches one as as the agent spends more time searching. Because of this, and because search is costly, we have the well-known trade-off between accumulating costs and increasing the probability of success. The use of the Exponential distribution to model search implies that the agent evaluates workers at a fixed rate. We make this assumption as we do not expect there to be major economies of scale, nor do we consider it appropriate to make the fairly common quadratic cost assumption when it comes to candidate screening within a given temporary work agency. The fixed cost can also be said
to be also supported by our assumption of a sufficiently large economy (i.e. pool of workers), as we then do not have to model search without replacement. Both the principal and agent are risk-neutral and they discount exponentially at the exogenous market interest rate \( r \), yielding the standard discount factor 

\[
   d(t) = e^{-rt}
\]

Because there is a positive interest rate the question the agent is faced with is not only if, but also when to search for a qualified worker. This adds a timing dimension to the problem and because the payment will be realized at \( T \), if a match has been found, then an agent that does not want to spend all available time, i.e. \((0, T)\), searching will defer search, as to incur its costs as late as possible. This paper will use graphs to illustrate the model mechanics, and in order to do this we will use some baseline parameter values. These have been chosen to generate clear and concise graphs and to make the comparison between the CA and MA more intuitive.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>—</td>
<td>Point in time</td>
</tr>
<tr>
<td>( r )</td>
<td>0.1</td>
<td>Interest/discount rate</td>
</tr>
<tr>
<td>( V )</td>
<td>150</td>
<td>Lifetime value to the principal (at ( T ))</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.1</td>
<td>Share of qualified workers in the economy</td>
</tr>
<tr>
<td>( T )</td>
<td>20</td>
<td>(Contract) duration</td>
</tr>
<tr>
<td>( \Theta )</td>
<td>50</td>
<td>Fee promised, conditioned on delivery</td>
</tr>
<tr>
<td>( c )</td>
<td>1</td>
<td>Search cost, per unit of time</td>
</tr>
<tr>
<td>( \omega )</td>
<td>—</td>
<td>Waiting</td>
</tr>
</tbody>
</table>

Table 1: Variable descriptions and baseline values

We will now introduce two types of agents that use different decision rules for when to start searching, as they differ in their time frames of evaluation. The first one, called the Cumulatist Agent (CA), is rational and will choose the optimal waiting \( \omega^*_CA \) based on the cumulative probability of success and the expected cost associated with its search. The second agent type, called the Marginalist Agent (MA), will only look at the immediate future and decide if a small (marginal) amount of effort is worth exerting at any given point in time, which will also result in a waiting rule \( \omega^*_MA \). These two types of agents are extremes at each end
of the planning spectrum; the CA plans as much as possible and the MA as little as it can.

We will now go through the behavior of both agents, one by one, starting with the CA.

### 2.1 The Cumulatist Agent (CA)

The CA is profit maximizing and its decision variable is how long it will wait until it starts searching, meaning that it will start searching at some point in time $\omega^*_C \in (0, T)$ and continues to until it finds a match, or the contract expires. The waiting decision is made at $t = 0$ and the purpose of this section is to find an expression for $\omega^*_C$. We start by using the definition of maximum effort from Equation (2) and the CDF from Equation (4) to define the probability of success, and conversely failure, after some amount of effort $\epsilon$ as

$$p_s(\epsilon; \lambda) \equiv 1 - e^{-\lambda \epsilon} \quad (6)$$
$$p_f(\epsilon; \lambda) \equiv e^{-\lambda \epsilon} \quad (7)$$

Standing at time 0, the total expected present value revenue and cost from searching from $\omega^*_C$ will therefore be

$$TR^e(\omega, \lambda, T) = p_s(\bar{\epsilon}; \lambda) \Theta d(T) =$$
$$= \left(1 - e^{-\lambda \bar{\epsilon}}\right) \Theta e^{-rT} =$$
$$= \left[1 - e^{-\lambda(T-\omega)}\right] \Theta e^{-rT} \quad (8)$$

$$TC^e(\omega, \lambda, T) = c \int_{\omega}^{T} p_f(\epsilon; \lambda) d(t) \, dt =$$
$$= c \int_{\omega}^{T} e^{-\lambda(t-\omega)} e^{-rt} \, dt =$$
$$= c \int_{\omega}^{T} e^{-\lambda(t-\omega)-rt} \, dt =$$
$$= \frac{c}{r + \lambda} \left(e^{-\omega r} - e^{-rT+\omega\lambda-T\lambda}\right) =$$
$$= \frac{c}{r + \lambda} \left[e^{-\omega r} - e^{\omega\lambda-T(r+\lambda)}\right] \quad (9)$$

The discounting of the revenue is based on the assumption that payment is made at time $T$. The cost on the other hand is a flow, and therefore needs to account for the fact that the agent may find a match during its time spent searching. As the agent searches, the probability of not having found a match goes down, which is why we integrate over the probability of failure, as it is also the probability of having to search (and accumulate costs) at any given point in time. All of the
expenditures also have to be discounted down to present value at \( t = 0 \), which is why the discount factor is also used in the integration.

The goal of the CA is to maximize its expected profit

\[
\pi_{CA} (\omega, \cdot) = TR^e (\omega, \cdot) - TC^e (\omega, \cdot) = \\
= \left[ 1 - e^{-\lambda(T-\omega)} \right] e^{-rT} \Theta - \frac{c}{r+\lambda} \left[ e^{-\omega r} - e^{\omega\lambda - T(r+\lambda)} \right] = \tag{10}
\]

by choosing how long to wait, so the maximization problem can be written as

\[
\max_{\omega \in (0, T)} \pi_{CA} (\omega, \cdot) = \max_{\omega \in (0, T)} TR^e (\omega, \cdot) - TC^e (\omega, \cdot) = \tag{11}
\]

where waiting \((\omega)\) is the decision variable.

We will now derive an expression for the optimal waiting, that balances the accumulation of costs and the probability of success.

**Proposition 1.** The CA’s optimal (i.e. profit maximizing) waiting will be \( \omega^*_A = \)

\[
\begin{cases}
T & \text{if } \Theta \leq \frac{\xi}{\lambda} \\
T - \frac{1}{r+\lambda} \log \left\{ \frac{c}{r} \left[ \Theta (r + \lambda) - c \right] \right\} & \text{if } \frac{\xi}{\lambda} < \Theta < \frac{c}{\lambda} \left[ \frac{e^{T(r+\lambda)_{r+\lambda}}}{r+\lambda} \right] \\
0 & \text{if } \Theta \geq \frac{c}{\lambda} \left[ \frac{e^{T(r+\lambda)_{r+\lambda}}}{r+\lambda} \right].
\end{cases}
\]

**Proof.** Using the first-order condition

\[
\frac{\partial \pi_{CA}}{\partial \omega} = \frac{1}{\lambda + r} \left\{ [c - \Theta (\lambda + r)] \lambda e^{\omega(\lambda+r)} + cre^{T(\lambda+r)} \right\} e^{-\omega r - T(\lambda+r)} = 0 \tag{12}
\]

and solving for \( \omega \) yields

\[
\omega^*_A = T - \frac{1}{r+\lambda} \log \left\{ \frac{\lambda}{cr} \left[ \Theta (r + \lambda) - c \right] \right\} \tag{13}
\]

which will be bounded in \((0, T)\). Setting \( \omega^*_A \) equal to \( T \) and 0 yields the contract range;

\[
(\Theta, \tilde{\Theta})_A = \left( \frac{c}{\lambda}, \frac{c}{\lambda} \left[ \frac{e^{T(r+\lambda)_{r+\lambda}}}{r+\lambda} \right] \right) \tag{14}
\]

\( \Box \)

The intuition of these bounds is that if the prize is too low, the TWA will not find it profitable to search, and at some higher prize the TWA will expend maximum time in order to succeed in finding a match.

This means that a CA has to be paid more than \( \Theta_{CA} \) in order to want to search and will start searching immediately if promised at least \( \Theta_{CA} \). While the lower
bound \( \varnothing_{CA} \) may be seen as a participation constraint, the CA may still accept such a contract, or any contract with a smaller prize, but not exert any effort. This is because the CA is indifferent between declining the contract and accepting it and not exerting effort, as they both yield \( \pi_{CA} = 0.3 \)

**Corollary 2.** The width of the contract range decreases as \( r \) becomes smaller; \( \lim_{r \to 0} \varnothing = \varnothing \).

The contract range shrinking as the interest rate decreases makes sense, as the waiting trade-off disappears; if there is no interest rate then there is no difference between spending today and tomorrow. The agent will then start searching immediately if it is paid (at least) the expected cost of search \( (\frac{c}{\lambda}) \).\(^4\) Note that when there is no interest rate, then this setup is equivalent to payment on delivery, because if the value of a future prize does not have to be discounted to present value, then there is no difference to the agent if it is paid when it finds a match, or at some future time \( T \). A positive interest rate is therefore what drives the waiting behavior.

The partial derivatives of the CA’s optimal waiting decision can be found in Table 2 and all but the partial derivative w.r.t. \( \lambda \) can be signed,\(^5\) the effect of the arrival rate on the CA is inconclusive, as the first and last terms in the numerator are negative, while the middle term is positive. The intuition is that we have two opposing effects at work; as it gets easier to find a match search is incentivized as its expected payoff increases, but on the other hand the agent does not have to spend as much effort to get the same expected result, disincentivizing search. This is illustrated in Figure 1 using the baseline parameters.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Waiting ((w_{CA}^*))</th>
<th>Partial derivative</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( \frac{1}{\varnothing} )</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>( c )</td>
<td>( \frac{\varnothing}{c</td>
<td>\varnothing(\lambda+r)-c</td>
<td>} )</td>
</tr>
<tr>
<td>( \varnothing )</td>
<td>( \frac{c-\varnothing(\lambda+r)}{c} )</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>( r )</td>
<td>( \frac{\varnothing}{c-\varnothing(\lambda+r)} - \frac{\log\left{ \frac{c(\varnothing(\lambda+r)-c)}{(\lambda+r)^2} \right} + \frac{\lambda}{r}} )</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>( \lambda )</td>
<td>( \frac{c-\varnothing(\lambda+r)}{c} - \frac{\log\left{ \frac{c(\varnothing(\lambda+r)-c)}{(\lambda+r)^2} \right} - \left( \frac{\lambda}{r} + 2 \right)} )</td>
<td>+/−</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Partial derivatives of the CA’s optimal waiting duration

\(^3\)While we do not model competition or varying degrees of efficiency among agents, this could be a strategy for a CA not to leave money on the table if there would be some competitor better at searching, i.e. with a higher \( \lambda \), who would get the contract and associated positive expected profit.

\(^4\)As mentioned above, \( \lambda^{-1} \) is the expected value of an exponentially distributed random variable.

\(^5\)See Appendix A.1 for proof of \( \frac{\partial w_{CA}^*}{\partial r} > 0 \).
To give the intuition to how the CA makes its maximization we will show how expected profit and total cost and revenue are affected by the agent’s waiting. Figure 2 shows plots of $TR^e(\omega, \cdot)$, $TC^e(\omega, \cdot)$ and $\pi_{CA}(\omega, \cdot)$, for various levels of waiting, given different prizes. Figure 2a uses the baseline $\Theta$ and yields an interior solution which can be verified using Equation (13), and the interpretation is that the CA will wait for 9 of the 20 units of time and then start searching. Figure 2b and 2c then show that corner solutions are obtained when the CF promises the prize $\Theta_{CA}$ or $\bar{\Theta}_{CA}$ respectively. Notice that the expected total cost curves are the same in all three graphs, as they do not depend on the prize.

We will now move on to the other agent, which does not look at all possible waiting durations, but instead makes instantaneous decisions of whether to exert a small amount of effort.

### 2.2 The Marginalist Agent (MA)

While the CA is rational and will objectively maximize the expected profit of any contract it is assigned, it can be argued that not everyone plans ahead. Because of the growing literature on behavioral economics and the computational limits of human cognition we now introduce a heuristic that represents the other extreme of the planning spectrum.

In the context of a temporary work agency trying to fill a vacancy an employee may consider if it is worth looking at an application at any given point in time, without thinking about all applications he/she is likely to have to go through before finding a match. To use another example, think of someone who is at a corner shop and decides to buy a lottery ticket. This person will think about the probability of winning when buying that one ticket, and compare this to the discounted value of the prize to be received at some later date. This behavior

---

6The plots use the baseline parameter values from Table 1 on page 5.
7See Kahneman (2013) for an introduction.
Figure 2: How expected total revenue, cost and profit depend on waiting
is probably more common than the person thinking about all the lottery tickets it may have to buy before it wins, in addition to also taking into account that it should wait to start buying lottery tickets, which is exactly the behavior of the CA.

To model the behavior of someone who acts as the lottery ticket buyer would be likely to do, i.e. evaluate buying a ticket at that given moment, we introduce the Marginalist Agent (MA), which only looks at the immediate future; instead of comparing expected profit of all possible waiting durations it evaluates, at every point in time, whether it can expect to make a profit by exerting a tiny amount of effort. It will do this by comparing the expected marginal revenue and expected marginal cost, defined as the revenue and cost of searching for a tiny amount of time \( s \).

Instead of choosing \( s \) we will have it approach zero, as that lets us keep the unit of time undefined, i.e. not specify if it is for example days or months, and also lets us interpret the expected revenue and cost as truly “marginal”. This is preferable to arbitrarily assuming e.g. \( s = 1 \), as using the limit lets us look at how the MA would act if it has no planning horizon, making it the polar opposite of the CA.

This means that the MA will not solve a maximization problem, but instead its evaluation is done continuously at every point in time \( t \), based on the following condition

\[
MR^e(t) \geq MC^e
\]  

where

\[
MR^e(t) = \left(1 - e^{-\lambda s}\right) \Theta e^{-r(T-t)}
\]

\[
MC^e = c \int_0^s p_f(t) d(t) \, dt = c \frac{1 - e^{-s(r+\lambda)}}{r + \lambda}
\]

Standing in some time period \( t \in (0, T) \) the expected marginal revenue will be based on the probability of success after a small amount of search \( s \) multiplied by the size of the prize \( \Theta \) promised at time \( T \) if the agent is successful, discounted down to the time at which the evaluation is made. The expected marginal cost will integrate over the probability of failure and the discount factor. This is similar to how the CA does calculates its expected (total) revenue and cost of search, but with the difference that the MA looks at the expected revenue and cost of searching for a tiny amount of time \( s \), instead of the expected cost of searching until it succeeds.

**Proposition 3.** *The MA’s waiting will be*
\[ \omega_{\text{MA}}^* = \begin{cases} T & \text{if } \Theta \leq \frac{c}{\lambda} \\ T - \frac{1}{r} \log \left( \frac{\Theta}{\bar{\Theta}} \right) & \text{if } \frac{c}{\lambda} < \Theta < \frac{c}{\lambda} e^{rT} \\ 0 & \text{if } \Theta \geq \frac{c}{\lambda} e^{rT} \end{cases} \]

**Proof.** We can see that the expected marginal cost is time-independent and constant, whereas the expected marginal revenue will be monotonically increasing as we get closer to the expiration of the contract \((T)\). This is because as time passes the time-invariant expected prize will occur closer in time, thereby raising its present value. Setting

\[ MR^e(t) = MC^e \]

and solving for \(t\) gives us the first and only time expected marginal revenue crosses the expected marginal cost from below. This \(t\) is therefore the first point in time that the agent will be indifferent between searching and not searching, after which it will find the expected marginal revenue to exceed expected marginal cost. This \(t\) is therefore the MA’s waiting, \(\omega_{\text{MA}}\), when planning for the amount of time \(s\) and solving for it yields

\[ \omega_{\text{MA}} = T - \frac{1}{r} \log \left( \frac{\Theta (\lambda + r) (1 - e^{-\lambda s})}{c \left[ 1 - e^{-s(\lambda + r)} \right]} \right) \]  

(19)

As we want the MA to plan as little as possible we take the limit (and use L’Hospital’s rule twice) to arrive at the MA’s waiting

\[ \omega_{\text{MA}}^* = \lim_{s \to 0} \omega_{\text{MA}} = \]

\[ = T - \frac{1}{r} \log \left( \frac{\Theta}{\bar{\Theta}} \right) \]  

(20)

This waiting will be bounded in \((0, T)\) and thus it follows that the contract range is found by setting \(\omega_{\text{MA}}^*\) equal to \(T\) and \(0\) respectively, and solving for the fee;

\[ (\Theta, \bar{\Theta})_{\text{MA}} = \left( \frac{c}{\lambda}, \frac{c}{\lambda} e^{rT} \right) \]  

(21)

The MA has the same lower bound as the CA, and here the waiting mechanism also disappears as the interest rate goes towards zero. Rearranging and taking the limit on Equation 15 gives us

\[ \Theta e^{-r(T-t)} \geq \frac{c}{\lambda} \]  

(22)

which we can plot as the MA’s expected marginal revenue and expected marginal cost to show its waiting in Figure 3. Notice that the LHS is the present value of
the prize at time \( t \) and the right hand side is the expected cost of search, absent of any interest rate. We have removed the numbers from the y-axis as they would be meaningless; they decrease as \( s \) gets progressively smaller, but it does not matter as the curves maintain their relative positions.\(^8\)

Looking at the partial derivatives of the MA’s waiting duration shown in Table 3 we see that they have the expected signs. All partial derivatives, except for \( \lambda \), have the same signs as the CA’s and have the same interpretations. The effect of the arrival rate is negative, showing that the easier it is to find a match, the more the agent will want to search. This means that since the MA does not plan, it does not face the same effect on \( \lambda \) as the CA; a higher probability of success will only make the MA start searching earlier, as it does not take into account that it could potentially save (interest rate) costs by starting to search later.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Waiting (( \omega_{MA}^{*} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Partial derivative</td>
</tr>
<tr>
<td>( T )</td>
<td>( \frac{1}{\sqrt{T}} )</td>
</tr>
<tr>
<td>( c )</td>
<td>( \frac{1}{\sqrt{c}} )</td>
</tr>
<tr>
<td>( \Theta )</td>
<td>( -\frac{1}{\sqrt{\Theta}} )</td>
</tr>
<tr>
<td>( r )</td>
<td>( \frac{1}{\sqrt{r}} \log \left( \frac{1}{c \Theta} \right) )</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>(-\frac{1}{\sqrt{\lambda}} )</td>
</tr>
</tbody>
</table>

Table 3: Partial derivatives of the MA’s optimal waiting duration

\(^8\)Plotting Equation (16) and (17) for very small \( s \) produces the same result.
2.3 Comparing the Agents

We have now described two types of agents; the CA which plans ahead and the MA that only looks at the immediate future.

The intuition for finding the optimal amount of waiting is similar to the standard model of a firm choosing its (deterministic) output level so that marginal revenue equals the marginal cost. Looking at Table 4, this is the behavior of the CA, as it chooses waiting, which implicitly sets the output (probability), while the MA instead uses a condition which, although “marginal”, does not maximize profits.

<table>
<thead>
<tr>
<th>Agent</th>
<th>Decision rule</th>
<th>Timing</th>
<th>Horizon</th>
<th>Resulting waiting, $\omega^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>$\frac{\partial TR^c}{\partial \omega} = \frac{\partial TC^c}{\partial \omega}$</td>
<td>At $t = 0$</td>
<td>$(0, T)$</td>
<td>$T - \frac{1}{r+\lambda} \log \left{ \frac{\lambda}{cr} [\Theta (r + \lambda) - c] \right}$</td>
</tr>
<tr>
<td>MA</td>
<td>$\frac{\partial TR^c}{\partial t} \geq \frac{\partial TC^c}{\partial t}$</td>
<td>At every $t$</td>
<td>$\rightarrow 0$</td>
<td>$T - \frac{1}{r} \log \left( \frac{\Theta \lambda}{c} \right)$</td>
</tr>
</tbody>
</table>

Table 4: Decision rule summary

Which of the two agents will start searching first? Figure 4 shows the waiting behavior $\omega^*_i \in \{MA, CA\}$ using the baseline parameters, where it is apparent that the CA waits longer than the MA. This can in fact be shown to hold true for all parameter sets.

![Agent waiting](image)

Figure 4: Agent waiting

**Proposition 4.** If effort is exerted by at least one agent, and unless the agents both exert full effort, the MA will start searching earlier than the CA.

**Proof.** Referring back to Proposition 3 and Proposition 1 we see that the agents both need to be paid more than $\Theta_{CA} = \Theta_{MA} = \frac{c}{\lambda}$ in order to exert effort. As
both agents’ waiting will be monotonically decreasing in the promised fee (but bounded in \((0, T)\)) we compare the partial derivatives of waiting, to show that

\[
\frac{\partial w^\star_{MA}}{\partial \Theta} < \frac{\partial w^\star_{CA}}{\partial \Theta}
\] (23)

which simplifies to

\[
(\Theta \lambda - c) [\Theta (\lambda + r) - c] > 0
\] (24)

and as all parameters are assumed to be positive Equation (23) holds true for any fee \(\Theta > \frac{c}{\lambda}\), i.e. for the ranges where the agents exert effort.

Having a longer time horizon when planning one’s actions will (on average) pay off, even-though it results in longer waiting, because the (true) expected profit will be higher. Now that we have looked at these two types of agents we will turn our attention to the behavior of the principal.

3 Principal

The principal will offer an agent a contract that takes the form \(<D, \Theta, \cdot>\) where \(D \in (0, T)\) is the deadline and \(\Theta\) is the prize to be paid by the principal conditioned on the agent having delivered a qualified worker by the deadline. Since the principal needs the worker by \(t = T\), it will always set \(T\) as its deadline when only contracting a single agent. This will however not be the case in subsection 3.4 where we will let the principal contract two agents sequentially, in essence replacing the first if it does not deliver and will thus need to be able to offer two different contracts with different deadline and prizes.

While we have two types of agents we will only have one principal, since the MA’s behavior does not translate over to the principal; as the principal’s decision is made at \(t = 0\) there is no appropriate heuristic for attempting to make a profit “as soon as possible”. Thus the principal acts similarly to the rational CA, in that it takes into account the end result when deciding how to construct the contract. The principal has full information on agent type \((i \in \{CA, MA\})\), its unit search cost \((c)\), the share of qualified workers in the economy \((\lambda)\) and the market interest rate \((r)\), so it will account for the agent’s behavior by selecting the profit-maximizing fee \((\Theta)\), and if perfect monitoring is possible, also the agent’s waiting \((\omega)\). Thus the general formulation of the principal’s present value expected profit function is

\[
\pi_{CF} |_{i \in \{MA, CA, \ldots\}} = p_{d, i} (\Theta, \omega, \cdot) (V - \Theta) e^{-rT}
\] (25)

where \(p_{d, i}^\star (\Theta, \omega, \cdot)\) is the probability of delivery by agent type \(i\), identical to the probability of success defined in Equation (6).
Before looking at how to construct the contracts we should mention that a trivial solution to this problem already exists (cf. Varian, 1992); set the prize to the value of the worker ($V$) and have the agent pay a participation fee equal to its expected profit, because that way the principal would get all the surplus. This is however not reasonable to implement in the context of the temporary work sector, which is why we only mention this the contract setup here.

We now will start looking at how the principal should construct a contract under perfect monitoring. This means that there will be no moral hazard as the principal is assumed to be able to see if the agent is exerting effort or not. It can therefore, in addition to deciding the promised payment also decide when the agent should start searching for a qualified worker.

### 3.1 Perfect Monitoring

Assuming that perfect monitoring of the agent by the principal is cost-less, then the principal may dictate the terms and effectively choose the waiting behavior of the agent. Although this is a strong assumption, not likely to be fulfilled in the real world, we will nevertheless use it as a benchmark.

The principal will not be able to make a MA search more than it would under imperfect monitoring, as it already starts searching when it expects to make a profit, but the principal will be able to affect the waiting behavior of the CA, as it if unmonitored consciously waits in order to maximize its expected profit.

The principal’s optimization problem under perfect monitoring will therefore be to set both the fee and how much (if any) waiting it will allow, while making sure that the agent expects to spend all its resources on search:

$$\max_{\Theta \in \left( \frac{e^{-\lambda}}{r}V, V \right), \omega \in (0,T)} \pi_{CA} = \left( 1 - e^{-\lambda(T-\omega)} \right) (V - \Theta) e^{-rT}$$

The first part is the probability of a qualified worker being delivered, the second is the value of the worker minus the amount the principal has to pay to the agent for delivering said worker, and the last term is the discount factor, as payment will occur at time $T$, conditioned on payment. There is also a zero-profit condition for the agent, as the principal will take all the surplus. We will now characterize the optimal contract.

**Proposition 5.** The optimal contract under perfect monitoring will be

$$\langle D, \Theta, \omega \rangle_{CF}^* = \left( T, \frac{\lambda V - c}{r \left( \frac{V(r+\lambda)-c}{r+\lambda} \right)^{\frac{1}{r+\lambda}} - 1}, T - \frac{1}{r+\lambda} \log \left( \frac{\lambda}{\lambda} \left( V (r+\lambda) - c \right) \right) \right).$$

**Proof.** The maximization problem in Equation (26) can be reduced to a maximization in one dimension by taking the zero-profit condition of the agent ($\pi_{CA} = 0$) and solving it for the prize $\Theta$, giving us

---

9Note that the effect will be the same as that of a participation fee.
\[ \Theta^* (\omega) = c \left\{ \frac{e^{-(\lambda+r)\omega} \left[ e^{(\lambda+r)T - e^{(\lambda+r)\omega}} \right]}{(\lambda + r)} \right\} = \frac{c}{\lambda + r} \left[ \frac{e^{(\lambda+r)(T-\omega)} - 1}{e^{\lambda(T-\omega)} - 1} \right] \]  

which is the fee that equals the agent’s expected total expenditure for a given level of waiting. Taking this fee and plugging it back into the objective function function in Equation (26) simplifies the maximization to

\[ \max_{\omega \in (0,T)} \pi_{CF}|_{CA} = \left[ 1 - e^{-\lambda(T-\omega)} \right] [V - \Theta^* (\omega)] e^{-rT} \]  

(27)

The objective function can therefore be written as

\[ \pi_{CF}|_{CA} = \left[ 1 - e^{-\lambda(T-\omega)} \right] [V - \Theta^* (\omega)] e^{-rT} = \left[ 1 - e^{-\lambda(T-\omega)} \right] \left\{ V - \frac{c}{\lambda + r} \left[ \frac{e^{(\lambda+r)(T-\omega)} - 1}{e^{\lambda(T-\omega)} - 1} \right] \right\} e^{-rT} \]  

(28)

and its partial derivative w.r.t. waiting becomes

\[ \frac{\partial \pi_{CF}|_{CA}}{\partial \omega} = \frac{e^{\lambda\omega -(\lambda+r)T}}{\lambda + r} \lambda \left\{ -c e^{(\lambda+r)(T-\omega)} + c + V (\lambda + r) \left[ e^{\lambda(T-\omega)} - 1 \right] \right\} + \frac{e^{\lambda\omega -(\lambda+r)T}}{\lambda + r} \left[ -c (\lambda - r) e^{(\lambda+r)(T-\omega)} - \lambda V (\lambda + r) e^{\lambda(T-\omega)} \right] = \frac{1}{\lambda + r} \left\{ c \left[ \lambda + r e^{(\lambda+r)(T-\omega)} \right] - \lambda V (\lambda + r) \right\} e^{\lambda\omega -(\lambda+r)T} \]  

(29)

Which by setting equal to zero and solving for the waiting yields

\[ \omega_{CF}^* = T - \frac{1}{r + \lambda} \log \left\{ \frac{\lambda}{c r} [V (r + \lambda) - c] \right\} \]  

(30)

and plugging that into the fee we get the optimal fee

\[ \Theta_{CF}^* = \frac{\lambda V - c}{r \left\{ \frac{\lambda}{c r} [V (r + \lambda) - c] \right\} ^{\frac{1}{\lambda + r}} - 1} \]  

(31)

\[ \Theta_{CF}^* = \frac{\lambda V - c}{r \left\{ \frac{\lambda}{c r} [V (r + \lambda) - c] \right\} ^{\frac{1}{\lambda + r}} - 1} \]  

(32)

\[ \square \]

Corollary 6. The contract range under perfect monitoring will be

\[ (\Theta, \tilde{\Theta})_{MO} = \left( \frac{c}{\lambda}, \frac{c}{\lambda+r} \left[ \frac{e^{(\lambda+r)T}}{e^{\lambda T}} - 1 \right] \right) \]  

(33)

17
Proof. The contract range can be found by using the limits on the fee coming from the zero-profit condition;

\[ \Theta_{MO} = \lim_{\omega \to T} \Theta^*(\omega) = \frac{c}{\lambda} \]

(33)

\[ \tilde{\Theta}_{MO} = \lim_{\omega \to 0} \Theta^*(\omega) = \frac{c}{\lambda + r} \left[ \frac{e^{(\lambda+r)T} - 1}{e^{\lambda T} - 1} \right] \]

(34)

The principal will face the same trade-off between probability of success and total costs as the agent, illustrated in Figure 5a. There we see that if the principal has the agent start searching immediately then the principal will get the profit in point A, while the agent gets zero profit in point D. If the principal instead pays according to the optimal contract, and has the agent wait, then the principal’s profit will be maximized in point B, while the agent still gets zero profit in point E. This is all assuming that there is perfect monitoring, so that the principal can decide the waiting duration. Notice that \( \pi_{CF}^e(\omega, \Theta^*(\omega)) \) and \( \pi_{CA}^e(\omega, V) \) are identical, as are \( \pi_{CF}^e(\omega, V) \) and \( \pi_{CA}^e(\omega, \Theta^*(\omega)) \). This is because they show what happens when one player gets the entire surplus, and none respectively.

In the next subsection we will introduce hidden action, where the principal cannot observe effort, and thus can only set the prize. In Figure 5 we can see what happens if the principal sets the fee \( \Theta^* \) but cannot control the agents waiting behavior. Instead of point E, the agent would chose point F, maximizing its profit, while the principal’s profit is lowered to point C. This means that the optimal fee is no longer \( \Theta^* \), as the principal has to adjust to the agent’s moral hazard. This is exactly what the principal will do next.

3.2 Hidden Action

Hidden action, as opposed to perfect monitoring, is where the principal cannot observe the actions of the agent. It is therefore not possible for the principal to add a clause in the contract stipulating \( \omega_{CF}^* \) as under perfect monitoring. Instead, the principals only decision variable is the fee it promises to pay at \( T \), if the agent delivers a qualified worker.

Hidden action is a common assumption in principal-agent models, where monitoring is impossible, or too costly to be practical. With the example of a CF contracting a TWA to fill a vacancy hidden action is a perfectly reasonable assumption, as the task is outsourced so that the CF does not have to do the searching itself.

We will now go through the optimization problems of a principal contracting first a CA, and then a MA.
Figure 5: The principal's optimization
3.2.1 Contracting a CA

Starting with the CA, the maximization problem now only contains the decision variable $Q$, and the principal can no longer enforce the zero-profit condition as it did under perfect monitoring. As the principal now has to account for the agent choosing its own waiting duration the maximization problem of the principal becomes

$$\max_{\Theta \in \left(\frac{r}{T}, V\right)} \pi_{CF} = \left[1 - e^{-\lambda(T - \omega)}\right] (V - \Theta) e^{-rT} \quad (35)$$

s.t. $\omega = \omega^*_CA(\Theta, r)$

Internalizing an unmonitored CA’s waiting behavior $\omega^*_CA(\Theta, r)$ from Equation (13) gives us the objective function

$$\pi_{CF|CA} = \left[1 - e^{-\lambda(T - \omega^*_CA(\Theta, r))}\right] (V - \Theta) e^{-rT} =$$

$$= \left[1 - e^{-\lambda[T - \frac{1}{\lambda} \log \left\{ \frac{1}{\Theta(r + \lambda) - c} \right\}]}\right] (V - \Theta) e^{-rT} =$$

$$= \left[1 - \left\{ \frac{cr}{\lambda \left[\Theta(r + \lambda) - c\right]} \right\} \frac{1}{\lambda} \right] (V - \Theta) e^{-rT} \quad (36)$$

and thus the principal’s maximization can be simplified to

$$\max_{\Theta \in \left(\frac{r}{T}, V\right)} \pi_{CF} = \left[1 - \left\{ \frac{cr}{\lambda \left[\Theta(r + \lambda) - c\right]} \right\} \frac{1}{\lambda} \right] (V - \Theta) e^{-rT} \quad (37)$$

where the first-order condition

$$\frac{\partial \pi_{CF|CA}}{\partial \Theta} = \left(\frac{c - r\Theta - \lambda V}{c - r\Theta - \lambda \Theta} \left\{ \frac{cr}{\lambda \left[\Theta(r + \lambda) - c\right]} \right\} \frac{1}{\lambda} - 1\right) e^{-rT} = 0 \quad (38)$$

is an in-decomposable polynomial because of the exponent $\frac{1}{\lambda r}$. This means that there is no global (closed form) solution, but we can still find local maxima by making assumptions about the relationship between $r$ and $\lambda$. If, for example, we assume that $r = \lambda$, then we can show that the optimal fee will be the first root of

$$8\Theta^3 r^3 - 13\Theta^2 cr^2 + \Theta \left(8c^2 r - 2cr^2 V\right) - 2c^3 + 2c^2 r V - cr^2 V^2 = 0 \quad (39)$$

for $r > \frac{c}{\pi}$. This would correspond to e.g. a 1% interest rate in an economy where 1% of the pool of workers would be qualified for the vacancy at the client firm. While this seems reasonable, changing the relationship between $r$ and $\lambda$ to say $r = 2\lambda$ creates a really large expression for which we need to find the first
root. The point is to show that, while there is no closed form solution to this problem, the model is nevertheless solvable. Before providing graphs to show the intuition, we will go through the principals pricing strategy when instead contracting a MA, so that we can compare the two.

### 3.2.2 Contracting a MA

The setup of this maximization problem is identical to that in previous subsection, except for the principal now accounting for a MA’s behavior, instead of a CA’s.

$$\max_{\Theta \in \left( \frac{c}{\lambda}, V \right)} \pi_{CF} = \left[ 1 - e^{-\lambda(T - \omega)} \right] (V - \Theta) e^{-rT}$$  \hspace{1cm} \text{s.t. } \omega = \omega_{MA}(\Theta, \cdot) \tag{40}$$

Internalizing the MA’s waiting behavior \( \omega_{MA}(\Theta, \cdot) \) from Equation (19) gives us the objective function

$$\pi_{CF\mid MA} = \left[ 1 - e^{-\lambda(T - \omega_{MA}(\Theta, \cdot))} \right] (V - \Theta) e^{-rT} =$$

$$= \left( 1 - e^{-\lambda \{T - \frac{1}{2} \log(\frac{\Theta}{\lambda}) \}} \right) (V - \Theta) e^{-rT} =$$

$$= \left[ 1 - \left( \frac{c}{\lambda \Theta} \right)^{\frac{1}{r}} \right] (V - \Theta) e^{-rT} \tag{41}$$

and thus the principals maximization problem above can be simplified to

$$\max_{\Theta \in \left( \frac{c}{\lambda}, V \right)} \pi_{CF\mid MA} = \left[ 1 - \left( \frac{c}{\lambda \Theta} \right)^{\frac{1}{r}} \right] (V - \Theta) e^{-rT} \tag{42}$$

with a first-order condition

$$\frac{\partial \pi_{CF\mid MA}}{\partial \Theta} = \frac{\lambda [\Theta r + \lambda (V - \Theta)] - cr}{r (c - \Theta \lambda) \left( 1 - \frac{\Theta \lambda}{c} \right)^{\frac{1}{r}}} e^{-rT} = 0 \tag{43}$$

that is unfortunately not explicitly solvable for \( \Theta \), because of the exponent \( \frac{1}{r} \); depending on the relationship between the arrival and interest rate the above polynomial will grow or shrink in size, similar to what happens when contracting a CA.

Figure 6a shows the contour plot for the principal’s optimal prize \( (\Theta_{MA}) \) when we let the interest rate \( (r) \) and arrival rate \( (\lambda) \) vary. As the iso-price lines are concave, by picking any point in the graph it must be the case that the optimal price decreases as the match is easier to find, and increases as the interest rate increases. Figure 6b is just as intuitive; the principal’s profit increases as the arrival rate increases (i.e. the share of qualified worker increases, making it easier to find a match) and decreases as the interest rate increases.
Figure 6: Optimal pricing and the principal’s profit, when contracting a MA

Just as we could find a solution to the optimal pricing problem when contracting a CA, we can find solutions for when contracting a MA, if we assume some relationship between $r$ and $\lambda$. If we again assume that $r = \lambda$, then the profit maximization simplifies to

$$
\pi_{CF|MA,r=\lambda} = \left(1 - \frac{c}{\lambda \Theta}\right) (V - \Theta) e^{-rT}
$$

and standard profit maximization yields

$$
\Theta_{MA}^* | r=\lambda = \sqrt{\frac{c}{\lambda} V}
$$

But this is only an example of a closed form solution that can be found. Assuming instead that $r = \frac{1}{2} \lambda$ gives us the optimal price as the first root of

$$
4\Theta^3 \lambda - \Theta^2 c - 2\Theta c V - cV^2 = 0
$$

Our purpose is not to characterize all possible solutions to these maximization problems, but instead we turn our attention to the difference between contracting a CA or a MA.

### 3.2.3 Summary and Comparison

As concluded earlier, in order to make a profit the principal needs to pay $\Theta \in (\xi, V)$ to incentivize effort, but not to make a loss if a match is delivered. Figure 7 shows the principals profit and optimal fee when contracting either agent under imperfect monitoring. From the graph it is clear that if the principal pays too little, then it will make zero profit, as none of the agents will conduct any search, and if it promises to pay more than the match is worth it will expect to make a
loss. We can also see that the profit from contracting a MA is higher from that of contracting a CA, simply from the fact that the MA will start searching earlier for any $\Theta$, as was shown earlier. This means that the principal will prefer to contract a MA if it is to be unmonitored, but a CA if perfect monitoring is possible.

When looking at Figure 7 the optimal prize under hidden action is (almost) the same for both agents, but this is not always the case; when comparing the optimal fee paid to either agent, we find an interesting relationship. Figure 8 shows that the optimal price is fairly similar when $r = \lambda$, but that the MA is more sensitive to changes in this relationship. The plot is for the baseline parameters, varying $\lambda$, but other parameter sets tested show the same pattern; if the interest rate is higher than the arrival rate then the MA should be paid more than the CA and the reverse is true if the interest rate is smaller than the arrival rate. It should be noted however that large enough arrival rates will make the CA wait more, because of effect shown in Figure 1, and thus the optimal prize offered to the CA will be bigger than that of the MA, even if the arrival rate is smaller than the interest rate. This will however happen at unreasonable large interest rates.

![Diagram](image.png)

Figure 7: The principal’s profit, given agent type and fee (baseline parameters)
What about social optimum? The gray points in Figure 8 are the fees that generate the effort level the principal would have chosen itself, had it not delegated search. From the graph it is apparent that the principal will not offer any of the two agents such a contract under hidden action, as it would not be profit maximizing for the principal. This will always be the case for the CA, as the standard result from the literature applies.

**Proposition 7.** The socially optimal prize when contracting a CA under hidden action is $\Theta_{CA}^s = V$

**Proof.** This follows from

$$\omega_{CA}^* (\cdot) = \omega_{CF}^* (\cdot) \Rightarrow T - \frac{1}{r + \lambda} \log \left\{ \frac{\lambda}{cr} [\Theta (r + \lambda) - c] \right\} = T - \frac{1}{r + \lambda} \log \left\{ \frac{\lambda}{cr} [V (r + \lambda) - c] \right\}$$

and solving for the fee ($\Theta$) yields

$$\Theta_{CA}^s = V$$

This is the problem of agency; if the agent is only rewarded partially (i.e. less than $V$) for its effort it will not search enough, but if it will be rewarded $V$ it then becomes the residual claimant and will exert the socially optimal amount of

---

These can be considered socially optimal, given that no player has an outside option.
effort, which is the standard result in the PA literature (see e.g. Lewis, 2012). The principal’s profit when promising to pay \( V \) will be zero, which is shown in point \( D \) in Figure 7, as well as point \( D \) in Figure 5a. The MA on the other hand will have a different socially optimal prize, and it will be possible to end up in social optimum when contracting a MA, even when the principal is profit maximizing.

**Proposition 8.** The socially optimal fee when contracting a MA is

\[
\Theta_{\text{MA}}^s = \frac{c}{\lambda} \left\{ \frac{\lambda}{cr} [V (\lambda + r) - c] \right\}^{\frac{r}{\lambda + r}}
\]

**Proof.** Setting

\[
\omega_{\text{MA}}^s (\cdot) = \omega_{\text{CF}} (\cdot) \Rightarrow
\]

\[
\Rightarrow T - \frac{1}{r} \log \left( \frac{\Theta}{\lambda} \right) = T - \frac{1}{\lambda + r} \log \left\{ \frac{\lambda}{cr} [V (r + \lambda) - c] \right\}
\]

and solving for the fee (\( \Theta \)) yields

\[
\Theta_{\text{MA}}^s = \frac{c}{\lambda} \left\{ \frac{\lambda}{cr} [V (\lambda + r) - c] \right\}^{\frac{r}{\lambda + r}}
\]

\[
(52)
\]

**Corollary 9.** The socially optimal prize when contracting a MA will not always be higher than the principal’s profit optimizing prize.

**Proof.** The principal’s objective function will only have one maximum, so evaluating \( \frac{\partial \pi_{\text{CF}} |_{\text{MA}}}{\partial \Theta} \) at \( \Theta_{\text{MA}}^s \) and checking its sign would allow us to say if the principal always wants to pay more or less than the socially optimal amount. While we are able to sign the derivative for some arbitrary relationships between \( r \) and \( \lambda \) (such as \( r = \lambda \)) we cannot sign it for all.

\[
\frac{\partial \pi_{\text{CF}} |_{\text{CA}}}{\partial \Theta} \big|_{\Theta=\Theta_{\text{MA}}^s} = e^{-rT} \left( 1 - \left\{ \frac{\Lambda \Gamma}{cr} \right\}^{-\frac{r}{\lambda + r}} \right) \cdot
\]

\[
\cdot \left( \frac{cr}{\lambda} \right)^{-\frac{\lambda}{\lambda + r}} \left\{ \frac{c - \lambda V}{\left( \frac{cr}{\lambda} \right)^{\frac{r}{\lambda + r}} - \Gamma} \right\}^{\frac{r}{\lambda + r}} - 1
\]

\[
(53)
\]

where \( \Gamma = V (\lambda + r) - c \).

\[\text{11}\]This notation is only used so that the expression does not overflow the page.
when it is easy to find a worker it is being paid less than socially optimal. Looking at Figure 7 we can see that a positive partial derivative means being to the left of the maximum, meaning that the fee paid by the profit maximizing principal is above the social optimum, while conversely a negative partial derivative means the contracted prize is less than the socially optimal. It is therefore possible that the environment is such that the contract the principal offers the MA is socially optimal.

This concludes our look at the principal’s maximization problem and social optimality. We will now move on to two extensions; the first being when the principal should delegate under hidden action, assuming it could do the recruiting itself, and the last looks at what happens if the principal can contract agents sequentially, i.e. contract another agent if the first one fails.

3.3 Extension: When to Delegate Search

So far we have assumed that the principal has already made the decision to delegate search, but we have said nothing about this decision. The standard reason for any delegation in principal agent literature is that the agent is more efficient than the principal. In our model this comes down to the two actors having different arrival rates ($\lambda$).

Looking at the perfect monitoring case the principal should delegate if $\lambda_{CA} > \lambda_{CF}$, because this means that the agent can be expected to find a match faster than the principal.

If there is hidden action, on the other hand, then it is not sufficient for the agent to have an arrival rate larger than the principal’s, because recruitment should then be delegated if

![Figure 9: The principal’s partial derivative of profit, given arrival rate (and base parameters)](image-url)
The principal’s iso-profit curves
The principal’s contract curve, $\lambda_{CF} = 0.1$
The agent’s contract curve, $\lambda_{CA} = 0.1$
The agent’s contract curve, $\lambda_{CA} = 0.173$

Figure 10: Delegation decision

\[ \pi_{CF}^{*}\omega_{CF},\lambda_{CF} > \pi_{CF}^{*}\omega_{CA},\lambda_{CA} \]  

If we let the principal search using the baseline parameters, and assume that $\lambda_{CF} = 0.1$, then finding the lowest efficiency of the agent that will make the principal delegate can be obtained numerically by finding the root of

\[ \pi_{CF}^{*}\omega_{CF},\lambda_{CF} - \pi_{CF}^{*}\omega_{CA},\lambda_{CA} = 0 \]  

The decision is shown in Figure 10 and Equation (55) yields that the CA’s arrival rate needs to be at least $\lambda_{CA} \approx 0.173$ for the principal to be indifferent between spending money searching itself (Point A), or paying the agent less, for a lower probability of success (Point C). If the CA would have had the same arrival rate as the principal (i.e. $\lambda_{CA} = 0.1$) then delegating search would have yielded the principal lower profit (Point B), as the agent’s contract curve is tangent to a lower iso-profit curve for the principal.

The conclusion here is that the agent needs to make up in efficiency what the principal loses in the agent’s rent-seeking.

3.4 Extension: Replacing Agents by Sequential Contracting under Hidden Action

As noted in Lewis (2012), when you formulate a similar model, but use optimal control and variable effort level the principal can make search more urgent by threatening to replace the agent if it does not complete search with by some given date. Since our model has a fixed effort level that the agent cannot affect (as $\lambda$ is a constant) a threat of replacement cannot work in the same way, but we will show that it will still be beneficial for the principal in our model to contract a second CA if the first one fails.
Using the definition of maximum effort in Equation (2) and the CA’s optimal waiting in Equation (13) we get that the maximum effort exerted by a CA will be
\[
\bar{\epsilon}_{CA} = \frac{1}{r + \lambda} \log \left\{ \frac{\lambda}{cr} \left[ \Theta (r + \lambda) - c \right] \right\}
\] (56)

As we can see, this is independent of \(T\), so setting the deadline earlier than \(T\) will not affect the time the agent allocates for search, since it will always start searching when there is \(\bar{\epsilon}_{CA}\) units of time left until the deadline. We showed in subsection 2.1 that the CA would wait less if the prize was increased, so conversely
\[
\frac{\partial \bar{\epsilon}_{CA}}{\partial \Theta} = \frac{1}{\Theta (r + \lambda) - c} > 0
\] (57)
meaning that if the agent promises a larger prize the maximum effort the CA will devote to search increases. However, we can also see that
\[
\frac{\partial^2 \bar{\epsilon}_{CA}}{\partial \Theta^2} = -\frac{\lambda + r}{\left[ c - \Theta (\lambda + r) \right]^2} < 0
\] (58)
i.e. the marginal effect of the size of the prize on maximum effort is is decreasing.\(^\text{12}\) For notational convenience we now drop the bar and subscript on the maximum effort variable, \(\epsilon\).

**Proposition 10.** For any given maximum effort duration, the principal’s expected cost will be smaller if it plans to replace the first agent with another one, in case the first one fails after some specified amount of time.

**Proof.** This proof will consist of comparing the expected cost of contracting two agents sequentially versus the expected cost of contracting a single agent for the same duration. The principal must therefore specify an earlier deadline for the first agent, and if it fails to deliver then the principal contracts another agent to search for the remainder of time until \(T\).

We will start by making the assumption that
\[
\epsilon (\Theta_1 + \Theta_2) = \epsilon (\Theta_1) + \epsilon (\Theta_2)
\] (59)
which says that the maximum effort level by the single agent (LHS) should be the same as the maximum effort of the sequentially contracted agents (RHS). Thus it follows that the expected total revenue of contracting one or two agents sequentially will be the same;
\[
TR_{n=1}^\epsilon = \left[ 1 - e^{-\lambda \epsilon (\Theta_1 + \Theta_2)} \right] = \left[ 1 - e^{-\lambda \left( \epsilon (\Theta_1) + \epsilon (\Theta_2) \right)} \right] = TR_{n=2}^\epsilon
\] (60)
but it does not mean that the total expected cost will be identical.

\(^{12}\)This can also be seen by looking at Figure 4.
Equation (59) gives us the payment to the second agent as a function of the payment to the first agent

$$\Theta_2 (\Theta_1) = \frac{c(\Theta_1(\lambda + r) - c)}{(\lambda + r)(\Theta_1\lambda - c)} > 0 \quad (61)$$

and we can now formulate our total expected costs. Starting with the total cost when contracting a single agent, this is just probability that the agent will succeed multiplied by the prize:

$$TC_{n=1}^e (\Theta_1, \Theta_2) = p_s (\Theta_1 + \Theta_2) (\Theta_1 + \Theta_2) =$$

$$= \left[ 1 - e^{-\lambda \epsilon (\Theta_1 + \Theta_2)} \right] (\Theta_1 + \Theta_2) \quad (62)$$

Note that we do not discount down to present value, as this expected total cost will be compared to the expected total cost of discounting. When contacting two agents the first agent will get a contract $\langle D, \Theta \rangle = \langle T - \epsilon (\Theta_2), \Theta_1 \rangle$ at time $t = 0$, and if it fails then the second agent will be offered a contract $\langle D, \Theta \rangle = \langle \epsilon (\Theta_2), \Theta_2 \rangle$ at $T = T - \epsilon (\Theta_2)$. We will therefore have an expected total cost of the form

$$TC_{n=2}^e (\Theta_1, \Theta_2) = p_s (\Theta_1) \Theta_1 e^{\epsilon (\Theta_2)} + p_f (\Theta_1) p_s (\Theta_2) \Theta_2 =$$

$$= \left[ 1 - e^{-\lambda \epsilon (\Theta_1)} \right] \Theta_1 e^{\epsilon (\Theta_2)} + e^{-\lambda \epsilon (\Theta_1)} \left[ 1 - e^{-\lambda \epsilon (\Theta_2)} \right] \Theta_2 \quad (63)$$

The payment to the first agent is multiplied by its probability of success, but is also multiplied by a factor $e^{\epsilon (\Theta_2)}$ since the payment will incur interest until $T$. The payment to the second agent is multiplied by the probability that the first agent fails, and then by its own probability of success.

We now compare the total expected costs to see if one of them will be larger

$$\Delta TC_{1,2}^e (\Theta_1, \Theta_2) = TC_{n=2}^e (\Theta_1, \Theta_2) - TC_{n=1}^e (\Theta_1, \Theta_2) =$$

$$= p_s (\Theta_1) \Theta_1 e^{\epsilon (\Theta_2)} +$$

$$+ p_f (\Theta_1) p_s (\Theta_2) \Theta_2 -$$

$$- p_s (\Theta_2 + \Theta_2) (\Theta_1 + \Theta_2) =$$

$$= \left[ 1 - e^{-\lambda \epsilon (\Theta_1)} \right] \Theta_1 e^{\epsilon (\Theta_2)} +$$

$$+ e^{-\lambda \epsilon (\Theta_1)} \left[ 1 - e^{-\lambda \epsilon (\Theta_2)} \right] \Theta_2 -$$

$$- \left[ 1 - e^{-\lambda \epsilon (\Theta_1+\Theta_2)} \right] (\Theta_1 + \Theta_2) \quad (64)$$

and inserting Equation (61) into Equation (59) and simplifying gives us that
\[ \Delta TC_{1,2}^e (\Theta_1, \Theta_2 (\Theta_1)) < 0 \] (65)

if

\[ \left( \frac{\Theta_1 \lambda}{\Theta_1 \lambda - c} \right)^{\frac{1}{1+r}} > 1 \] (66)

assuming all parameters to be positive and \( \Theta_1 > \Theta = \frac{c}{\lambda} \), from which it also follows that \( \Theta_1 > \frac{c}{r+\lambda} \).\(^{13}\) Since Equation (66) will always hold true we can conclude that sequential contracting gives the principal a higher profit than only writing a single contract.

Even if the principal has to pay the first agent at the expiration of its contract \( T - \epsilon_2 \) it will be more profitable than hiring one single agent for the full duration. Note that we have said nothing about the optimal size of the contracts when presenting them in sequence, as our proof only showed that there will always be a better sequential contract, for any single contract.

4 Summary and Discussion

This paper showed, in accordance with mechanisms known from earlier literature, that an agent hired to complete a project may wish to postpone its effort, at the expense of the project’s probability of success (cf. Varian, 1992). In this model the waiting was time-consistent and driven by discounting, and since the exponential form as shown to be sufficient to produce the behavior, the model did not need to rely on a hyperbolic discounting.

The first novel contribution was the introduction of a Marginalist Agent that does not plan ahead, and as a result it did not wait as long to start exerting effort as the rational Cumulatist Agent. The first-best under perfect monitoring was characterized, and while the optimal amount of search would be achieved by giving the CA \( \Theta = V \), it was shown to be possible to achieve optimum by giving an MA less - sometimes even less than the principal’s profit maximizing amount. Either way, since the MA starts searching earlier it would be strictly preferred by the principal to contracting a MA instead of a CA.

The principal’s optimization covered both perfect and imperfect monitoring, and some closed form solutions were found for the hidden action case. Furthermore, intuition was given using numerical examples, graphs and simulations, were the calculus was too complicated to provide any useful guidance.

The extensions contributed two insights; the agent must be better than the principal by some factor for the principal to want to delegate under hidden action, and it is profit maximizing to sequentially contract two agents, than only contracting one.

\(^{13}\)See Appendix ??

30
In the literature on procrastination, whoever is impulsive is more likely to procrastinate, but here the “impulsive” MA will be the first to start searching. While a TWA most likely would not reveal to what extent it engages in waiting behavior, suggestions for future work include empirically testing if people are more like the Cumulatist, or the Marginalist.

References


Appendix

A  The Cumulatist Agent’s Waiting

A.1  Partial Derivative w.r.t \( r \)

\[
\frac{\partial \pi_{CA}}{\partial \omega} = \frac{1}{\lambda + r} \left\{ [c - \Theta (\lambda + r)] \lambda e^{\omega (\lambda + r)} + cre^{T(\lambda + r)} \right\} e^{-\omega r - T(\lambda + r)}
\] (67)

\[
\frac{\partial \pi_{CA}}{\partial \omega} = 0 \Rightarrow
\]

\[
\Rightarrow \omega_{\ast CA}^r = T - \frac{1}{r + \lambda} \log \left\{ \frac{\lambda}{c r} \left[ \Theta (r + \lambda) - c \right] \right\}
\] (69)

\[
\frac{\partial \omega_{\ast CA}^r}{\partial r} = \frac{\log \{ \lambda \left[ \Theta (\lambda + r) - c \right] \} + (\lambda + r) \left[ \frac{\Theta}{c - \Theta (\lambda + r)} + \frac{1}{r} + T \right]}{(\lambda + r)^2}
\]

\[
= \frac{- \log (c) + T (\lambda + r) + \log (r)}{(\lambda + r)^2} = \frac{c - \Theta (\lambda + r)}{(\lambda + r)^2} - \log \left( \frac{c}{\lambda (\Theta (\lambda + r) - c)} \right) + \frac{1}{r}
\] (71)

Signing this partial derivative we will see that it is positive.

Proposition 11. \( \frac{\partial \omega_{\ast CA}^r}{\partial r} = \frac{c}{\Theta (\lambda + r) - c} \log \left( \frac{c}{\lambda (\Theta (\lambda + r) - c)} \right) + \frac{1}{r} > 0 \)

Proof. The first term in the numerator will be negative, while the second and third positive. The second term will be positive, as \( \frac{c}{\lambda (\Theta (\lambda + r) - c)} < 1 \Rightarrow \Theta > \frac{c}{\lambda} \), assuming positive parameters, which is satisfied by Proposition 1.

Digging deeper we see that

\[
\frac{\partial^2 \omega_{\ast CA}^r}{\partial r \partial\Theta} = \frac{\lambda + r}{\Theta (\lambda + r) - c} + \frac{c (\lambda + r)}{(c - \Theta (\lambda + r))^2} = \frac{\Theta}{(c - \Theta (\lambda + r))^2} > 0
\] (73)

and

\[
\lim_{\Theta \to \frac{c}{\lambda}} \frac{\partial \omega_{\ast CA}^r}{\partial r} = 0
\] (74)
Young Adults in the Swedish Temporary Agency Sector: Implications of Family Experience

Morgan Westéus
Department of Economics
USBE
Umeå University

Tomas Raattamaa
Department of Economics
USBE
Umeå University

Urban Lindgren
Department of Geography
and Economic History
Umeå University

Abstract
A persons first experience of working life is not the individuals actual first job, but rather the perception conveyed by his or her family and other reference groups. Using Swedish register data on young adults (aged 18-34), and controlling for personal characteristics, we find that individuals with family members or partners with work experience from the temporary agency sector are highly over-represented in the sector. The peer-groups previous experience is also found to be among the most influential variables determining the relative probability that an individual will work in the temporary agency sector.

Acknowledgements: The authors wish to thank Thomas Aronsson, Karl-Gustaf Lofgren, discussants Magnus Wikstrom and Sofia Tano as well as the seminar participants at the Department of Economics for their comments and suggestions. We would also like to thank the anonymous reviewer(s) from LABOUR for the good comments and revision suggestions, which have improved this paper.
1 Introduction

An increasing body of research explores the effect of the parents working life experience on the labour market outcome of their children. Whiston and Keller (2004) survey the literature on family impact on youth career development and find that the family, in their role of “informal role models” (Morningstar, 1997), have a stronger influence than any other reference group – either through structural variables (education, occupation and socio-economic status) or process variables (relationships, aspirations and support; see also Davis-Kean 2005; Penick and Jepsen 1992; Young and Friesen 1992).

While having a job provides economic opportunities, self-sufficiency and relative independence for the worker – the stress of an increased risk of becoming unemployed does not only relate to the potential negative impact on individual wealth, and the opportunity for temporal reallocation of that wealth (Kalleberg, 2009). It also includes significant non-pecuniary costs (Winkelmann & Winkelmann, 1998) due to that employment is the feature – other than family affiliation – that individuals associate their identity with the most (Skans, 2009). Job insecurity has been shown to result in lower private consumption (Benito, 2006) and Darity, Jr and Goldsmith (1996) claim that living with job insecurity might even induce more stress than actually losing the job. This since the former implies a state of uncertainty, whereas the latter forces the individual to adapt to a realised situation (cf. De Cuyper & De Witte, 2005; Sverke, Hellgren, & Näswall, 2002).

Barling, Dupre, and Hepburn (1998) find that seeing their parents go through downsizing, lay-offs and job insecurity will also have a negative effect on the children's preconceptions, perceptions and attitudes towards working life. Similarly, Kind and Haisken-DeNew (2012) find a lasting negative impact to the subjective well-being of sons of fathers who became (involuntarily) unemployed, and Gregg, Macmillan, and Nasim (2012) find that children of displaced fathers run a higher risk of getting lower grades, earning lower wages, or even becoming unemployed themselves.

A person's first experience of working life is therefore arguably not the individual's first employment per se, but rather the experience conveyed by various reference groups through their experiences and economic circumstances (see also Barling et al., 1998; Kanfer, 1993).

It is important to note that atypical employment arrangements are not always negative per se. Individuals who actively choose this type of employment (due to their personal preferences and skills) are more likely to experience the positive aspects of such work arrangements Loughlin and Barling (1999). Some authors also note that there could even be important cultural differences (Connolly & Rogers, 2001), or differences in norms and expectations (De Cuyper & De Witte, 2007; De Cuyper et al., 2008), that influence how the individual will perceive atypical work. It is when the flexibility is perceived as being precarious that the negative effects arise (Sverke et al., 2002).

The main purpose of this paper is to estimate the relative effect on the probability
of working in the Swedish temporary agency sector for a cohort of young adult workers (aged 18-34 in 2007)\(^1\) of having parents, siblings and/or a partner with previous experience of working for a temporary work agency (TWA). We also control for multiple individual characteristics that influenced the relative probability of working in the Swedish temporary agency sector in 1999 (see Joona & Wadensjö, 2008).

Younger workers were chosen due to the Swedish temporary agency sector’s relative youth (having been deregulated in 1993), and also that this age group cohort is the largest within the Swedish temporary agency sector Andersson-Joona and Wadensjö (2012). That these workers are in the process of forming their own experiences and expectations of the labour market makes this cohort a particularly relevant object to study – especially as Macky, Gardner, and Forsyth (2008) suggests that these traits will remain stable into adulthood.

TWA employment was chosen since most register data delineates this specific type of atypical employment. The surveyed literature in the following section also suggests that the temporary agency sector is characterised by a significant wage penalty and other adverse working conditions. There is also no evidence of temporary agency employment being a stepping-stone into regular employment in Sweden (except for some groups of immigrants). Stressors such as these contribute to poorer well-being and could also constitute a precarious job situation (Sverke et al., 2004). As a consequence, partners and family members of temporary agency workers might (either actively or passively) become dissuaded from working in the sector themselves.

Conversely, the close personal networks of friends and family constitute one of the most common ways of obtaining employment in Sweden; 15-20 percent of the open-ended contracts, and 20-25 percent of the fixed-term contracts, while fewer than 10 percent obtain their employment through the Swedish Public Employment Service (SCB, 2013). This could possibly imply an effect opposite to the one suggested above – perhaps even regardless of the surveyed negative aspects\(^2\).

The paper also includes a detailed analysis of both primarily gainfully employed workers (förvärvsarbetande)\(^3\) and a group that we denote as students. We argue that this is a potentially important distinction as these groups could differ in their incentives of working through a TWA. The individuals in the student-group have been largely excluded in previous studies of the Swedish temporary agency sector (see e.g. Andersson & Wadensjo, 2004; Andersson-Joona & Wadensjö, 2010; Joona & Wadensjö, 2008) which makes this an important addition.

---

\(^1\)This is the largest cohort in the Swedish temporary agency sector (Andersson-Joona & Wadensjö, 2012).

\(^2\)The Swedish TWA employer organisation presents a survey according to which more than two thirds of the workers would recommend work in the temporary agency sector to their friends and family, whereas only about 20 percent state that they would not recommend employment in the agency sector (Bemanningsföretagen, 2012).

\(^3\)The classification of gainfully employed workers mainly follows ILOs definition (working at least one hour/week). SCB uses a model based approach to triangulate this information from official income statements (wages, income and employment transfers), gender, age and qualitative answers from the Labour Force Survey into the RAMS data base.
The remainder of the chapter surveys the research on the characteristics of the temporary agency sector as well as the effect of family experience on working life. Chapter 2 describes the data and outlines the model. Chapter 3 presents the results for the full sample and the two sub-samples; students and gainfully employed. The following chapter (4) summarises the results, and the final chapter (5) includes a discussion of the results and issues that remain to be investigated in future research.

1.1 Work in the Temporary Agency Sector

The temporary agency sector in Sweden has grown considerably after having been deregulated in the early 1990s (Andersson-Joona & Wadensjö, 2010; Johnson, 2010), and has in previous studies been found to consist mainly of workers with a traditionally weak position on the labour market; youths, low-educated or unskilled workers, immigrants and women (Andersson-Joona & Wadensjö, 2010; Joona & Wadensjö, 2008).

The worker in the temporary agency sector is employed by the TWA, but continuously leased to a client firm to work under the supervision and guidance of that firm (either with specific tasks, or side-by-side with the regular workforce of the client firm). This creates the sector’s characteristic tripartite relationship where the liabilities between the worker and the TWA are regulated by labour law, while the arrangement between the TWA and the client firm is delimited general contract law statutes.

Employment protection for the temporary agency worker is arguably lower than for direct hire employees (cf. Håkansson, Isidorsson, & Strauss-Raats, 2013) as claiming just-cause for dismissing the worker due to a lack-of-work becomes easier when the assignment with the client firm has in fact been discontinued. There is also evidence that temporary agency workers are indeed used as a buffer to more easily adjust the size of the labour force while protecting the core of regularly employed workers (Heery, 2004; Spermann, 2011). This practice is also found to be more common in companies that are experiencing a more volatile demand for their products (Salvatori, 2009; Thommes & Weiland, 2010).

A more liberal view by the labour unions with respect to the use of temporary agency workers could also allow for a reduced level of employment protection by allowing for deviations through a collective agreement. This could e.g. allow for other types of atypical employment that does not apply to direct hire workers, or potentially even remove the Swedish norm of seniority-based employment protection.

With regard to the characteristics of temporary agency employment, Andersson-Joona and Wadensjö (2012) show a significant, and (mostly) increasing, wage penalty for Swedish temporary agency workers from 1998-2008. The wage penalty is persistent even when controlling for individual characteristics. In a previous study, the authors find the same pattern even when controlling for sample selection biases (such as the over-representation of physicians and highly qualified
IT personnel), and also conclude that the workers who are leaving the temporary agency sector for a job outside the sector enjoy the highest wage increase (Andersson-Joona & Wadensjö, 2010).

Studies in other European countries have also yielded similar wage penalties for temporary agency workers (Autor, 2009; Forde & Slater, 2005; E. J. Jahn, 2008; Tijdens, van Klaveren, Houwing, van der Meer, & van Essen, 2006), and Niennhuser and Matiaske (2006) even find that the negative wage penalties exist regardless of national legislation requiring the temporary agency worker’s wages to correspond to that of similar workers employed directly at the client firm. Forde and Slater (2005), Salvatori (2009) and Heery (2004) even state that carrying out an equivalent job for a lower wage, and under worse working conditions, may facilitate a view that temporary agency workers are a type of second-tier employees.

Håkansson et al. (2013) survey the literature on the temporary agency sector’s physical and psychosocial work environment and find that workers in the sector experience the lowest degree of autonomy and overall job satisfaction compared to other types of employees. They also find that workers in the temporary agency sector exhibit a significantly higher risk of feeling depressed and that working conditions for temporary agency workers are significantly worse than for similar workers employed directly at the client firm (see also Fabiano, Curro, Reverberi, & Pastorino, 2008; Tijdens et al., 2006).

Another stated rationale for working in the temporary agency sector is that it is assumed to facilitate a transition into regular employment (compared to remaining unemployed). However, using data from 2000-2008 in a difference-in-difference model, Hveem (2013) finds no evidence of a stepping-stone effect in Sweden. The results rather suggest that temporary agency employment increases the time until the worker obtains employment in the regular sector (although not so for non-western immigrants, cf. Andersson & Wadensjo, 2004).

Güell and Petrongolo (2007) argue that the probability to transition from temporary employment into what we in this paper denote the regular sector is mainly affected by outside options, and the ability of the worker to credibly threaten to quit the temporary job (cf. also Loughlin & Barling, 2001). Koslowsky (1998) and De Witte and Nåswall (2003) note that the combination of (involuntary) temporary employment and job insecurity is the worst combination for the worker, and that the associated negative effects of these two factors could very well strengthen each other multiplicatively rather than additively.

The reviewed characteristics of the sector furthermore correspond to several of the stressors that could constitute a precarious job situation for the worker (Sverke et al., 2004). However, the results in De Cuyper and De Witte (2007) also suggests that the impact of similar characteristics could be less negative for temporary

---

4See e.g. Tijdens et al. (2006) or Spermann (2011). Bemanningsföretagen (2012) report that more than sixty percent of those employed by a TWA in Sweden 2012 would rather work directly at the client firm. However, the same study also shows that very few are offered to make this transition and that only about one third got a new assignment with the TWA when their current assignment expired.
workers than for permanent employees.

The theories on psychological contracts (Argyris, 1960; Koslowsky, 1998) suggest that the worker’s subjective perception of a job is based on the compliance of the prior expectations and the realized outcome. A worker that expects adverse working conditions may consequently not experience the situation in the same way that another worker with different expectations. One possible explanation for the results may therefore lie in different labour market cultures, norms and expectations (Connolly & Rogers, 2001; De Cuyper & De Witte, 2007; De Cuyper et al., 2008).

1.2 The Family Experience Effect on Working Life

A growing body of research emphasises the importance of family and other peer group’s conveyed experience on the outcomes of the individual; e.g. criminal activity, education choice and labour market outcome – especially in the presence of job insecurity (see e.g. Davis-Kean (2005), Newman (2005), Gregg et al. (2012) Kind and Haisken-DeNew (2012), or see Whiston and Keller (2004) for a review). Macky et al. (2008) write in an editorial introduction and overview on generational differences at work that the development of beliefs, expectations and values is theorised to be influenced by early environmental stimuli and human socialisation – and that these features remain stable into adulthood. Major socio-economic events such as important changes to the family work pattern, pervading unemployment rates and the deterioration of job security through downsizing and offshoring/outsourcing, are mentioned specifically as highly influential factors (cf. Egri & Ralston, 2004). The reviewed findings by Penick and Jepsen (1992), Young and Friesen (1992), Barling et al. (1998), Whiston and Keller (2004), Davis-Kean (2005), Gregg et al. (2012) and Kind and Haisken-DeNew (2012) lend further support to these theories.

Loughlin and Barling (2001) similarly argue that a persons first contact with working life is the experience, perceptions and opinions conveyed by his or her family and other immediate reference groups. The impact of these reference groups with regard to (atypical) employment, occupational health and safety, management, and labour unions thus becomes especially important to include when studying the labour market outcome for the cohort of young adults.

There is a wide literature on occupational inheritance, but Aldrich and Kim (2011) argues that the biggest interest among researchers has been on more prestigious and highly rewarded occupations (e.g. lawyers, doctors and business owners). There is however also a fairly large interest in intergenerational transmission of entrepreneurship; some papers focus on values and intention (see e.g. Laspita, Breugst, Heblich, & Patzelt, 2012; Wyrwich, 2015), but Schölin, Broomé, and Ohlsson (2016) used Swedish register data to look at the family effect on self-employment in Sweden. They found that the influence that family factors have on an individuals choice of company type is strong, especially for entrepreneurs of limited liability companies. They also reference seven other papers on labour
market outcomes and summarizes their results as the probability that a child ends up as self-employed approximately doubles by having a self-employed parent. Statistics Sweden (SCB, 2013) report that personal connections and recommendations by family and friends constitute the most common recruitment channel in Sweden, regardless of employment type, which further motivates our focus on family effects.

2 MODEL AND DATA

To examine the effect of family experience on young people’s transition probability into temporary agency work we use register-based data, compiled from a number of Swedish register databases – primarily LOUISE/LISA (Longitudinal Database on Education, Income and Occupation) and RAMS (Labour Statistics Based on Administrative Sources). Our dichotomous dependent variable is being employed in the TWA sector in 2007 and by controlling for a set of background characteristics, such as immigration status, education, etc., we will estimate the transition probability for all young adult workers (aged 18-34) in Sweden. Our full dataset contains over 1.5 million individuals, of which 3.1% worked in the TWA sector, and while the number of observations is large, the large proportion of zeroes must be taken into consideration when choosing a regression model.

We therefore estimate a bias-reduced logistic model, as proposed by Firth (1993) and Kosmidis and Firth (2009). While some contend that this is not necessary for large datasets, it is known that the standard logistic model is biased for finite sample sizes; while Schaefer and Richardson (1985) argue that bias correction would be insignificant with a sample size above 200 observations, and Nemes, Jonasson, Genell, and Steineck (2009) show that the bias decreases in sample size, the results of Monte Carlo simulations in King and Zeng (2001b), however, suggest that the finite sample bias, amplified by what in their paper is called rare events (below 5% or so) may warrant some concern.

In order to see if the a standard logistic regression would have had problems in our case, due to being biased, we have compared its results to that of the adjusted-score approach and there were only marginal differences. We, however, choose to present the bias-reduced regression results, since its estimator is second-order unbiased and by construction has a smaller variance than the maximum likelihood estimator (cf. Kosmidis & Firth, 2009).67

6 The data is collected during a week in November which could potentially underestimate the total number of workers in the sector during any given year due to large seasonal variations and the high turnover rate in the temporary agency sector. The appendix contains detailed information and descriptive statistics on all variables, as well as some cross tables.

7 We have also run maximum penalized likelihood estimation, where penalization is by Jeffreys invariant prior (see Firth, 1992), and the results were identical to those of the adjusted-score approach adopted.

Another possible model choice would be to use multi-level analysis, which was used in Schölin et al. (2016) to look at the family experience as a predictor for self-employment. That approach would
Our controls include a set of dichotomous variables that capture the work experience of parents, siblings and partners (where applicable). Either parent or sibling is regarded as having experience of the Swedish temporary agency sector if they have been registered as working in the sector during any year from 2000 to 2007. The data on the current partner is more limited and the effect is instead estimated on the experience during 2004 to 2007.

The country of birth classification is the most disaggregate that the data allows, and corresponds largely to that used by Joona and Wadensjö (2008). Second-generation immigrants are defined as individuals born in Sweden, but with at least one parent having a different country of birth. More than two years of upper secondary school is chosen as the reference level for education as it corresponds to the norm in Sweden (even though school is only compulsory through primary school).

TWA workers are identified through the Swedish Standard Industrial Classification (SNI02) if being specifically involved in either labour recruitment or the provision of personnel activities. The former are workers employed directly at the TWA, whereas the latter are workers being leased to external client companies. While there could potentially be some differences in the two group’s experience of the TWA sector, aggregating them allows for a better comparison with earlier studies (Joona & Wadensjö, 2008, utilise the SNI92 classification which does not distinguish between these groups).

To be identified as a student, the individual must have obtained financial study assistance (grants or loans) from the Swedish National Board of Student Aid (CSN), while also not being registered as a gainfully employed worker. An individual may otherwise have received financial study assistance for the first part of the year but transitioned into primarily gainful employment during the late part of the second half when the data was collected.

A similar set of background variables have been found by Joona and Wadensjö (2008) to influence the probability of being employed in the Swedish temporary agency sector in 1999.

### 3 Results

When presenting our results, we use the odds ratio (O.R.) instead of the regression coefficients, as O.R. is simply a transformation into over the relative probability of TWA employment.8

While controlling for a number of characteristics found to be influential by Joona and Wadensjö (2008), we may conclude from the results in the top of Table 1 that there is indeed a significant effect from the work experience of family members and partners.

8The standard errors are thus odds-ratio adjusted using the Delta method (cf. Greene, 2008).
### Table 1

Odds ratios (O.R.) of being employed in the temporary agency sector

(*Full sample, n = 1,532,879*)

<table>
<thead>
<tr>
<th></th>
<th>O.R.</th>
<th>S.E.</th>
<th>Sig.</th>
<th>L.O.</th>
<th>H.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Family TWA Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>1.802</td>
<td>0.055</td>
<td>***</td>
<td>1.695</td>
<td>1.909</td>
</tr>
<tr>
<td>Father</td>
<td>1.867</td>
<td>0.060</td>
<td>***</td>
<td>1.749</td>
<td>1.984</td>
</tr>
<tr>
<td>Sibling</td>
<td>1.972</td>
<td>0.031</td>
<td>***</td>
<td>1.910</td>
<td>2.033</td>
</tr>
<tr>
<td>Partner</td>
<td>2.749</td>
<td>0.129</td>
<td>***</td>
<td>2.497</td>
<td>3.000</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-20 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21-22 years</td>
<td>1.168</td>
<td>0.019</td>
<td>***</td>
<td>1.131</td>
<td>1.205</td>
</tr>
<tr>
<td>23-24 years</td>
<td>1.004</td>
<td>0.017</td>
<td>0.801</td>
<td>0.970</td>
<td>1.038</td>
</tr>
<tr>
<td>25-26 years</td>
<td>0.881</td>
<td>0.016</td>
<td>***</td>
<td>0.849</td>
<td>0.913</td>
</tr>
<tr>
<td>27-28 years</td>
<td>0.722</td>
<td>0.014</td>
<td>***</td>
<td>0.693</td>
<td>0.750</td>
</tr>
<tr>
<td>29-30 years</td>
<td>0.576</td>
<td>0.013</td>
<td>***</td>
<td>0.551</td>
<td>0.600</td>
</tr>
<tr>
<td>31-32 years</td>
<td>0.473</td>
<td>0.011</td>
<td>***</td>
<td>0.450</td>
<td>0.495</td>
</tr>
<tr>
<td>33-34 years</td>
<td>0.403</td>
<td>0.010</td>
<td>***</td>
<td>0.384</td>
<td>0.423</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school (&lt; 9 years)</td>
<td>1.094</td>
<td>0.082</td>
<td>0.228</td>
<td>0.934</td>
<td>1.254</td>
</tr>
<tr>
<td>Primary school (9-10 years)</td>
<td>0.722</td>
<td>0.012</td>
<td>***</td>
<td>0.698</td>
<td>0.746</td>
</tr>
<tr>
<td>Upper secondary (&lt; 2 years)</td>
<td>1.191</td>
<td>0.022</td>
<td>***</td>
<td>1.147</td>
<td>1.235</td>
</tr>
<tr>
<td>Upper secondary (&gt; 2 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher education (&lt; 3 years)</td>
<td>1.531</td>
<td>0.021</td>
<td>***</td>
<td>1.489</td>
<td>1.572</td>
</tr>
<tr>
<td>Higher education (&gt; 3 years)</td>
<td>1.377</td>
<td>0.020</td>
<td>***</td>
<td>1.338</td>
<td>1.417</td>
</tr>
<tr>
<td>Postgraduate education</td>
<td>0.390</td>
<td>0.072</td>
<td>***</td>
<td>0.248</td>
<td>0.531</td>
</tr>
<tr>
<td><strong>Country of Birth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Nordic country</td>
<td>1.565</td>
<td>0.096</td>
<td>***</td>
<td>1.376</td>
<td>1.774</td>
</tr>
<tr>
<td>Central Europe</td>
<td>1.634</td>
<td>0.126</td>
<td>***</td>
<td>1.386</td>
<td>1.881</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>1.552</td>
<td>0.112</td>
<td>***</td>
<td>1.332</td>
<td>1.772</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>2.142</td>
<td>0.052</td>
<td>***</td>
<td>2.040</td>
<td>2.243</td>
</tr>
<tr>
<td>Former Soviet Union</td>
<td>2.029</td>
<td>0.156</td>
<td>***</td>
<td>1.723</td>
<td>2.334</td>
</tr>
<tr>
<td>USA, Canada, Australia, New Zealand</td>
<td>1.262</td>
<td>0.180</td>
<td>0.104</td>
<td>0.908</td>
<td>1.615</td>
</tr>
<tr>
<td>Other North-, Central- or South America</td>
<td>1.984</td>
<td>0.085</td>
<td>***</td>
<td>1.817</td>
<td>2.151</td>
</tr>
<tr>
<td>North Africa or Middle East</td>
<td>2.337</td>
<td>0.056</td>
<td>***</td>
<td>2.227</td>
<td>2.448</td>
</tr>
<tr>
<td>Other Africa</td>
<td>2.971</td>
<td>0.127</td>
<td>***</td>
<td>2.722</td>
<td>3.220</td>
</tr>
<tr>
<td>Other Asian countries or other Oceania</td>
<td>1.588</td>
<td>0.070</td>
<td>***</td>
<td>1.451</td>
<td>1.724</td>
</tr>
<tr>
<td><strong>Other Attributes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second-generation immigrant</td>
<td>1.627</td>
<td>0.030</td>
<td>***</td>
<td>1.568</td>
<td>1.686</td>
</tr>
<tr>
<td>Metropolitan municipality</td>
<td>1.200</td>
<td>0.012</td>
<td>***</td>
<td>1.177</td>
<td>1.222</td>
</tr>
<tr>
<td>Student</td>
<td>0.963</td>
<td>0.013</td>
<td>***</td>
<td>0.937</td>
<td>0.988</td>
</tr>
<tr>
<td>Children (at least one child)</td>
<td>0.762</td>
<td>0.012</td>
<td>***</td>
<td>0.740</td>
<td>0.785</td>
</tr>
<tr>
<td>Female</td>
<td>0.657</td>
<td>0.006</td>
<td>***</td>
<td>0.644</td>
<td>0.670</td>
</tr>
</tbody>
</table>

*** = sig. < 0.01, ** = sig. < 0.05, * = sig. < 0.1
The results show that previous experience of temporary agency sector employment in any of these peer-groups greatly increases the probability of the young adult also being employed in the sector. The most prominent effect in this category comes from the experience of the partner – suggesting that an individual is between two and a half to three times as likely to work in the sector if his or her partner has experience from the temporary agency sector.

As stated in the introductory section of this paper, and as we will see for some of the personal characteristics that we control for, there could potentially be several important differences between young adult workers who are gainfully employed and the group defined as students. The results for the family experience variables in Table 1 are nevertheless among the variables that show very similar effects for the two groups (Table 2). Another interesting aspect is that the effects of these variables are strikingly large compared to most other control variables, and only some effects relating to the immigration background are of a similar magnitude.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Gainfully employed (n = 1 236 610)</th>
<th>Student (n = 227 453)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% C.I.</td>
<td>95% C.I.</td>
</tr>
<tr>
<td></td>
<td>O.R.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Mother</td>
<td>1.827</td>
<td>0.064***</td>
</tr>
<tr>
<td>Father</td>
<td>1.922</td>
<td>0.072***</td>
</tr>
<tr>
<td>Sibling</td>
<td>1.952</td>
<td>0.037***</td>
</tr>
<tr>
<td>Partner</td>
<td>2.871</td>
<td>0.146***</td>
</tr>
</tbody>
</table>

*) All applicable variables in Table 1 were also included when estimating the above effects.

The results for the age-group of the individual (Table 1) show a small hump shape in the relative propensity to work in the TWA sector; the odds ratio for the lower age cohorts initially increases and thereafter diminishes steadily in the higher ranges. This indirectly justifies our choice to focus our study on the cohort of young adults and corroborates the findings from the Swedish temporary agency sector in 1999 (Joona & Wadensjö, 2008). However, the Andersson-Joona & Wadensjö study uses a different definition of the temporary agency sector, combined with another reference age cohort (41-45 years) that is not available in our sample. This makes any exact comparison difficult and we may only establish that almost a decade later the sector is still seemingly constituted primarily of relatively young workers.

---

9For cross tables see Appendix A.3.
Disaggregating the age effects into gainfully employed and students (Table 3) shows that the hump-shaped effect is attributable to the student group; showing two peaks at 21-22 years and 25-26 years, and then diminishing until there is no statistical difference for the older age cohorts. Gainfully employed workers on the other hand show a steadily diminishing odds ratio from the reference level (corresponding to one) down to just over thirty percent.

<table>
<thead>
<tr>
<th>AGE*</th>
<th>GAINFULLY EMPLOYED</th>
<th>STUDENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O.R.</td>
<td>S.E.</td>
</tr>
<tr>
<td>18-20 years</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>21-22 years</td>
<td>1.014</td>
<td>0.019</td>
</tr>
<tr>
<td>23-24 years</td>
<td>0.835</td>
<td>0.017</td>
</tr>
<tr>
<td>25-26 years</td>
<td>0.686</td>
<td>0.015</td>
</tr>
<tr>
<td>27-28 years</td>
<td>0.569</td>
<td>0.013</td>
</tr>
<tr>
<td>29-30 years</td>
<td>0.460</td>
<td>0.011</td>
</tr>
<tr>
<td>31-32 years</td>
<td>0.385</td>
<td>0.010</td>
</tr>
<tr>
<td>33-34 years</td>
<td>0.329</td>
<td>0.009</td>
</tr>
</tbody>
</table>

*) All applicable variables in Table 1 were also included when estimating the above effects.

This pattern arguably reflects the temporary agency sectors intermediary characteristics as the workers seemingly move on to employment in the regular sector; implying that the worker indeed prefers working in the regular sector. However, working in the Swedish temporary agency sector does not necessarily increase the probability of getting a job in the regular sector compared to if the individual had remained unemployed (Andersson & Wadensjo, 2004; Hveem, 2013). The results in Table 1 regarding highest education level attained show that young adult workers with a relatively high education (with the exception of those with a postgraduate degree) are over-represented in the temporary agency sector. Table 4 similarly show that even the education level among TWA workers in the student group is relatively high – even though they might not yet have attained their final education level. Flexible working hours could potentially be a coveted feature for this group rather than entailing a high degree of stress-inducing precariousness. However, the results also show that there is an even stronger over-representation of workers with a relatively high education level (except for those with a postgraduate degree) among the gainfully employed. The individuals of student with postgraduate education is potentially a statistical anomaly as a consequence of there being very few with that characteristic within the sample.
students with a postgraduate degree in the TWA sector is too small for us to draw any conclusions, as it has lead to a very wide O.R. range in the 95% confidence interval.

**Table 4**

**Education**

<table>
<thead>
<tr>
<th></th>
<th>Gainfully employed</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O.R.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Primary school (&lt; 9 years)</td>
<td>1.108</td>
<td>0.114</td>
</tr>
<tr>
<td>Primary school (9-10 years)</td>
<td>0.869</td>
<td>0.018</td>
</tr>
<tr>
<td>Upper secondary (&lt; 2 years)</td>
<td>1.186</td>
<td>0.026</td>
</tr>
<tr>
<td>Upper secondary (&gt; 2 years)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Higher education (&lt; 3 years)</td>
<td>0.460</td>
<td>0.011</td>
</tr>
<tr>
<td>Higher education (&gt; 3 years)</td>
<td>0.385</td>
<td>0.010</td>
</tr>
<tr>
<td>Postgraduate education</td>
<td>0.329</td>
<td>0.009</td>
</tr>
</tbody>
</table>

*) All applicable variables in Table 1 were also included when estimating the above effects.

The results in Table 1 on the country of birth show that there is a relative over-representation of workers born outside Sweden in the temporary agency sector – with a weak exception for those in the group consisting of USA, Canada, Australia and New Zealand. There are, however, such a small number of TWA workers from USA, Canada, Australia and New Zealand, that we cannot draw any conclusions for that group.

The group with the highest relative odds ratio comes from the sub-Saharan parts of Africa, followed by workers from countries in North Africa, Eastern Europe, the former Soviet Union, and Latin American countries, respectively. There are also relatively many TWA workers who come from other European countries. The relative over-representation of workers born outside of Sweden is also consistent for both the student group and the gainfully employed, as shown in Table 5. The disaggregated results show mostly minor variations, except for workers from the former Soviet Union and Eastern Europe.

11Joona and Wadensjö (2008) find that there is an over-representation of workers from North America while there is an under-representation of similar proportion of workers from Oceania (which in large is constituted by Australia and New Zealand). Unfortunately, the data in this study does not allow for any additional disaggregation.

12Guyana, Surinam and Jamaica are not Latin American countries, but are included in this subgroup due to their geographical proximity to the Latin American countries.
Table 5

COUNTRY OF BIRTH*

<table>
<thead>
<tr>
<th>Country of Birth</th>
<th>Gainfully employed</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O.R.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Sweden</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Other Nordic country</td>
<td>1.478</td>
<td>0.109</td>
</tr>
<tr>
<td>Central Europe</td>
<td>1.522</td>
<td>0.15</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>1.507</td>
<td>0.133</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>1.986</td>
<td>0.059</td>
</tr>
<tr>
<td>Former Soviet Union</td>
<td>2.224</td>
<td>0.218</td>
</tr>
<tr>
<td>USA, Canada, Australia, New Zealand</td>
<td>1.320</td>
<td>0.234</td>
</tr>
<tr>
<td>Other North-, Central- or South America</td>
<td>1.987</td>
<td>0.106</td>
</tr>
<tr>
<td>North Africa or Middle East</td>
<td>2.270</td>
<td>0.069</td>
</tr>
<tr>
<td>Other Africa</td>
<td>2.787</td>
<td>0.156</td>
</tr>
<tr>
<td>Other Asian countries or other Oceania</td>
<td>1.457</td>
<td>0.085</td>
</tr>
</tbody>
</table>

*) All applicable variables in Table 1 were also included when estimating the above effects.

Table 1 shows that there is an over-representation of second-generation immigrants, and that temporary agency workers are utilised to a larger extent within the municipalities of the three largest cities in Sweden. This could possibly be explained by a higher cost of living (supply side) and that new ventures and businesses are both concentrated to, and created primarily in, the metropolitan municipalities – which could arguably create a higher demand for (TWA) workers. Contrary to the situation for all age groups in 1999 (Joona & Wadensjö, 2008), the young adult cohort in 2007 is shown to predominantly consist of men.

Table 6

OTHER ATTRIBUTES*

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Gainfully employed</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O.R.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Second-generation immigrant</td>
<td>1.497</td>
<td>0.034</td>
</tr>
<tr>
<td>Metropolitan municipality</td>
<td>1.155</td>
<td>0.013</td>
</tr>
<tr>
<td>Children (at least one child)</td>
<td>0.762</td>
<td>0.013</td>
</tr>
<tr>
<td>Female</td>
<td>0.704</td>
<td>0.008</td>
</tr>
</tbody>
</table>

*) All applicable variables in Table 1 were also included when estimating the above effects.
The odds ratios for the remaining variables (Table 6), establish that there are some differences between the gainfully employed and the student group, even though the overall effects follow the same pattern with regard to over and under-representation as in the full sample. In the student-sample, there is both a relatively larger representation of second-generation immigrants and workers in metropolitan municipalities, but also an even more accentuated under-representation of women compared to workers classified as gainfully employed.

4 Conclusion

In a survey of the literature, Håkansson et al. (2013) establish characteristics of working in the temporary agency sector that impact the psychosocial and physical work environment of the worker in several negative ways, and these include many of the stressors associated with a precarious job situation. We argue that this could potentially impact the propensity to (actively) recommend this type of employment to family, friends and peers – or even make individuals opt out of working in the temporary agency sector after having (passively) experienced their peers working under these conditions.

The results nevertheless show that there is a significant positive effect from the previous work experience of temporary agency work for each included peer group. These effects are also among the most influential of all explanatory variables in the model in determining the relative probability of an individual working in the temporary agency sector. In addition, the sizes of the effects are more or less equal for gainfully employed workers and for the individuals in the student group. This could potentially corroborate that family members and partners constitute an important recruitment channel, regardless of the many negative aspects associated with temporary agency work. However, the established correlation between the labour market experience of the included peer-groups does not automatically imply causality and supplementary future research is therefore required following these initial findings.

The results for gainfully employed workers and for the student group show that there are some other important differences that have not been captured by previous studies. For instance, there are relatively many temporary agency workers in some of the lower age cohorts in the student sample, whereas the gainfully employed show an almost linear decay in the relative probability of being employed in the agency sector as they grow older. A noteworthy result that is very similar in all samples, but quite different to the findings in 1999 by Joona and Wadensjö (2008), is the relatively high education level among the younger cohorts of the temporary agency sector (cf. also Andersson-Joona & Wadensjö, 2012; Petersson, 2013).

The overall results of this study further establish that there is still an over-representation of individuals with an immigrant background, but also that there is a predomina-
ance of men who are employed in the younger cohorts of the temporary agency sector.

5 Discussion

The results presented in this paper show that having parents, siblings, or partners who have worked in the temporary agency sector will greatly increase the probability that the observed individual will also work in that sector. The norm in Sweden is that all employment contracts are assumed to be open-ended (if the parties have not explicitly stated otherwise), which may not be the case in all countries. Sweden also has a seniority principle where the total length of employment in a firm is directly related to the level of employment protection. Such a principle is given as an example by Sverke et al. (2002) of a social and cultural norm that affects the perception of working life. Immigration status could thus potentially influence an individual's expectation and experience of flexible (or even precarious) employment.

Loughlin and Barling (2001) suggest that contemporary young workers do not consider work to be an investment in their future with the particular company they are currently working for, but rather that they seek immediate pay-off. This could constitute a rationale for choosing temporary agency employment if the individual was compensated for any negative aspects. However, the temporary agency worker seemingly does not only receive a wage penalty (rather than compensation) for the reviewed associated negative physical and psychosocial aspects, but temporary agency employment in Sweden has even been found to prolong the time until the worker gets a regular job rather than being a stepping-stone (Hveem, 2013).

Another plausible explanation for our results is that unemployment is considered to represent even lower status and opportunities than any type of employment – regardless of its content (Sverke et al., 2004). Workers thus utilise any and all recruitment possibilities, where family and close personal networks represent an important channel, to obtain any job rather than holding out for a specific job with certain characteristics. That individuals might be accepting rather than actively choosing to work within the sector is a recurring theme in the literature (see e.g. Andersson & Wadensjo, 2004; Bernasek & Kinnear, 1999; Forde & Slater, 2005; Hveem, 2013; E. J. Jahn, 2008).

This theory has some support in that in the temporary agency sector workforce show an over-representation of groups that traditionally have had a weaker position on the labour market; e.g. young people and immigrants. It could also help explain the findings of Barling et al. (1998), Gregg et al. (2012), and Kind and Haiksen-DeNew (2012), that the children of parents who have experienced job precariousness run a higher risk of earning lower wages, or even experiencing precarious job situations themselves. The relatively high education level in the temporary agency sector among both gainfully employed workers and the student group is an interesting result that is
not only seemingly different from the situation in 1999 (Joona & Wadensjö, 2008), but also somewhat challenges the argument of accepting rather than choosing temporary agency employment. An increased education level should allow for more opportunities and outside options (cf. Güell & Petrongolo, 2007) for the temporary agency worker. This could in turn also allow for a higher transition rate into regular employment if the CF would primarily utilise the TWA in order to screen potential workers. However, for this to hold true it is arguably also important, and perhaps even a prerequisite, that the temporary agency worker is used for such tasks that he/she may accumulate firm-specific human capital (Forde & Slater, 2005). It is nevertheless interesting to note that the education level seems to have increased during the same time that Andersson-Joona and Wadensjö (2012) finds that the temporary agency sectors negative wage difference, compared to the regular sector, has grown.

Having to accept (rather than choosing) temporary agency employment, even though having obtained a relatively high education, may further augment the difference between expectation and outcome. Indeed, Loughlin and Barling (2001) caution that many young adults with an education that required them to think for themselves and who are anticipating motivating work may find themselves at a loss given the increased use of TWAs by the hiring firms. Evidence of this is found by de Graaf-Zijl (2012) who finds that temporary agency workers with the highest education also experience the largest negative difference in job satisfaction compared to similar workers on regular contracts. Walter (2012) suggests that the increased education level could be a result of the TWAs needs to continuously market themselves as providers of the most skilled labour. Westéus and Raattamaa (2014) on the other hand suggests that the increased education level could be a client firm response to the recruitment behaviour of the TWA in the presence of misaligned incentives.

Parents could arguably be more inclined to accept temporary agency employment (rather than to remain unemployed) in order to provide for their offspring. However, both Table 1 and Table 6 suggest that this is not the case. If the relatively few individuals with children in the temporary agency sector is a result of a conscious decision to abstain from having children until the job situation is more stable (cf. Benito, 2006), then the growth of the sector (and atypical employment in general) could have negative effects on society as a whole – and not only for the individual\textsuperscript{13}.

Through a quantitative approach this study has shown that the previous labour market experience of certain close family peer-groups appears to impact the labour market outcome of the young adult. This concluding discussion has debated some of the results and their conformity with previous findings, and put forward some plausible explanatory theories. However, the specific reasons and underlying mechanics of why (and if) family members explicitly recommend this type of

\textsuperscript{13}The data unfortunately does not allow us to estimate any effect regarding the size of the family since the data for the children-variable is only categorically coded for the presence of children or not (and not their actual number), and this possibility is therefore left for future research.
employment despite the surveyed negative characteristics of the sector is left for future research.

Another important question is whether individuals in general (and young adults in particular) choose or accept temporary agency work – especially in the light of an increased relative education level, the sectors adverse working conditions and decreasing relative wages. It would also be interesting to make comparative studies with other countries; e.g. Spain (low probability to transition into regular employment; (low probability to transition into regular employment; Amuedo-Dorantes, Malo, & Muñoz-Bullón, 2008), or Denmark (high transition probability; E. J. Jahn & Rosholm, 2014).

References


doi: 10.1177/0143831X03024002002
doi: 10.1353/mpq.0.0030
doi: 10.1002/j.2161-0045.2009.tb00171.x
doi: 10.1016/j.ssci.2007.05.004
doi: 10.1007/978-3-662-26811-7
doi: 10.1111/j.1475-5890.2012.00160.x


Loughlin, C., & Barling, J. (2001). Young Workers’ Work Values, Attitudes, and


Spermann, A. (2011). *The new role of temporary agency work in Germany* (Discussion
APPENDIX

A.1 CLASSIFICATION OF VARIABLES

[Dependent]

Binary response variable (dummy) on whether or not the individual is working in the Swedish agency sector in 2007 (according to SNI02). Missing values (incl. individuals that are either unemployed or outside of the workforce) are excluded.

Mother/Father/Sibling

Binary response variables on whether or not any of these relatives are recorded as having worked in the Swedish agency sector during some point in time between 2000-2007. Missing values (e.g. omitted information, working outside of Sweden, being unemployed, or not being in the workforce) are regarded as indicating that the individual did not have agency employment experience for that year.

Partner

Binary response variable on whether or not the individual has a partner recorded as working in the agency sector at any point in time from 2004 to 2007 (due to limited data availability). Missing values are classified as that the individual does not have a partner with prior experience.

AGE GROUP-category

The age group to which the individual belongs.

EDUCATION-category

The highest level of education attained at the time of measurement.

COUNTRY OF BIRTH-category

The recorded country of birth of the individual, aggregated into a standard classification.

Second-generation immigrant

Binary response variable. The individual is regarded as a second-generation immigrant if the individual is born in Sweden, and has at least one parent who was born outside of Sweden.
Metropolitan municipality

Binary response variable indicating whether the individual is registered in a municipality belonging to one of Sweden’s three largest cities (Stockholm, Gothenburg and Malmö).

Student

Binary response variable. The worker is defined as a student if he/she obtained financial study aid in 2007, and is not registered as being gainfully employed.

Children

Binary response variable on whether the individual is recorded as having at least one child.

Female

Binary response variable on the gender of the individual.
### A.2 Descriptive Statistics

<table>
<thead>
<tr>
<th>Table A</th>
<th>Descriptive statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>FULL SAMPLE</strong></td>
</tr>
<tr>
<td></td>
<td>No.</td>
</tr>
<tr>
<td><strong>1 532 879</strong></td>
<td><strong>47 450</strong></td>
</tr>
<tr>
<td><strong>FAMILY INFLUENCE</strong></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>20 877</td>
</tr>
<tr>
<td>Father</td>
<td>17 243</td>
</tr>
<tr>
<td>Sibling</td>
<td>73 609</td>
</tr>
<tr>
<td>Partner</td>
<td>8 892</td>
</tr>
<tr>
<td><strong>AGE</strong></td>
<td></td>
</tr>
<tr>
<td>18-20 years</td>
<td>274 024</td>
</tr>
<tr>
<td>21-22 years</td>
<td>184 529</td>
</tr>
<tr>
<td>23-24 years</td>
<td>172 401</td>
</tr>
<tr>
<td>25-26 years</td>
<td>172 094</td>
</tr>
<tr>
<td>27-28 years</td>
<td>176 671</td>
</tr>
<tr>
<td>29-30 years</td>
<td>172 359</td>
</tr>
<tr>
<td>31-32 years</td>
<td>182 924</td>
</tr>
<tr>
<td>33-34 years</td>
<td>197 877</td>
</tr>
<tr>
<td><strong>EDUCATION</strong></td>
<td></td>
</tr>
<tr>
<td>Primary school (&lt; 9 years)</td>
<td>4 045</td>
</tr>
<tr>
<td>Primary school (9-10 years)</td>
<td>203 920</td>
</tr>
<tr>
<td>Upper secondary (≥ 2 years)</td>
<td>124 637</td>
</tr>
<tr>
<td>Higher education (≥ 3 years)</td>
<td>215 708</td>
</tr>
<tr>
<td>Postgraduate education</td>
<td>4 578</td>
</tr>
<tr>
<td><strong>COUNTRY OF BIRTH</strong></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>1 420 765</td>
</tr>
<tr>
<td>Other Nordic country</td>
<td>7 621</td>
</tr>
<tr>
<td>Central Europe</td>
<td>3 925</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>5 157</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>31 881</td>
</tr>
<tr>
<td>Former Soviet Union</td>
<td>2 977</td>
</tr>
<tr>
<td>USA, Canada, Australia, New Zealand</td>
<td>1 371</td>
</tr>
<tr>
<td>Other North-, Central- or South America</td>
<td>11 042</td>
</tr>
<tr>
<td>North Africa or Middle East</td>
<td>28 360</td>
</tr>
<tr>
<td>Other Africa</td>
<td>7 216</td>
</tr>
<tr>
<td>Other Asian countries or other Oceania</td>
<td>12 564</td>
</tr>
<tr>
<td><strong>OTHER ATTRIBUTES</strong></td>
<td></td>
</tr>
<tr>
<td>Second-generation immigrant</td>
<td>71 488</td>
</tr>
<tr>
<td>Metropolitan municipality</td>
<td>644 494</td>
</tr>
<tr>
<td>Student</td>
<td>227 453</td>
</tr>
<tr>
<td>Children (at least one child)</td>
<td>407 792</td>
</tr>
<tr>
<td>Female</td>
<td>744 535</td>
</tr>
</tbody>
</table>
### A.3 Cross Tables

#### A.3-1 Gainfully Employed and Students

<table>
<thead>
<tr>
<th>TWA</th>
<th>Stud.</th>
<th>Emp. 0</th>
<th>Emp. 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>64,928</td>
<td>1,201,720</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>218,781</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>3,888</td>
<td>34,890</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>8,672</td>
<td>0</td>
</tr>
</tbody>
</table>

Emp. Gainfully employed
Stud. Student
TWA Working in the TWA sector

#### A.3-2 Family

<table>
<thead>
<tr>
<th>Partner</th>
<th>Mother</th>
<th>Father</th>
<th>Sibling</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mother</th>
<th>Father</th>
<th>Sibling</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
List of dissertations at the Department of Economics, Umeå University

Avhandlingar framlagda vid Institutionen för nationalekonomi, Umeå universitet

Holmström, Leif (1972) Teorin för företagens lokaliseringsval. UES 1. PhLic thesis
Stage, Jörn (1973) Verklighetsuppfattning och ekonomisk teori. UES 4. PhLic thesis
Löfgren, Curt (1998) Time to Study Students: Two Essays on Student Achievement and Study Effort. UES 466. PhLic thesis
Berglund, Elisabet (1999) Regional Entry and Exit of Firms. UES 506. PhD thesis
Sörensson, Robert (2010) Marshallian Sources of Growth and Interdependent Location of Swedish Firms and Households. UES 815. PhD thesis
Heidrich, Stefanie (2016) Essays on Intergenerational Income Mobility, Geographical Mobility, and Education. UES 932. PhD Thesis